Lecture Notes in Economics
and Mathematical Systems 602

Norman Ehrentreich

Agent-Based Modeling

The Santa Fe Institute Artificial Stock Market Model Revisited

Springer

Dr. Norman Ehrentreich
RiverSource Investments, LLC
262 Ameriprise Financial Center
Minneapolis, MN 55474
USA
norman.x.ehrentreich@ampf.com

ISBN 978-3-540-73878-7 e-ISBN 978-3-540-73879-4

DOI 10.1007/978-3-540-73878-7

Lecture Notes in Economics and Mathematical Systems ISSN 0075-8442

Library of Congress Control Number: 2007937522

Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover-design: WMX Design GmbH, Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Meinen Eltern gewidmet.

Foreword

When the original Santa Fe Institute (SFI) artificial stock market was created in the early 1990's, the creators realized that it contained many interesting new technologies that had never been tested in economic modeling. The authors kept to a very specific finance message in their papers, but the hope was that others would pick up where these papers left off and put these important issues to the test. Tackling the complexities involved in implementation has held many people back from this, and many parts of the SFI market remain unexplored. Ehrentreich's book is an important and careful study of some of the issues involved in the workings of the SFI stock market.

As Ehrentreich's book points out in its historical perspective, the SFI market was intended as a computational test bed for a market with boundedly rational learning agents replacing the standard setup of perfectly rational equilibrium modeling common in economics and finance. These agents exhibit reasonable, purposeful behavior, but they are not able to completely process every aspect of the world around them. This can be viewed much more as a function of the complexity of the world, rather than the computational limitations of agents. In a financial world out of equilibrium, optimal behavior would require knowledge of strategies being used by all the other agents, an information and computational task which seems well out of reach of any trader. The SFI market's main conclusion was that markets where agents were learning might not converge to traditional simple rational expectations equilibria. They go to some other steady state in which a rich set of trading strategies survives in the trader population. In this steady state the market demonstrates empirical signatures that are present in most financial time series.

This book is an excellent reference to both the learning, and empirical literature in finance. It stresses the difficult empirical facts that are out of reach of most traditional financial models including persistent volatility and trading volume, and technical trading behavior. However, Ehrentreich's main mission is to dig deeply into the SFI market structure to understand what is actually going on. Computational economic models can often be explored at three levels. There is sort of a big picture level where concepts such as rational expectations and bounded rationality are explored. There is also the very low level where researchers discuss the nuts and bolts of different modeling languages. In between these sits a region where many of the computational learning technologies are implemented. This is where technologies such as genetic algorithms, classifier systems, and neural networks drive much of what is going on. This is Ehrentreich's area of exploration, and it is critically important to agent-based modelers since one needs to know the sensitivity of the higher level results to changes in the learning structures used beneath them.

The SFI market uses two learning mechanisms extensively: the genetic algorithm, GA, and classifier system. Both of these are developments of John Holland, one of the SFI market coauthors. The GA is a type of general evolutionary learning mechanism, and it is used in both computer science and economics. Its properties have been studied, but it is still not completely understood. In computer science it is often studied in difficult optimization problems. These are problems with well defined objectives, and are quite different from the more open ended coevolutionary problems in economics where agents are competing with each other. The classifier system is an interesting learning structure that allows agents to dynamically find relevant states in the world around them. For example, actions might be conditioned on whether a stock is currently priced above a certain multiple of dividends. The classifier has the power to endogenously slice up a stream of empirical information into states of the world. Very few learning mechanisms are able to do this. With this generality comes a lot of model complexity, and many implementations of the classifier seem computationally unwieldy. They also involve many implementation questions that need to be explored.

In several chapters Ehrentreich explores some of the more important aspects of the SFI classifier implementations. He shows that the SFI classifier is sensitive to certain design characteristics. Under different assumptions about evolution the classifier system behaves very differently from the original SFI model. Ehrentreich carefully modifies and explores his own operation on mutating trading strategies. Using this modified mutation causes a situation in which the SFI market is much more likely to converge to the rational expectations equilibrium, and the rich technical trading dynamic does not emerge. The results in the original SFI market are clearly sensitive to how mutation is implemented. The book goes on to do a comparative study between mutation operators. A key issue is how many technical trading related rules are evolved, and whether the system is likely to generate lots of technical rules by chance in the evolutionary process. The modified mutation operator does not generate many of these rules, so they never really get a foot in the door of trading activity. The SFI structure facilitates their formation, but it is possible this could be driven more by genetic drift than selection. The original SFI studies never really answered these questions, and it only looked at trading strategy formation in an indirect level by looking at aggregate numbers. This was a clear weakness. Ehrentreich does some careful checks to see if technical rules are adding value at the agent level. It appears that they are, so many of the SFI indirect conclusions are sound.

The dynamics of wealth was never part of the original SFI market. It is an interesting omission that the SFI market never really considered long term wealth in a serious way in its implementation. This is strange since many arguments about efficient markets thrive on the relative dynamics of trader wealth. Ehrentreich concludes that this is a complex problem, and there may be difficulties with some of the other studies that try to tag a wealth dynamic onto the SFI market. In my opinion this is one of the biggest limitations of the actual SFI market structure.

This book is an important piece of work for understanding the dynamics of models with interacting learning agents. I think researchers in the future will find it critical in helping them to understand the dynamics of evolutionary learning models. Most importantly, it sets an important standard for doing careful internal experiments on these markets and the learning mechanisms inside them.

Brandeis University, Waltham, MA
September 2007

Blake LeBaron

Preface

The road of science is filled with surprises. When embarking on a scientific journey, we probably have a specific destination in mind, but we never know whether the road will take us there nor what places we may encounter along the way.

This trip was no exception. Before anyone starts reading this travelogue, I think that I should briefly mention a few places that I visited, but decided to pass over while writing this book. I originally aimed at converting the well-known Santa Fe Institute Artificial Stock Market (SFI-ASM) into a two stock version to study portfolio decisions of individual investors. My early forays into this unknown territory yielded some interim results, but until now they are still waiting to be further examined.

Instead, my road took a sudden and unexpected turn. One of the most important findings of the original SFI-ASM was the emergence of technical trading for faster learning speeds. Yet a thorough analysis of the agent's learning algorithm suggested that this might have been caused by an ill-designed mutation operator. For a couple of years, many tests confirmed this supposition. For instance, even though technical trading rules emerged in the original SFI-ASM, they were rarely acted upon. Most importantly, though, was that agents with an alternative mutation operator discovered the homogeneous rational expectations equilibrium, a result that found immediate approval by neoclassically inclined economists.

I traveled a long way down this road. Since I considered the existence of technical trading to be an empirical fact of financial markets, I tried to unearth the necessary ingredients to reintroduce it into my model. Nothing that I devised, neither social learning nor explicit herding mechanisms, succeeded in that endeavor. There was, however, another surprise waiting behind the supposedly final turn of my journey. One newly designed test showed a slight superiority of technical trading rules in the original model. A side-trip all the way down to population genetics finally proved that my agents were committing a mistake by deciding to ignore technical trading rules. Again, parts of my prior research were discarded, and a new chapter was written explaining why I and previous researchers went wrong in interpreting the simulation results. I hope that this chapter will prove most useful for any research involving genetic algorithms. My prior belief that technical trading was an artificially introduced model artifact had also caused me to visit some previous studies about wealth levels. I was able to show that the SFI-ASM was not designed to address any questions related to wealth. Fortunately, this part was unaffected by the breakdown of the initial motivation to look into the wealth generation process.

A long journey with such detours was certainly not easy. I could not have arrived at the final destination without the tremendous support and encouragement that I have found along the way. Above all, I wish to thank my parents Werner and Ellinor Ehrentreich, for without them, I would not have had the opportunity to embark on this journey. I would also like to thank Reinhart Schmidt for letting me choose my destination and for giving me the freedom to follow my own path. Among the numerous friends, colleagues, and conference participants who have contributed in many ways are Manfred Jäger, Ulrike Neyer, Ralf Peters, Martin Klein, Heinz-Peter Galler, Joseph Felsenstein, Alan Kirman, and James Stodder. Of course, this book would not have been finished without the contributions by Blake LeBaron. Not only did he play a major role in the creation of the model that I set out to extend, then critiqued, and finally confirmed, he also often helped and clarified many questions that I was pondering. Many thanks also go to Lars Schiefner, Doris Storch, and Klaus Renger, especially for their help during the final stages of this project. Last, but not least, I thank Tanya Novak for her patience and help, especially for her proofreading. Nonetheless, I absolve her from all remaining mistakes and typos and credit them to my cats, Zina and Francesco, who stubbornly insisted on their input by jumping on the keyboard.

I now hope that the reader will find it useful to visit the places that I have found worthwhile to mention in this book.

Minneapolis, MN
September 2007 *Norman Ehrentreich*

Contents

Part I

Agent-Based Modeling in Economics

1

Introduction

What I cannot create, I do not understand.
Richard P. Feynman

In addition to deduction and induction, simulation is sometimes seen as a third methodology for doing research. Even though simulation does not prove theorems, it can enhance our understanding of complex phenomena that have been out of reach for deductive theory. Tesfatsion defines *agent-based simulations* as the computational study of economies that are modeled as evolving systems of autonomous interacting agents [421]. In the last decade, they have become a widely accepted tool for studying decentralized markets.

A major advantage is that agent-based models allow the removal of many restrictive assumptions that are required by analytical models for tractability. For instance, all investors could be modeled as heterogeneous with respect to their preferences, endowments, and trading strategies.

Among the numerous agent-based simulations of financial markets [247, 268, 81], the Santa Fe Institute Artificial Stock Market (SFI-ASM) is one of the pioneering models and thus, probably the most well-known and best studied. It was created by a group of economists and computer scientists at the Santa Fe Institute in New Mexico to test whether artificially intelligent agents would converge to the homogeneous rational expectations equilibrium or not. The original SFI-ASM has been described in a series of papers [331, 11, 330, 244].

In agent-based simulations, replication of existing models from scratch is an important, but often neglected step. Axelrod emphasizes that without this outside confirmation, possible erroneous results based on programming errors or misinterpretation of simulation data can go undetected [13]. Skepticism towards computational models is sometimes voiced since their results are seen as counterintuitive and incomprehensible because the computer program remains an impenetrable “black box”. Judd considers the black box criticism to be more a result of poor exposition of computational work and a lack of sufficient sensitivity analyses. Since third party replications need to open that black box widely, that criticism can effectively be addressed [208].

This book starts by presenting the reasons that led to the adoption of the agent-based programming approach in economics. Since one of these reasons, the desire to replace the representative agent approach with heterogeneous agents, leads to a break down of rational expectation formation, chapter 3 contrasts several concepts of agent rationality. When modeling less than fully rational agents, researchers have to equip their agents with learning algorithms which are discussed in chapter 4.

The replication of selected stylized facts of financial markets through agent-based simulations of financial markets is the focus of chapter 5. It starts by introducing the Efficient Market Hypothesis (EMH). Much of our empirical knowledge about financial markets, often summarized as stylized facts, stems from attempts to either prove or disprove the EMH. Several competing market hypotheses that strive to better explain the stylized facts than the EMH are subsequently discussed. In the final part of chapter 5, a selection of agent-based models of financial markets is briefly introduced.

The main part of this book analyzes a particular Java-replication of the SFI-ASM model which was originally programmed in Objective-C. The goal is to assess whether the SFI-ASM's result of emergent technical trading is robust to changes in the model design. To this end, the simulation results are supplemented by theoretical analyses of certain model features. The replication results of the reprogrammed Java-version are presented in chapter 6. A Markov chain analysis of the original mutation operator delivers the motivation to develop an alternative mutation operator in chapter 7. Because of the differences in simulation results, chapter 8 needs to reexamine the model structure with respect to wealth accumulation. The final chapter analyzes and compares the two mutation operators in more detail. By interpreting the insights gained from the theoretical analysis of the two mutation operators, chapter 9 concludes by offering an explanation as to why the validity of the EMH is entirely consistent with the simultaneous existence of technical trading rules.

2

The Rationale for Agent-Based Modeling

Imagine how hard physics would be if electrons could think.
Murray Gell-Man[1]

2.1 Introduction

The ascent of powerful and affordable microcomputers and the availability of huge economic data sets have sparked the development of a rather recent branch in economic research. The rapidly growing field of computational economics is a broad concept and encompasses many different areas. An eclectic and not exhausting overview of all these different activities[2] can be found in the "*Handbook of Computational Economics*" by Amman et al. [8]. Since a precise definition of computational economics is not being offered by its editors, Riechmann goes as far as to suggest that every economist who uses a computer for more than mere typewriting engages in computational economics [356].

But if we emphasize the aspect of computability of economic problems, i.e., problems that allow for numerical results, the roots of computational economics could easily be dated back before computers were actually used by economists.[3] Several contributions in this handbook

[1] Attributed to Gell-Mann, 1969 Nobel laureate in physics and co-founder of the Santa Fe Institute, cited by Page [329].

[2] Among those are, for instance, the numerical computation of Nash equilibria, deterministic and stochastic simulations, or numerical dynamic programming problems.

[3] Nagurney [316], as well as Kendrick [217], mention the early contributions to computational economics in the 1950ies by Koopmans, Samuelson, or Solow. In finance, the work by Markowitz about efficient portfolio diversification needs to be mentioned at this point [288, 289].

focus on algorithms and numerical methods for finding Nash-equilibria or the solutions to dynamic nonlinear systems of equations, yet some of them were developed even before computers were. The access to computer technology just reduced the sometimes prohibitive computational costs of these algorithms and allowed them to be used for practical purposes.

Within computational economics, the field of agent-based modeling or simulation (ABM, ABS), sometimes also called microscopic simulation (MS, Levy et al. [248]) or agent-based computational economics (ACE, Tesfatsion [420]), seems currently to be the most rapidly growing discipline. This is acknowledged, for instance, by the appearance of a second volume of the "*Handbook of Computational Economics*" which will be solely dedicated to this approach [209].

In the ABS approach, model economies are built from the bottom up, i.e., they consist of many autonomous and interacting agents. It had its first breakthrough with the influential models by Schelling [375, 376] about endogenous neighborhood segregation.[4] These models are populated by two types of agents who only care about the composition of their own small neighborhood. In particular, they do not tolerate more than a certain fraction of agents of the other type in their vicinity, however, they do not care about integration or segregation on the city level. Unsatisfied agents are allowed to move to a neighborhood that they are happy with. Schelling showed that for a wide range of the agents' neighborhood tolerance parameter, an initially integrated city emerges to an almost completely segregated entity.[5] Thus, Schelling demonstrated how stable macrobehavior may emerge from strictly local motives on the agents' level, a macrobehavior that would be hard to predict by exclusively looking at individual motives. Nowadays, this phenomenon is known as emergence, a key concept of the theory of complex adaptive systems.

The agent-based approach is currently considered by many researchers as the latest revolution in economic methodology. However,

---

[4] A comparison of the different versions by Schelling is given by Pancs and Vriend [332]. The most important distinction is the dimensionality of neighborhoods. In his 1969 model, a neighborhood is defined only in one dimension, i.e., agents populate a line, while his later models use two dimensional lattices. A two-dimensional web-based simulation example in NetLogo by Wilensky [436] can be found at http://ccl.northwestern.edu/netlogo/models/Segregation.

[5] Pancs and Vriend [332] extend the framework by asking how individual preferences may alter the outcome. Surprisingly, the results are very robust to changes in preferences. Even if all agents strictly prefer perfect integration, neighborhood segregation will still occur.

in order to understand and properly evaluate this shift in methodology, Bankes points out that one should not focus on the progress in the computer sciences which made this whole development possible, but on the inappropriateness of traditional methods which made it necessary [22].

The deficiencies of analytical modeling can only be sought in its assumptions. Thus, the first part of this chapter discusses some of these and contrasts them with the requirements of agent-based modeling. The main culprits from an agent-based point of view are the widespread assumption of representative agents and that of rational expectations. First, it is obvious that fictitious representative agents in macroeconomic models are incapable of generating emergent phenomena. Financial markets would also be characterized by the absence of any trading activity. Secondly, fully rational and perfectly informed agents have essentially no ability to exercise free choice. Despite some suggestive rhetoric, individuals can follow only one, i.e., the rational course of action [235].

The agent-based modeling approach, on the other hand, requires neither of these two assumptions. The ability to cope with heterogeneous and boundedly rational agents makes it a perfect tool to study decentralized markets. Instead of a reductionist approach, agent-based models treat the economy as an evolving complex adaptive system consisting of many heterogeneous and interacting agents. The development of artificial financial markets has thus become a major application for the agent-based paradigm.[6]

2.2 The Representative Agent Modeling Approach

The notion of representative agents appeared in the economic literature already in the late 19th century. Edgeworth used the term '*representative particular*' [104, p. 109], while Marshall introduced a '*representative firm*' in his *Principles of Economics* [290].[7] However, only after Lucas [261] had published his article about econometric policy evaluation—the famous Lucas-critique—they became the dominant macroeconomic

[6] According to [227], scientific paradigms share two essential characteristics: First, their achievements must have enough novelty to attract a permanent group of scientists away from competing modes of scientific activity. Secondly, their open-endedness must allow for addressing many different kinds of problems. Whether the agent-based modeling approach already is or might become the next paradigm in the economic sciences is left to the reader to be answered.

[7] A discussion of the origins of representative agents can be found in [172].

approach. Today's representative agent models are characterized by an explicitly stated optimization problem of the representative agent, which can be either a consumer or a producer. The derived individual demand or supply curves are then, in turn, used for the corresponding aggregate demand or supply curves.

A series of papers in which the rationale for using representative agent models is convincingly set forth does not exist according to [172]. From the "*large set of introductions, paragraphs, and parenthetical asides*", however, Hartley identifies several motives for their use. They were thought to avoid the Lucas critique, to provide rigorous microfoundations to macroeconomics, and to help build powerful Walrasian general equilibrium models.

2.2.1 Avoiding the Lucas-Critique

Before [261], macroeconomic models were often defined in terms of three vectors: y_t being the set of endogenous variables, x_t the vector of exogenous forcing variables, and ϵ_t the set of random shocks. Fixed parameters are subsumed in a vector θ.

$$y_{t+1} = F\left(y_t, x_t, \theta, \epsilon_t\right). \tag{2.1}$$

Lucas, however, criticized that it is likely that some of the the parameters contained in θ change due to a shift of policy regime λ. Aggregate quantities and prices might react differently than predicted since agents may change their behavior in a way which is not captured in the aggregate equations. For instance, agents could adapt their expectations about future inflation rates or, in the case of rational expectations, change them even before an anticipated policy shift is implemented. An attempt to exploit a potential trade-off between unemployment and the inflation rate through an expansionary monetary policy may thus be foiled. Taking the Lucas-critique into account, equation (2.1) should be rewritten as

$$y_{t+1} = F\left(y_t, x_t, \theta(\lambda), \mu, \epsilon_t\right), \tag{2.2}$$

where $\theta(\lambda)$ contains regime dependent parameters, while μ is thought to consist of truly invariable taste and technology parameters.

While Lucas offered no solution to this fundamental problem, representative agent models were soon to be thought of offering an easy escape from it. Going beyond simple aggregate relationships and analyzing the economy at a deeper level than before, macroeconomists pretended to know the structural equations from which the aggregate supply and demand curves are derived [371]. If a policy change is announced or implemented, the representative agent simply recalculates his optimization problem, given his objective function and budget constraints. This approach also satisfied the desire for a microfoundation in the new classical sense since behavior is derived from a utility maximization problem.[8]

Hartley argues though that it is impossible to identify truly invariable taste and technology parameters [172]. Acknowledging this, economists should realize that the Lucas critique imposes a standard that no macroeconomic model can probably ever fulfil. Since representative agent models suffer from the same deficiencies as old style Keynesian macroeconomic models, the justification for their use is greatly undermined.

2.2.2 Building Walrasian General Equilibrium Models

The desire to build Walrasian general equilibrium models of the economy provides another strong motivation for using representative agents. Not only are economists interested in the existence of equilibria within these types of models, they should furthermore be unique and stable. The Arrow-Debreu framework as the modern embodiment of Walrasian models, however, is far too complex to be solved for millions of heterogeneous consumers and firms. Using a representative agent instead makes it easy to find the competitive equilibrium allocation for a model economy [370].

However, for Walrasian models to be true, their structural assumptions have to be true. False structural assumptions lead to false conclusions.[9] Since real individuals are obviously heterogeneous with respect to their preferences and cognitive abilities, the assumption of a representative agent cannot be considered structural since it is not true. Thus, it must be superficial, i.e., irrelevant to the underlying structure of the economy. When using representative agents instead of heterogeneous individuals in a Walrasian model, the modeler must believe

[8] According to [172], there is another view of what constitutes an appropriate microfoundation. Keynesian models, for instance, backed up their macroeconomic relationships with some explanatory story which is thought to be entirely sufficient.

[9] Hartley contrasts the requirement of true structural assumptions in Walrasian equilibrium models with Friedman's famous dictum that the realism of assumptions is irrelevant [172]. For Friedman, the validity of a theory is based on how well its prediction match reality, no matter how realistic its assumptions.

that the actual world would not look very different from the model if populated by identical clones [172, p. 66].

There is mounting evidence, however, that this is not the case. First of all, representative agent models are usually characterized by a complete absence of trade and exchange in equilibrium [225], one of the most basic activities in a market economy.[10] For instance, in the classical CAPM [389, 252, 313], there is no trade after agents have completed their initial portfolio-diversification.

2.2.3 Representative Agents and the Fallacy of Composition

It has long been known in economics that what is true for individual agents may not hold for the aggregate economy. This phenomenon is called the *fallacy of composition* [61, 172]. Together with its logical counterpart, the fallacy of division, it highlights the tension between micro- and macroeconomics. The economy as a whole is formed by many consumers and firms whose interactions may cause emergent behavior at the macroeconomic level. Correct policy recommendations for individual economic units may not work for the aggregate economy or vice versa. For instance, in times of recession, a profit maximizing firm is likely to lay off workers in order to survive, while a similar action by the government as an aggregate player will aggravate the economic downturn.

Representative agent models usually commit the fallacy of composition by ignoring valid aggregation concerns. Kirman, for instance, provides a graphical example based on [203] in which the representative agent disagrees with all individuals in the economy [225]. Policy recommendations based on the representative agent, a common practice in today's macroeconomics, are illegitimate in this case.

A rigorous treatment of this logical fallacy can be found in the literature on exact aggregation.[11] Gorman [160] was the first who derived general conditions under which the aggregation of individual preferences is possible. He showed that aggregate demand is dependent on income distribution unless all agents have identical homothetic utility functions. Only when their Engel curves are parallel and linear, a redistribution of income will leave the aggregate demand unaffected. Other authors [412, 203, 249] derived similar conditions for exact aggregation,

[10] In the literature, there are as many no-trade theorems [306, 12, 422] as attempts to solve this apparent contradiction with economic reality. These attempts are usually characterized by a relaxation of the assumption of strict homogeneity of market participants [83, 230, 428].

[11] An introduction to the problem of exact aggregation can be found in [161].

the least restrictive ones given in [249]. In any case, those conditions are still so special that no economist would ever consider them to be plausible [225]. In the unlikely event of them being satisfied, no-trade situations would be the result.

For apparent reasons, the literature on exact aggregation has largely been ignored by representative agent modelers. In summarizing this body of research, Kirman concludes that the reduction of a group of heterogeneous agents to a representative agent is not just an analytical convenience, but is "*both unjustified and leads to conclusions which are usually misleading and often wrong*". Hence, "*the 'representative' agent deserves a decent burial, as an approach to economic analysis that is not only primitive, but fundamentally erroneous*" [225, p. 119].

2.2.4 Expectation Formation in Markets with Heterogeneous Investors

In asset pricing models with homogeneous investors, asset prices typically reflect the discounted expected payoffs and follow a martingale, i.e., today's expectation of next period's price $E_t\left[p_{t+1}\right]$ just equals the current price p_t [365]. This martingale property of asset prices implies that the expectations of the representative investor satisfy the *law of iterated expectations*: $E_t\left[E_{t+1}(p_{t+2})\right] = E_t\left[p_{t+2}\right]$, that is, his expectation today of tomorrow's expectation of future payoffs equals his current expectation of these future payoffs [6]. Since all investors are alike, individual and average expectations of future asset prices coincide.

Yet, when giving up the concept of a representative investor, the knowledge of average expectations of future payoffs becomes important for an individual investor. With differential private information and public information, an individual's expectation and the average expectation about future payoffs are likely to diverge. Furthermore, for average expectations, the law of iterated expectations is not satisfied anymore. Thus, according to Keynes [218], real financial markets with heterogeneous agents resemble more a beauty contest in which the competitors have to choose the six prettiest faces from a hundred photographs, the winner being the one whose choice most nearly corresponds to the average preferences of the other competitors. Instead of choosing the face that one considers prettiest, participants devote their intelligence to anticipate "*what average opinion expects the aver-*

age opinion to be" with some participants forming even higher order beliefs about other participants' beliefs.[12]

Asset pricing models with higher order beliefs, however, are prone to an excess reliance on public signals [6]. Instead of always tracking the fundamental value, those models may exhibit bubbles and sunspots, i.e., equilibria that are different from the fundamental equilibrium simply because agents believe that a purely extrinsic random variable, a so-called sunspot[13] variable, has an effect on the equilibrium allocation. Bubbles and self-fulfilling sunspot equilibria constitute a kind of *market created uncertainty* [418] or *behavioral uncertainty* [334], which is caused by the direct or indirect interaction between heterogeneous market participants.

The prevailing hypothesis in expectation formation since the early seventies is that of rational expectations [315]. It assumes that agents have perfect knowledge of the underlying market equilibrium equations and that they use these equations to determine their rational expectations forecast. As a result, the agents' subjective expectations are identical to the objective mathematical expectations given the available information set. Any errors in expectations must be random with a mean of zero. In other words, a consistent bias in expectations is inconsistent with the presumed rationality of the agents.

When forming their expectations of future payoffs, investors with rational expectations do not stop with third or fourth order beliefs about other beliefs. Their perfect deductive logic requires them to do an infinite regress about the expectations of every other investor. Only for homogeneous expectations among investors are these expectations well defined. The expectations of heterogeneous investors, on the other hand, that their expectations remain indeterminate under the Rational Expectations Hypothesis [334, 11]. Thus, heterogeneous investors cannot use deductive logic in forming their expectations about future payoffs. Perfect rationality and rational expectations in particular cannot be well defined in analytical asset pricing models with many heterogeneous investors.

---

12 A formalization of Keynes' beauty contest can be found in [334, p. 276–281]. Further discussions of expectation formation among heterogeneous investors are [425, 33].

13 Instead of "sunspots", the terms "animal spirits" [218] or "self-fulfilling prophecies" are sometimes used in the literature. The term "sunspot" is used to point out the arbitrariness of the variable. A model solution depends on a sunspot variable only because everyone believes in its importance. Reviews on sunspot equilibria and endogenous fluctuations can be found in [29, 14].

2.3 Rational Expectations and Disequilibrium Dynamics

If perceptions of the environment, including perceptions about the behavior of other people, were left unrestricted, then economic models in which peoples' actions depend on their perceptions remain indeterminate and can produce many possible outcomes. Such models are useless as instruments for generating predictions as Sargent [372] pointed out. Because of the requirement of consistency of beliefs, the Rational Expectations Hypothesis has proved fruitful in restricting the number of possible outcomes of economic models. Yet, there are still many rational expectations models with multiple equilibria, e.g. [284, 364], and the problem of indeterminacy remains.

Rational expectations models also implicitly assume that agents already have acquired all relevant information such as the various probability distributions they face. The process of gathering and processing this information to optimally choose their perceptions and actions is intentionally left out from the analysis.[14] Lucas [262] thus emphasizes that the Rational Expectations Hypothesis is just an equilibrium concept and not a behavioral assumption.[15] It omits any description of information perception, information acquisition, and information processing. As an equilibrium concept, it can at best describe how a system might eventually behave if it ever settles down to a situation in which all agents have successfully acquired all necessary information [372]. It can neither describe whether the system would actually arrive at one specific equilibrium nor tell us anything about the out-of-equilibrium dynamics of the system. These questions are out of the scope of rational expectations models and cannot be answered without any additional concepts.

A basic proposition of the Rational Expectation Hypothesis is that expectations are unbiased estimates of the true underlying stochastic process. However, if information acquisition is costly, agents will

---

[14] This concept of rationality in economics is classified by Simon [395] as substantive as opposed to his own concept of procedural rationality. Simon used bounded rationality and procedural rationality as synonyms since the focus is on the reasoning process itself rather than on identifying a goal-maximizing action in a given situation.

[15] Much of the criticism from behavioral economists and psychologists is directed towards the lacking behavioral foundations of the Rational Expectations Hypothesis [397]. According to [140], however, this criticism is misdirected because it is often raised disregarding the predictive qualities of the respective rational expectations models. In his famous dictum, the descriptive accuracy of assumptions is of no importance; they even have to be descriptively false. What matters is only the accuracy of predictions.

only slowly adapt their expectations to changes in their economic environment. During this adjustment process, their expectations will be systematically in error which seems to violate the definition of rational expectations. [368] point out though, that in fact it does not. Rational expectation models such as [260] assume that agents learn about changes of a stochastic process by the beginning of the next period. While the length of this period is never explicitly defined, it is implicitly understood as whatever time it takes for complete learning to occur. Thus, if economists are about to address the short-run phenomena during the adjustment process towards a new long-run rational expectations equilibrium, they will have to create evolutionary models of the economy.

2.4 The Economy as an Evolving Complex Adaptive System

The evolutionary character of the economy consisting of many independent and interacting agents led Holland [183], from the Santa Fe Institute in New Mexico, to describe it as an adaptive nonlinear network (ANN), a term which has largely been replaced by the notion of complex adaptive systems (CAS). The origins of this multidisciplinary concept are in mathematics and the natural sciences [430, 20, 322]. However, when trying to identify the distinguishing properties of complex adaptive systems, one realizes that no commonly agreed upon definition of complexity exists.[16] It ranges from simple statements about complicated behavior to exploding computational or informational requirements for solving certain problems. Simon [398], for instance, points out that complexity can arise from sheer numbers. In statistical mechanics, for instance, physicists study systems made of large numbers of interacting, i.e., colliding and energy exchanging, particles.[17] While these particles are often assumed to be homogeneous, the 'particles' in social or economic interactions—consumers, investors, firms, private or governmental organizations—are not. Contrary to elementary particles in physics, economic agents have preferences and a free will. They thus engage in decision-making which increases the degree of complexity found in social systems. However, local interactions between particles is not a sufficient condition to form a complex system. In his introduction to

[16] A discussion of these various definitions can be found, for instance, in [146] in the first issue of the journal *Complexity* and on [361].

[17] [248] point out that one of the first applications of the microscopic simulation method was for nuclear fission in the 1950's.

dynamic complex systems, Bar-Yam emphasizes that the distinguishing property of thermodynamic and complex systems in physics is the concept of interdependency [23]. In an economic context, interdependencies arise because the decisions made by one agent usually affect other agents in the economy.

The "Santa Fe Perspective" of complexity has been defined by [186] and by [10, p. 3-4] who list six features of the economy which may cause traditional linear mathematical methods to fail when studying CAS.

1. There are dispersed parallel interactions between many heterogeneous agents.
2. There is no global entity that controls the agent interactions. Instead of a centralized control mechanism, agents compete with each other and coordinate their actions.
3. The economic system has many hierarchical levels of organizations and interactions. Lower levels serve as building blocks for the next higher level.
4. Behaviors, strategies, and products are continuously adapted as agents learn and accumulate experience.
5. Perpetual novelty leads to new markets, new behaviors, and technologies. Niches emerge and are filled.
6. The economy works far away from any optimum or global equilibrium with constant possibilities for improvement.

Because of these difficulties for traditional analytical methods, agent-based modeling has now become a major tool to study complex adaptive systems. The predecessors of agent-based simulation of CAS in economics—cybernetics, catastrophe theory, and chaos theory—share a common thread of nonlinear relationships and an interest in discontinuities.[18] Recessions, market bubbles, or crashes are just simple examples of such discontinuities in economic systems. The most famous representative of the cybernetics approach in economics has been Jay Forrester [133, 134, 135] who created the *World3* model, the foundation for the influential Club of Rome report "*The Limits to Growth*" [295].[19] Forrester's *World3* is an aggregate model of the world economy consisting of a system of nonlinear difference equations. Even though feedback loops were modeled, learning and, thus, shifting aggregate relationships were completely ignored in the *World3* model.

[18] For a comparison of these four concepts see [193, 361]. Horgan [193] also fires a harsh criticism at complexity theory and all its predecessors.

[19] A technical documentation of the *World3* model is found in [296]. There is a 20 and 30 year update of the original study available [294, 297]. Critiques of the theoretical foundations were formulated by Nordhaus [323, 324].

An economic example of catastrophe theory can be found in [444] where stock market crashes are seen as a result of too many chartist traders relative to fundamentalist traders. Chaos theory was popularized in [149] and is characterized by a sensitive dependence on initial conditions and a deterministic system behavior which appears to be stochastic. It remains a lively area in economic research [335, 336, 54, 55] and agent-based complexity models are nowadays often tested to see whether they also exhibit chaotic behavior [266, 145].

While sharing the thread of nonlinearities and discontinuities with its predecessors, the main theme of complexity theory is that of emergence. The complicated behavior at the aggregate level is explained as the result of the interactions of constantly adapting agents whose microbehaviors are qualitatively different. Loosely speaking, emergence is the appearance of some patterns, structures, or properties at the macro level of a complex adaptive system which are not peculiar to the constituting parts.[20] Emergence is also an example of the fallacy of composition (see section 2.2.3) and is also closely related to the concept of "self-organized criticality" [20].

2.5 Some Methodological Aspects of Agent-Based Simulations

The ability to study high-dimensional complex adaptive systems in economics through agent-based simulations, however, comes at a cost. Agent-based simulations not only require inductive methods on the level of the agents, but also require inductive reasoning on the level of the modeler. Unlike most analytical models, an agent-based simulation does not produce theorems and existence proofs. It usually generates time series of state variables both on the agent level and on the macro level. However, one simulation run is just one particular realization of the potential infinite set of all possible realizations. To gain an understanding of the model behavior, researchers have to analyze the gen-

[20] A commonly accepted formal definition of emergence in CAS does not yet exist to my knowledge. Bankes [22] observed that the declaration of emergent behavior seems to be at the discretion of the researcher who intuitively relies on visual observations of differences between the micro- and macrolevel. Thus, he calls for an operational definition of emergence such as the specification of some standardized threshold level of micro- and macrolevel differences above which emergence can be declared.

erated time series with econometric methods, and general conclusions can only be made by means of inductive reasoning.[21]

Pheby [337] explains that induction is ampliative, i.e., its conclusions go beyond the available evidence. It involves reasoning from specific statements towards more general ones, while deductive arguments move from general to more specific statements.[22] Judd [208] then points to a trade-off between deductive and inductive methods in economics. One can either achieve logical purity with deductive methods in low-dimensional models of the economy while sacrificing realism, or analyze more realistic high-dimensional models and accept numerical imprecision.[23] The real issue is not whether to approximate, but where. In comparing the strengths and weaknesses of each approach, Judd concludes that computational models with inductive methods and deductive reasoning in economic theory are not substitutes, but rather complements [208]. Pheby [337] also asserts that the problem associated with induction is not so much one of rationality, but one of reliability.

The reliability issue is also connected with another objection that is often raised in connection with computational methods in economics. The black box criticism is the argument that the results of computational models are counterintuitive and incomprehensible because the computer program generating the results is only seen as an impenetrable black box. For Judd, this objection points to a more general problem of how computational studies are presented. Opening the black box by making the source code publicly available is one way to address this problem, extensive sensitivity analysis is another. The current fragmentation of the agent-based community in economics with respect to programming languages and simulation platforms may limit the impact of

---

[21] Axelrod [13] views simulation as a third research method in science. Like deduction, it starts with explicit assumptions, but it does not prove theorems. The simulated data are then analyzed by means of indexinduction induction. But unlike induction, these data come from a specified rule set and not from real world observations.

[22] According to Popper [341], the long-standing debate about the problem of induction in science was initiated by David Hume in the eighteenth century. Because past evidence tells us only about past events, to base expectations about future events on them is simply irrational. Popper then claimed to have solved Hume's induction problem by positing his principle of falsification as a direct antithesis to induction. In economics, the tension between inductive and deductive methods culminated in the famous "Methodenstreit" (conflict of methods) between Gustav Schmoller, a representative of the Younger Historical School in Germany, and Carl Menger as a member of the Austrian School of Economics.

[23] A similar trade-off between rigor and relevance in economics is presented in [291].

making source codes available.[24] But as with the lack of sufficient sensitivity analyses and standardized testing methods for computational models, it should not be a severe obstacle for the usefulness of agent-based simulation in economics.

[24] The Santa Fe Artificial Stock Market, for instance, was first programmed in Objective-C. Later, it was reprogrammed to utilize the SWARM-simulation library. Wilpert [437] ported the model to Borland-C++, while Ehrentreich [106] reprogrammed the market under Java and the RePast-library. Other artificial stock markets use mathematical program packages such as GAUSS [268, 269] or Matlab [240].

3

The Concept of Minimal Rationality

I have had my results for a long time: but I do not yet know how I am to arrive at them.

Karl Friedrich Gauss

3.1 Introduction

For a long time economists have felt that the idealization of rationality for modeling purposes needed theoretical justification. The first evolutionary arguments for rationality appeared during the 1930's in the works of Schumpeter [382] or in the famous marginalism controversy by Harrod [171] who compared new business procedures to "*mutations in nature*". Alchian [3] later claimed that competition will lead to a survival of firms with positive profits, whether or not they consciously maximize them or not.[1]

The debate whether forces of competition will assure that only the "fittest", i.e., most rational actors survive in the market was further spurred on by Friedman's famous "as-if" argument [140]. He claimed that individuals who do not behave completely rationally, for instance, by not consciously maximizing expected returns, could still be considered *as if* they do. Business behavior that is consistent with the maximization of returns hypothesis, even though it may only be by chance, will lead to prospering firms, while "natural selection" will weed out those firms whose behavior contradicts the maximization hypothesis.

[1] A discussion of the marginalist controversy can be found in [310]. Evolutionary ideas in economics, however, can be traced back even further than this. Hodgson [180] gives a comprehensive account of the history of evolutionary thought in economics.

Instead of creating economic models which are based on more realistic assumptions of individual behaviors, the arguments above suggest that one can easily take a shortcut by outright assuming rational behavior. This is also in line with Friedman's other famous claim that the realism of an assumption does not matter as long it yields correct predictions. This argument was also used by Becker [28], who showed that negatively inclined market demand curves can result both from rational and irrational behavior. In a way, this could be labeled as emergent rationality. Rationality on the macro-level, i.e., coherence with the aggregate results predicted by models populated by rational agents, would come about in spite of irrationality on the micro-level. Emergent rationality as a property of a complex adaptive system is also investigated in [69], who formally interpret both the Efficient Market Hypothesis and the Rational Expectations Hypothesis as an example of emergent behavior in the context of an artificial stock market. Evidence that indexstrategic complementarity "strategic complementarity" and "strategic substitutability" are important in determining aggregate outcomes is provided in [125]. Strategic substitutability describes a situation in which an increase in the action of one individual results in a decrease in the action of other individuals. A minority of rational agents can be sufficient to generate aggregate rationality. Under strategic complementarity, i.e., a situation in which an increase in i's action causes increases in the actions of other agents, a small amount of individual irrationality may suffice to generate large deviations from the predictions of rational models. Herding behavior in financial markets is an example for strategic complementarity. The possibility of strategic complementarity, however, is a strong argument against the unreflective use of rational agents.

The generality of the competition argument for assuming rational agents was critically examined by Winter [438] who argued that in evolution not the best conceivable, but the best available individual survives. Similarly, Blume and Easley illustrate that the link between fitness, i.e., a quantitative measure of success in a specific environment, and rationality in repeated financial markets may be weak and that rationality does not necessarily produce fit rules [41]. In general, the competition argument for rationality turns out to be highly conditional, and counterexamples are easily found [124, 98, 34]. The *as-if* argument was critiqued by Conlisk [79] and Sunder, for whom it is unsatisfactory since scientific models serve not only to predict, but also to convey an understanding to others: "*Understanding of phenomena is crucial to* *science; prediction without understanding does not build science* [413, p. 4].[2]

Despite vague references to competition and natural selection, however, the specific mechanism by which aggregate and/or individual rationality would come about had largely been left unanswered and remained a black box concept. Starting with the contribution by Nelson and Winter [319], the now burgeoning literature on modern evolutionary economics tries to open this black box by rigourously modeling dynamic processes of change in the economy. In its broadest sense, Witt characterizes evolutionary economics "*by its interest in economic change and its causes, in the motives and the understanding of the involved agents, in the processes in which change materializes, and in the consequences*" [440, p. xiii]. A good starting point for evolutionary economics is the works of [439, 49, 101].

Evolutionary economics is primarily concerned with the equilibrium selection problem [284, 435, 364]. Its other focus is to model how agents learn to become 'rational' in the economic sense [40, 137, 175]. The primary concern with equilibrium selection, however, is criticized by Börgers [48]. Since there is mixed empirical evidence whether and under which situations agents behave rationally or not, he urges first studying how rationality comes about. He emphasizes that this logically precedes the question of equilibrium selection.

3.2 Economic, Bounded, and Situational Rationality

Learning in an evolutionary context is often considered to be an adaptive process enabling agents to become rational. The prevailing concept of rationality in economics is an operational refinement of the former popular term *self-interest* and includes *perception*, *preference*, and *process rationality* [292]. Perception rationality assumes that agents behave *as if* they process information to form perceptions and beliefs by using Bayesian statistical principles. Preference rationality requires the acceptance of some well-defined axiom sets, for instance, the axioms of Herstein and Milnor [178] and Savage [374]. Typically, these

[2] An extensive discussion and critique of Friedman's evolutionary argument can be found in the chapter "Optimization and Evolution", in [180, ch. 13]. Celebrating the fiftieth anniversary of Milton Friedman's 'The methodology of positive economics', the 2003 annual meeting of the Allied Social Science Associations devoted a complete symposium to it. The papers delivered at that session, e.g., [353, 273], are contained in a special issue of the Journal of Economic Methodology.

axioms require a complete, continuous, and transitive preference ordering which should be invariant to framing and independent of irrelevant alternatives. Finally, process rationality in economics implies the maximization of an objective function subject to some constraints. I will refer to this rationality concept as economic rationality.[3]

Even though this is a widely accepted perception of rationality, attacks on it from within the economic profession or from other fields do not cease. One strand of criticism is directed towards its axiomatic foundations. By now, there is abundant empirical evidence that real decision makers often deviate from these rationality axioms [4, 109, 212, 258].[4]

The competing concept of bounded rationality in economics primarily criticizes the neoclassic notion of process rationality and calls for replacing the maximization principle through satisficing procedures.[5] Its founder, Herbert A. Simon, also refers to it as procedural rationality. He argues that economics has largely been preoccupied with the results of rational choice rather than processes of choice and thus, defines rational behavior when it is the outcome of appropriate deliberation [395, 396]. A central feature of bounded rationality is non-optimizing procedures such as simple search heuristics and decision rules [388].

However, I follow Langlois in defining rationality in a situational context [231, 233]. The method of situational analysis goes back to Max Weber, Alfred Schütz, and most recently, Karl Popper .[6] It "*insists that agents act not optimally but merely reasonably under the circumstances*" [231, p. 693]. While the concept of economic rationality emphasizes solely the agent's preferences as the basis for his best course of action, situational analysis focuses on the circumstances in which these preferences are embedded.[7]

---

[3] The term "economic rationality" might not be optimal since it is by no means "economical." It always requires the use of all our cognitive and computational capabilities—sometimes even beyond that—to perform simple tasks for which simple heuristics would suffice. In this respect, procedural or situational rationality, to be defined below, are more economical concepts of rationality.

[4] A survey about violations of the basic postulates of economic rationality can be found in [64].

[5] Herbert A. Simon revived the Scottish word "satisficing" (=satisfying) to denote problem solving and decision making that sets an aspiration level, searches until an alternative is found that is satisfactory by the aspiration level criterion, and elects that alternative [394, Part IV].

[6] For a discussion of Weber's and Schütz's contributions to rational choice theory, see [379].

[7] A typical definition of economic rationality by Hirshleifer [179] illustrates the lack of any situational components. "*In the light of one's goals (preferences), if the means chosen (actions) are appropriate, the individual is rational; if not,*

Langlois agrees with the proponents of bounded rationality that there are many situations in which it would be quite unreasonable to solve foot-long Lagrangians to determine an optimal response [231]. The culprit, however, is not the optimization approach as such, but its misapplication to ill-defined situations. Langlois thus criticizes that behaviorists threw out the baby with the bath water when the optimization approach was completely replaced with ad-hoc rules about satisficing behavior. Thus, not the logic of the situation, but ad-hoc imposed behavioral rules determine an agent's behavior. In cases in which models of bounded rationality offer a theory from where these rules come from, they become an instance of situational analysis.[8] The same can be said about neoclassical models when optimization is reasonable under the given circumstances. Thus, situational analysis combines both approaches to rationality.

It was pointed out in chapter 2 that the assumption of heterogeneity leads to indeterminate expectations of fully rational agents. Does this mean that heterogeneous agents must form their expectations in an irrational way? Barro, for instance, dubbed the appropriation of the term *rational* by the rational expectations revolution as one of its cleverest features since opponents of this approach were then "*forced into the defensive position of either being irrational or of modeling others as irrational, neither of which are comfortable positions for most economists*" [25, p. 179]. Simon [396], however, points out that the economic notion of rationality does not correspond to any classical criterion of rationality.[9] The label "rational expectation models" provides a rather unwarranted legitimization. In his view, a more neutral name such as "consistent expectations" would have been the better choice.[10]

---

irrational. 'Appropriate' here refers to method rather than result." Note that rationality is solely defined in terms of one's preferences, not in terms of the specific environment the decision maker faces.

[8] Langlois [232] points out that models in the field of New Institutional Economics are able to explain the evolution of rules. Norms and conventions emerge through a process of repeated games. Eventually, agents will discover "evolutionary stable strategies." These strategies are simple rules which then become institutionalized.

[9] A dictionary definition simply asserts that being rational means "*agreeable to reason, reasonable, sensible*". In psychology, an appropriate deliberation process is viewed as a key ingredient of rationality. It is in this tradition that Herbert Simon used bounded rationality and procedural rationality as synonyms since the focus is on the reasoning process itself rather than on identifying a goal-maximizing action in a given situation.

[10] However, the term *consistent expectations equilibria* has now been claimed by Hommes and Sorger [191] and Sögner and Mitlöhner [406]. It is an informationally less demanding concept than the Rational Expectations Hypothesis. Agents do

Since economic rationality may imply unlimited cognitive and computational capabilities on the agent's part, it thus seems to be the limiting extreme in a spectrum of possible degrees of rationality. On the other end of that spectrum, one would find "perfect" irrational behavior. This, however, would probably be far from being random since a 'perfectly irrational' agent would first figure out a rational or optimal solution to a problem and then, for some pathological reason, systematically choose to do the opposite (if an opposite action can be defined). Random behavior, on the other hand, may not be as irrational as one would think. A situation that somehow eludes analytical investigation, mixed strategies over all possible actions might be the best an agent can do. For practical purposes, the spectrum of behaviors that a modeler could chose from is most likely limited between pure random behavior and perfect economic rationality.

From these different levels of rationality, which one should he choose though? This book espouses the notion of *minimal rationality* [70], which is an approach where the agent's actions typically fall between randomness and perfection. Instead of positing unbounded rationality, why not investigate how much intelligence or rationality is necessary to explain the observed data? With a few notable exceptions, e.g., [150, 356, 413, 123], this principle has seldom been applied to economics.

3.3 Situational Analysis, Minimal Rationality, and the Prime Directive

The principle of minimal rationality is the antithesis to what Rubinstein [362] referred to as the *Prime Directive* in financial economics: "*Explain asset prices by rational models. Only if all attempts fail, resort to irrational investor behavior.*" He laments that the burgeoning behavioralist literature seems to have lost all constraints to adhere to this directive by always seeking an explanation for a financial anomaly in systematic irrational behavior.

Unfortunately, Rubinstein does not explicitly state whether he subsumes models with lesser degrees of rationality into the group of rational models or not. However, since he defines rationality as the adherence to the Savage axioms [374], it is implied that investors *maximize* expected utility using unbiased subjective probabilities. Rubinstein's insistence on the maximization principle thus implies that he does not

not need to know the underlying market equilibrium equations. They only need to have beliefs which turn out to be consistent with actual observations.

grant the rationality status to models with lesser rationality. Hence, it can be assumed that in his view the Prime Directive and the principle of minimal rationality are indeed irreconcilable positions.[11]

Friedman [140] points out that the validity of a hypothesis is not a sufficient criterion for choosing among alternative hypotheses.[12] If several hypotheses perform equally well for predictive purposes, Friedman's instrumentalist answer does not help in deciding which level of rationality should be preferred. Since descriptive validity of the model's assumptions does not matter either in his view, what is then a sufficient criterion for choosing among alternative levels of agents' rationality? The Prime Directive asserts always picking the (most) rational model. The philosophical basis for the preferential treatment of (economic) rationality is sought by Rubinstein in the long cultural and scientific tradition dating back to the ancient Greeks who spoke of "reason" as the guide to life. However, it has already been mentioned that the definition of rationality in economics is rather new and much more specific than it used to be in the scientific tradition invoked by Rubinstein.

Proponents of the bounded rationality approach to economics often invoke a principle known as Occam's razor to justify their deviance from full rationality. This principle is attributed to the Franciscan monk William of Occam (ca. 1285–1349, sometimes also known as William of Ockham) and is nowadays usually stated as follows: "*Of two competing theories or explanations, all other things being equal, the simpler one is to be preferred.*"[13] Scientific parsimony, for which Occam's razor stands, tells us to prefer the postulate that men are merely reasonable over the postulate that they are supremely rational [396]. However, in his Nobel Prize lecture, Simon [397] points out that Occam's razor has a double edge. Even though satisficing models allow for weaker rationality assumptions, optimization models can be stated more briefly than the former. The two edges thus cut in opposite directions.

The philosophical justification for the principle of minimal rationality, however, is not sought in an application of Occam's razor.[14] Minimal rationality is seen as a straightforward application of situational

---

[11] Rubinstein, however, allows for investor heterogeneity by allowing differences of opinion and uncertainties about other investors' characteristics.

[12] Friedman also claims that if there is one hypothesis that is consistent with the available evidence, then there is an infinite number that are consistent (p. 9).

[13] In Occam's own words it reads as "*Pluralitas non est ponenda sine neccesitate*", which translates into English as "*Causes are not to be multiplied beyond necessity.*"

[14] Karl Popper points out that the application of Occam's razor may be used inappropriately, which may lead to bad problem reductions [341, p. 308].

analysis combined with the economic principle. Since human agents have only limited resources, e.g., memory, time, and computational abilities, efficient reasoning strategies are the result of minimizing the involved costs in decision making given a certain aspiration level. Solving foot-long Lagrangians for complicated decisions under time pressure might not be reasonable given the fact that simple heuristics might lead to satisfactory results. The economic principle contains the idea of achieving the most with a given amount of input or minimizing the input level to achieve a given output level. If we view cognitive and computational abilities as scarce resources that the agent economizes on, an analysis of the circumstances lead the agents to employ just as much intelligence and reasoning as necessary to perform the task they face. Because of the economic principle, situational analysis and minimal rationality are two concepts that necessitate one another.[15]

3.4 Minimal Rationality and the Phillips-Curve

It is surprising that the final breakthrough of the Rational Expectations Hypothesis seemingly adhered to the principle of minimal rationality. Even though the Rational Expectation Hypothesis had been developed by John F. Muth [315] as early as 1961, the rational expectations revolution was triggered only by the influential contributions of Robert E. Lucas [260, 261]. Its power of persuasion was derived from the discussion revolving around the famous Phillips curve, for which rational expectations turned out to be the appropriate level of rationality.[16]

The Phillips-curve discourse was opened by Arthur W. Phillips [339], who found a negative relation between wage inflation $\hat{w}$ and unemployment U, a result that was verified and refined by Lipsey [254]. The Phillips curve received its political explosiveness from Samuelson [367], who took productivity growth into account and substituted wage inflation with price inflation $\hat{p}$. Since price inflation is an exogenous parameter that can largely be controlled by the monetary authorities, a society, in principal, should be able to choose from any of the combinations on this modified Phillips curve.

Both Phillips and Lipsey argued that the negative slope of the Phillips curve is caused by nominal wage pressure due to excess demand or supply on the labor market, the unemployment rate being

[15] There is, however, a possible infinite regress in determining the optimal decision procedures. A solution to this problem is suggested by Lipman [253] in an article with the telling name "*How to decide how to decide how to ...*" .

[16] An overview of this discussion can be found in [368, 142].

a proxy for the latter. However, instead of purely negotiating nominal wages, employers and employees bargain over real wages w/p for a contract period. Since the nominal wages are usually fixed during the term of the wage contract, both sides have to form expectations about the future price level within the contract period. In surveying the literature on the Phillips curve, though, Santomero and Seater conclude that the early authors did not consider expectations as a reason to include price changes as an explanatory variable [368]. In hindsight, one realizes that this first stage of the Phillips curve discussion was characterized by an implicit assumption of static inflation expectations $E\left[p_{t+1}\right] = p_t$. Since it was obviously incapable of accounting for mounting empirical evidence of shifting Phillips curves and loops in the time series, static inflation expectations turned out to be an unsatisfactory level of rationality on the agents' part.

It was Edmund Phelps [338] and Milton Friedman [141] who drew attention to the role of inflation expectations and posited the natural rate hypothesis. Search frictions and other market imperfections would determine an equilibrium rate of unemployment—the natural rate of unemployment—that is independent of nominal phenomena such as the price level or expected changes in the price. Since this natural rate hypothesis is inconsistent with a long-run trade-off between inflation and unemployment, it implies a vertical long-run Phillips curve. If the the monetary authorities increase the inflation rate, agents would soon begin to adjust their inflation expectations. Thus, instead of moving along a stable, negatively-sloped Phillips curve, the short-run Phillips curve shifts upwards, and after completion of the expectations adjustment process, the economy ends up with a higher inflation rate at the same equilibrium unemployment level.

While Friedman did not specify the exact mechanism for adjusting inflation expectations, the following second stage in the Phillips curve discussion typically employed some kind of higher order autoregressive inflation expectations

$$E\left[p_{t+1}\right] = \sum_{n=0}^{T} c_n p_{t-n}. \tag{3.1}$$

Negatively-sloped Phillips curves emerged only because agents were slow to adapt their inflation expectations to the actual change.[17] For

[17] For static inflation expectation, i.e., for $c_0 = 1$ and all other $c_n = 0$ in equation 3.1, agents do not adapt their inflation expectations at all.

positive inflation rates, individuals' inflation expectations will thus consistently be downwardly biased. While the use of extrapolative expectations

$$\begin{aligned} E\left[p_{t+1}\right] &= p_t + \alpha\left(p_t - p_{t-1}\right) \\ &= (1+\alpha)p_t - \alpha p_{t-1} \\ &= c_0 p_t + c_1 p_{t-1} \end{aligned} \tag{3.2}$$

or adaptive expectations [62, 321]]

$$\begin{aligned} E\left[p_{t+1}\right] &= E\left[p_t\right] + \alpha\left(p_t - E\left[p_t\right]\right) \\ &= \sum_{n=0}^{T} \alpha(1-\alpha)^n p_{t-n} \end{aligned} \tag{3.3}$$

certainly increased the level of rationality in expectation formation, these concepts are oriented completely backwards. Agents do not consider any information they might have about future events affecting the inflation rate. Neglecting this information is certainly less rational than using it. The systematic downward bias of autoregressive inflation expectations contradicts the rationality assumption of economic agents since, for rational expectations, errors in expectation must be random with a mean of zero. In other words, rational expectations forecasts are unbiased estimates of the true underlying stochastic inflation process. Under rational expectations, agents are using all available information. If they correctly anticipate a monetary expansion through the central bank, the expansion itself will have no economic effects. Monetary policy is effective only to the extent that it deceives economic agents.

The development of the Rational Expectations Hypothesis thus closely followed the principle of minimal rationality. Only when resorting to fully rational individuals, the models were able to explain the observed empirical relationships between unemployment and price inflation. It is thus not surprising that the assumption of rational expectations became the standard assumption in economic theory. It is, however, likely that in other problem fields, lesser degrees of agent rationality may suffice to equally explain the empirical facts. The principle of minimal rationality would then prescribe the use of the model that requires the least amount of rationality.

4

Learning in Economics

> *You can always count on Americans to do the right thing — after they've tried everything else.*
>
> Winston Churchill

4.1 Introduction

Replacing fully rational economic agents with boundedly or minimally rational agents leads to the necessity of learning. Another motivation to endow agents with learning routines is to predict unique outcomes when a model possesses multiple equilibria. Being less than perfectly rational or omniscient does not mean that agents are irrational. Instead, they act according to the knowledge and skills they are equipped with. Learning agents then have the potential to gradually discover an optimal or at least satisfactory course of action.

After defining learning in more detail, this chapter briefly discusses the rationality-based Bayesian learning model. At the center of interest, however, are biologically inspired learning models. While replicator dynamics are extensively used in game theory, genetic algorithms or classifier systems are widely used in agent-based simulations.

4.2 Definitions of Learning

Modeling learning processes requires us to define learning more precisely. In the psychological literature, a behaviorist approach is usually taken. Burrhus F. Skinner, perhaps the most famous among the behaviorists, defines learning as a "*change in probability of response*" [399, p. 199]. Kimble set forth a widely accepted definition according to which "*learning is a relatively permanent change in a behavior potentiality which occurs as a result of reinforced practice*" [219, p. 82]. Bower and Hilgard add that it is necessary that this change "*cannot be explained on the basis of the subject's native response tendencies, maturation, or temporary states (such as fatigue, drunkenness, drives, and so on)*" [50, p. 11].

The emphasis on change in response probability or behavior potentiality indicates that learning may not be translated into behavior until some time after the learning has taken place [177]. Because of this, learning may not be immediately observable. According to Brenner, this specification does not matter in an economic context. Economists are not interested in the study of learning processes themselves, it is only their consequences, i.e., changes in economic behavior, that are of interest to them [52].

Through the principle of reinforcement, i.e., rewards or punishments that are associated with certain actions, individuals identify successful behaviors and attach higher activation probabilities to them. Failing behaviors will be less likely to be acted upon in the future. Brenner, however, considers this reinforcement learning, or conditioning as it was originally named by psychologists, as too restrictive in an economic context. While humans share this simple way of learning with animals, we are, in addition, able to reflect on our actions and their consequences. The psychology literature labels this as cognitive learning. Thus, Brenner's learning definition includes "*any cognitive or non-cognitive processing of experience that leads to a direct or latent change in economic behaviour, or to a change of cognitive pattern that influences future learning processes*" [52, p. 3]. Because our cognitive resources are scarce and thus, cannot be applied to every situation, Brenner [53] argues that many of our actions are a result of reinforcement learning. He considers the effect of cognitive learning on behavior, however, stronger than that of reinforcement learning.

The simultaneous existence of at least two kinds of learning processes implies different ways of how to model them. Brenner [53] notices an increasing number of learning models in economics in the past few years and points out that there are many possible ways to classify them.[1] Based on their origin, he distinguishes them as psychology-based models such as reinforcement learning, rationality-based models such as Bayesian or least-square learning, or models inspired from computer science and biology such as evolutionary algorithms or neural networks. He remarks, however, that different learning models are usually applied

[1] Other surveys on learning theories in economics can be found in [400, 405].

in different economic fields and, therefore, might be classified accordingly. While Bayesian and least-square learning dominate in macroeconomics, game theorists seem to prefer fictitious play, replicator dynamics, and other adaptive learning models. Experimental economists tend to employ reinforcement learning, fictitious play, and indexlearning direction theory learning direction theory, while evolutionary algorithms and genetic programming are prominent in agent-based computational models. Besides the fact that mathematical economists are limited in their choice to analytically treatable models, these different preferences within the economics profession mainly seem to reflect historical developments rather than inevitable design decisions.

4.3 Rationality-Based Learning Models

In the beginnings of modeling learning in economics, normative approaches such as Bayesian or least-square learning have been the most dominant. These rationally-based models assume that agents behave optimally, first by taking the best action regarding the available information, and second, by rationally updating their prior opinion as soon as new evidence becomes available.[2] Descriptions of both learning types can be found in [52, 53]. In the case of Bayesian learning, individuals start out with an initial set of hypotheses $\mathcal{H}$ about the situation they face. Each hypothesis $h \in \mathcal{H}$ is characterized by a list of subjective probabilities $P(e|h)$ with which mutually exclusive events e in the set of possible events $\mathcal{E}$ occur, i.e., for each hypothesis $\sum_{e\in\mathcal{E}} P(e|h) = 1$. In period 0, agents typically assign to each hypothesis h an equal probability $p(h, 0)$. If better information is available, these so-called prior probabilities might differ from each other. The set of possible events has to be complete and complementary, i.e., $\sum_{h\in\mathcal{H}} p(h,t) = 1$ must hold for all periods t.

As agents learn about new evidence in the next period, they update their initial subjective probabilities associated with each hypothesis according to

$$p(h, t+1) = \frac{P(e(t)|h) \cdot p(h,t)}{\sum_{h\in\mathcal{H}} P(e(t)|h) \cdot p(h,t)}. \tag{4.1}$$

Through this equation, an observed event increases the probability of that hypothesis for which the occurrence of this event is most likely.

[2] Blume and Easley, however, consider the term “rational learning” as a poorly chosen euphemism. In their view, Bayesian learning does not deserve this normative judgement [42].

The normalization through the denominator ensures that the condition $\sum_{h\in\mathcal{H}} p(h,t) = 1$ is fulfilled for each period. As time proceeds, the probability $p(h,t)$ for the correct hypothesis should converge to 1, while the probabilities for the other hypotheses should approach zero. Given these updated probabilities for each hypothesis, agents chose an action that maximizes their expected utility.

While Brenner points out that the basic mechanism of Bayesian learning roughly corresponds to the psychological notion of cognitive learning, he admits that it lacks empirical or experimental justification. It is very doubtful whether people are indeed able to do the necessary calculations that Bayesian updating requires. He also refers to the psychology literature which finds that people tend not to simultaneously consider many competing hypotheses [52]. Experimental evidence such as Monty Hall's three-door anomaly, too, suggests that agents consistently deviate from the rational Bayesian solution [139].

4.4 Biologically Inspired Learning Models

Evolutionary algorithms such as evolutionary strategies [351] and genetic algorithms [182] were initially developed to optimize technical systems.[3] They are designed to mimic natural evolutionary processes since they have been very successful in creating well-adapted species to changing environments. Evolution is sometimes simply characterized as "transmission with adaptations", whereby adaptations simply refer to differences. John Holland defines adaptation more precisely as any "*process whereby a structure is progressively modified to give better performance*" [182, p. xiii]. Since these adapted structures may be as diverse as protein molecules, human brains, or investment rules, adaptation is a rather universal principle applicable to many different fields.

Because learning theories are concerned with how behavioral decision rules are modified, evolutionary algorithms were soon thought to describe individual and social learning processes [51, 355, 401]. The connection is usually made by analogy [354]. For instance, the genotype in biology is thought to correspond to behavioral decision rules. The process of *mutation* of a gene could be translated into the creation of new decision rules through experiments or unintended mistakes. New decision rules can also be the result of communication between agents who derive them by combining previously successful strategies. This would correspond to the concept of *crossover*. Contrary to mutation

[3] An overview of the history of evolutionary computation can be found in [93].

and crossover, *replication*, or *reproduction* as it is often referred to, is variety preserving and can be interpreted as the imitation of successful decision rules. *Selection* is a variety restricting process of choosing individuals for reproduction, while *fitness* is the quality that is selected for [162]. As such, fitness only represents a propensity for evolutionary success, which is a retrospective measure of the relative increase or decrease in the whole population of possible solutions or decision rules.[4] As Hodgson [180] points out, even though fitness implies an expectation of success, it does not necessarily means survival. The fittest solution can go extinct while bad solutions have a positive probability of survival.[5]

It has repeatedly been pointed out that the analogies between learning and evolutionary algorithms are not isomorphic. Brenner [51] and Slembeck [401] emphasize that in spite of similar structure, evolutionary algorithms and learning processes possess crucial differences. First, individuals are likely to improve their decision rules less stochastically than suggested by the genetic operators , crossover, and selection, which are essentially stochastic processes. Brenner finds that motivational aspects cannot be adequately captured by evolutionary algorithms. Second, they generally have no memory, thus, past experiences of individuals are hard to model.[6] Third, the definition of a fitness function is easier and more objective than the way individuals attribute success to certain decision rules. In spite of these concerns, evolutionary algorithms continue to be widely used to model learning processes. A perfect analogy of evolutionary algorithms with learning processes might not be necessary as long as the results adequately describe changes in human behavior.[7]

---

[4] In population genetics , fitness is often defined as reproductive success which is only an ex post measure. Since species or members of a specific species do not come with a fitness function attached to them, evolutionary biology has tremendous difficulties in defining fitness. Lewontin admits that "*although there is no difficulty in theory in estimating fitnesses, in practice the difficulties are insuperable. To the present moment no one has succeeded in measuring with any accuracy the net fitness of genotypes for any locus in any environment in nature*" [250, p. 236].

[5] In evolutionary algorithms, this depends on the specific implementation of the selection operator. *Elitism*, for instance, retains some of the fittest individuals in each generation [308].

[6] Even though most evolutionary algorithms have the Markov property, past experiences could be implicitly captured in an appropriate definition of the fitness function.

[7] The application of "Darwinism" to economics and the use of biological analogies, however, is vehemently defended by Hodgson [181]. He finds that most criticisms of biological analogies are unfounded and that Darwinism involves a basic philo-

4.4.1 Learning Through Replicator Dynamics

In the marginalism debate, Alchian [3] and Friedman [140] suggested that irrational behavior would be driven out of the market. In a way, they must have considered markets to act as efficient evaluation and selection mechanisms. In modern evolutionary economics, their competitive argument could be modeled through what is now known as replicator equations[8] [383]. Replicator equations are based on Ronald A. Fisher's [128] *Fundamental Theorem of Natural Selection* (also known as Fisher's Law) which formalized the notion of the survival of the fittest in evolutionary biology.[9] Consider a population of n competing "species" which are characterized by their genetic fitness f. In evolutionary economics the "species" may simply be taken as a synonym for competing technologies, strategies, or trading rules. In the marginalism debate, the equivalent to a species is a group of firms exhibiting the same level of rationality. The higher their degree of rationality, the higher their fitness level. The overall population of firms is described by a vector of relative frequencies $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $\sum_{i=1}^{n} x_i = 1$ with which these types of firms exist. For each firm i, its next period share within the total population calculates as

$$x_i(t+1) = x_i(t)\frac{f_i(t)}{\overline{f(t)}} \tag{4.2}$$

with

$$\overline{f(t)} = \sum_{j=1}^{n} x_j(t) f_j$$

sophical commitment to detailed, cumulative, and causal explanations. While Darwinism has by now made substantial inroads into economics, more than 100 years earlier, Thorstein Veblen [429] expressed dismay that economics had not become an evolutionary science yet.

[8] The term *replicator equations* for Fisher's selection equation was suggested by Schuster and Sigmund [383], who pointed out the close relationship to the concept of replicators which Richard Dawkins [90] proposed in 1976. Replicators refer to a collection of molecules capable of either producing a copy of themselves, or replicating by using raw materials either from themselves or from other sources. Schuster and Sigmund also showed that Fisher's selection equation is a special case of the famous predator-prey equations (Lotka-Volterra equations).

[9] The term "survival of the fittest" had originally been coined by Herbert Spencer, who anticipated many of Darwin's evolutionary ideas. Darwin, who first used the expression "struggle for existence", adopted this expression only later, i.e., in his sixth edition of "*The Origin of Species*". Spencer became known as the father of "social Darwinism", which is widely seen today as a misapplication of evolutionary ideas to the social realm.

being the average fitness of the whole population.[10] The evolutionary process described by this kind of replicator dynamics (4.2) maximizes the average fitness of the population since firms with above average fitness prosper, while the share of below average fitness firms shrinks. In the limit, only the most rational type of firm in the initial population survives.

However, this result holds only for constant fitness functions, i.e., if fitness is independent of the frequency distribution in the population and not otherwise influenced by a changing environment. Most often, and possibly in the marginalism debate as well, fitness is time and frequency-dependent and equation 4.2 should be rewritten as

$$x_i(t+1) = x_i(t)\frac{f_i(\boldsymbol{x},t)}{f(\boldsymbol{x},t)}, \quad i = 1,\ldots,n. \tag{4.3}$$

In this case, multiple equilibria are possible, and the mean fitness of the population may or may not increase [234]. That is, under the conditions likely to exist in competitive markets, the formalization of Friedman's argument with replicator dynamics shows that evolutionary market forces do not ensure emerging rationality as an inevitable result.[11]

Replicator equations are often used to model social learning in which agents learn from others by imitating other, more successful agents [377, 378]. However, replicator equations predetermine the benchmark level of efficiency to that of the most efficient firm contained in the initial population [102]. Since replicator dynamics account for selection only and contain no element of discovery or invention, a superior behavior missing in the initial population of behavioral rules cannot be found. Novelty can be included with the so-called selection-mutation equation, sometimes also referred to as the Fisher-Eigen equation.[12] This is an extension of simple replicator dynamics by which novelty is accounted for by adding a mutation term to equation 4.2.[13] Since replicator dynamics and the selection-mutation equation allow for a

---

[10] The continuous time version of the replicator equation is more widely known and reads as $\dot{x}_i = x_i\left(f_i(t) - \bar{f}\right)$. However, since computer simulations can only evolve in discrete time steps, the discrete versions of any formula will usually be given.

[11] In a more detailed analysis, Blume and Easley show on the one hand that market selection favors profit maximizing firms. The long-run behavior of evolutionary market models, however, is not consistent with equilibrium models based on the profit-maximization hypothesis [43].

[12] This equation is named after two evolutionary biologists, Ronald A. Fisher from Great Britain, and Manfred Eigen from Germany.

[13] More details are given in [52].

rigorous mathematical treatment, its domain is mainly in evolutionary game theory. In simulation based approaches, genetic algorithms, classifier systems, and neural networks prevail.

4.4.2 Learning Through Genetic Algorithms

Genetic algorithms (GAs) were developed in the 1970ies by John Holland [182] at about the same time that evolutionary strategies were designed by Ingo Rechenberg [351]. The main difference between these two approaches is that GAs were designed to work on binary coded solutions (so-called genetic individuals) while evolutionary strategies work directly with real numbers. While the issue of binary versus real-valued representation caused a heated debate between Holland and Rechenberg1973, this distinction currently seems to be vanishing. As we will see, the GA in the SFI-ASM uses a combination of binary representation for the condition parts and a real valued representation for the forecast parameters. Brenner considers it a puzzle why modelers have not moved to use evolutionary strategies, which, he claims, seem more suited for economic problems [53]. The reason for this might be better understood from a historical perspective, possibly reflecting a path dependency. Classic textbooks on GAs have been written by Goldberg [154] and Mitchell [308], while Dawid [89] and Riechmann [356] have written newer textbooks with a special focus on learning in an economic context.

The basic element in a GA is a *population* or *gene pool* consisting of different *genetic individuals*. Since the gene pool in GAs most often contains only *haploid* genetic individuals, it has become commonplace in the GA literature to refer to these genetic individuals as *chromosomes* or, less often, as *genomes* or *genotypes*.[14] Chromosomes are made up of *genes* and the position i of a gene on a chromosome is known as its *locus*. A binary gene has two *allele* values, for instance, 0 and 1. The *cardinality* of a gene refers to the number of different alleles that a gene can take. A gene with cardinality 3 has three alleles, e.g., 0, 1, and 2.

[14] In population genetics , one usually distinguishes between haploid, diploid, or polyploid species. Haploid species have only one copy of each chromosome while diploid species, the most common case in nature, possess two copies of each chromosome. For haploid individuals with only one chromosome, the terms "genetic individual", "chromosome", "genotype", and "genome" coincide, yet one should be aware that they have slightly different meanings for diploid and polyploid species in evolutionary biology. For a definition of these terms in the context of population genetics see, for instance, [174, 80]. In the following, the terms genetic individual, chromosome, genotype, and genome will be used interchangeably.

In most GA representations, all genes have the same cardinality and a common allele set.

GAs are inherently parallel. Since a GA has to work on a population of several genetic individuals, in each time step several new solutions are simultaneously created and then tested for their performance. Little synchronization is needed and parallelization by means of distributing the numeric calculations onto multiple processors can be easily accomplished.

By now, there have been many different variants of GAs developed, some of which are are discussed in [112]. Most of them descended from the so-called canonical GA developed by Holland [182] and described in [154]. The basic structure of the canonical GA is as shown in figure 4.1.

BEGIN
    randomly create initial population P_0
    WHILE NOT stopping condition DO
        $t := t + 1$
        Evaluation of P_{t-1}
        Selection from P_{t-1} and Reproduction into P_t
        Recombination (Crossover) on P_t
        Mutation on P_t
    END
END

Fig. 4.1. Basic structure of the canonical GA.

The canonical GA is a non-overlapping populations approach in the sense that the offspring generation created through selective reproduction, mutation, and crossover completely replaces the parent generation. In the canonical GA, the selection probability Π_i of genetic individual i to be reproduced into the next generation is proportional to its relative fitness

$$\Pi_i = \frac{f_i}{\sum_{j=1}^{N} f_j}, \tag{4.4}$$

N being the population size. In other words, i's expected number of offspring equals its relative fitness in the population.

In contrast to the canonical GA, generation gap methods implement an overlapping generations such that the parent and offspring generations compete with each other.[15] The "generation gap" parameter

[15] Generation gap methods are discussed in [373].

G refers to the amount of overlap between parents and offspring. For $G = 1$, we have a complete generational replacement as in the canonical GA. A generation gap of $G = 0.2$ would replace 20% of the parent generation, while the so-called "steady-state GA" usually replaces only one or two individuals, i.e., $G = 1/N$ or $G = 2/N$.

Social Versus Individual GA Learning

There is an ongoing debate about whether GAs represent learning on the social or individual level. Riechmann, for instance, conceives of GA learning only as a way of social learning. "*As social learning always means learning from others, there simply is no GA learning by single, isolated agents*" [354, p. 226]. Whether GA learning corresponds to social or individual learning depends on the way it is implemented. When constructed as social learning, each agent is represented as one solution string. The population of agents and the population of possible solutions coincide in this case. It is also popular to model social learning by having a common pool of behavioral strategies to which all agents have equal access [241]. When all agents are equipped with a rule set of their own, and agents do not communicate about their favorite rules, GA learning takes place only on the individual level [9, 244]. Vriend [431] has demonstrated in a standard Cournot oligopoly game that social or individual GA learning may yield sharply different results. He showed that profit maximizing firms that used a social GA learning algorithm converged to the Cournot-Nash output level while individual learning moved the firms to the competitive Walrasian output level. The reason for this divergence is the so-called *spite effect* in social learning, i.e., agents may choose an action because it hurts others more than it hurts themselves. In spite of having the same objective function, social learning firms compare their single production rule to those of other firms, all of which are active in the same period. Production rules of individual learners, on the other hand, compete with each other in the learning process, but not in the Cournot market. Since an individual firm uses only one of its rules in any given period, the generated payoff of that rule is not influenced by the payoffs of other production rules in the firm's rule set. While there is still a spite effect in the market, it does not affect the learning process.

When employing learning algorithms, one has to be aware of whether the aggregate outcome can be a result of a possible spite effect. The decision about whether one should choose an individual or social learning algorithm is not a matter of convention or computational convenience, but should depend on the specific problem to be solved. Vriend also points the possibility of hybrid approaches in which elements of both learning variants are combined.[16]

Binary Encoding and Decoding

Historically, the binary encoding of objects as genetic individuals has been the predominant approach in GAs. Consider a binary chromosome of length L that represents a possible action or solution to an optimization problem. Such an action might refer to quantities such as how much to produce or to invest. This binary representation for the GA and the use of real-valued numbers outside the GA requires a constant encoding and decoding of the object variables.[17] The following is a short example of how to decode a binary chromosome. Let us assume that a choice variable can take on values in the interval $[D^{min} = -10, D^{max} = 40]$. Consider, for instance, a chromosome i with a length of 16 bits

$$S_i = \{\, 0\,1\,1\,0\,1\,0\,1\,0\,1\,1\,0\,0\,1\,0\,1\,1\}.$$

A direct transformation into the decimal system would result in a numerical value of

$$B(S_i) = 2^{14} + 2^{13} + 2^{11} + 2^9 + 2^7 + 2^6 + 2^3 + 2^1 + 2^0 = 27,339.$$

The numerical decimal value of i is then determined as

$$\begin{aligned} V(S_i) &= D^{min} + \frac{D^{max} - D^{min}}{2^L - 1} \times B(S_i) \qquad (4.5) \\ &= -20 + \frac{60}{65,535} \times 27,339 \\ &\approx 5.03. \end{aligned}$$

The maximum attainable precision for an object variable is given by

$$\frac{D^{max} - D^{min}}{2^L - 1}. \qquad (4.6)$$

To achieve higher numerical precision, longer binary chromosomes have to be used.

[16] A combination of social and individual learning would correspond to migration between different populations. In population genetics this is analyzed with so-called island or stepping-stone models [126].

[17] Often, the adherents of the competing school of evolutionary strategies criticized the necessity of constantly coding and decoding the choice variables [351].

Instead of the pure binary coding as described above, Gray codes[18] are frequently used in GA applications. Gray coded numbers are characterized by the "adjacency property", i.e., adjacent integers differ only in one single bit and have a Hamming distance of one. In pure binary representations, adjacent integers may lie many bit flips apart, which makes it less likely for a single bit flip to cause only small changes. Hollstien found that a GA with Gray code representation worked slightly better for optimizing functions than a pure binary encoding [188]. He attributed this to the adjacency property since it improves the chances that mutation, i.e., the flipping of a single bit, causes only incremental improvements. Although fewer mutations may cause large changes, those which do may cause even bigger changes than the flipping of the most significant bit in a pure binary representation.[19]

Evaluation

The definition of an objective function $f : A_x \to \mathbb{R}$, A_x being the state space of the optimization problem, is perhaps the most important element in any evolutionary algorithm. It evaluates the quality of genetic individuals and thus, specifies the direction towards which the GA should try to improve them. Generally, this definition is dependent on the specific optimization problem. In an economic context, usually some costs are to be minimized or profits to be maximized. If the objective function produces negative values or has to be minimized, it has to be mapped to a fitness function $\Phi : A_x \to \mathbb{R}_+$. Often, an additional scaling function is applied.[20]

It is sometimes helpful to imagine the objective function as defining a fitness landscape over the underlying state space. Adaptation could then be visualized as a process of "hill-climbing" through minor modifications towards "peaks" of high fitness on this fitness landscape [215]. Flat fitness landscapes with a single peak, so-called "needle-in-a-haystack" landscapes, or random fitness landscapes, are notoriously hard to solve, yet they are often used to analyze the performance of different evolutionary algorithms.[21]

---

[18] Gray codes are named after Frank Gray, who developed them in 1953 for use in shaft encoders. More on the history and the use of Gray codes can be found in [173, 144, 343].

[19] An overview of different encoding and decoding functions for GAs is given in [94].

[20] More information on fitness scaling and possible mappings of objective values into fitness functions can be found in [162].

[21] The analogy between the fitness function and a landscape is due to Sewall Wright who considered forces other than selection in his shifting balance theory, in par-

Unlike with simple function optimization or many combinatorial problems such as the travelling salesman problem, many GAs work with state dependent fitness functions.[22] Riechmann [355] defines such GAs as economic GAs. A state dependent fitness function means that the fitness of an agent does not only depend on his strategy, but also on the strategies of all other agents.[23] Actually, state dependent fitness functions are even more common in biological evolution where species are constantly adapting to each other. Stuart Kauffman , for instance, speaks of coupled fitness landscapes and species that constantly coevolve [215]. Any step of one species on its fitness landscape deforms the landscape for another species which will react accordingly. This is also known as an evolutionary arms race.

Selection and Reproduction

The genetic operator of selection chooses genetic individuals for reproduction, i.e., chosen solutions will create identical offspring who become members of the next population. Conceptually, selection for reproduction can be divided into two different steps. First, a selection algorithm assigns each individual a real-valued expected reproduction value, i.e., its expected number of offspring. The distribution of expected offspring could be visualized as a "wheel of fortune" where the selection algorithm determines the size of the section assigned to each individual. In a second step, a sampling algorithm chooses an integer number of actual offspring for each parent. The two most used sampling algorithms are roulette wheel (RW) and stochastic universal sampling (SUS).

There are numerous selection schemes. One of the most popular selection schemes is proportional selection [182] where an individual's i selection probability $\Pi_{RW}(i)$ is proportionate to its relative fitness, i.e.,

$$\Pi_{RW}(i) = \frac{\Phi(i)}{\sum_{j=1}^{N} \Phi(j)}, \tag{4.7}$$

N being the population size. Since fitness proportionate selection requires positive fitness values, it is sometimes necessary to add a positive

ticular, genetic drift, as a means of moving a population across valleys in the fitness surface [442].

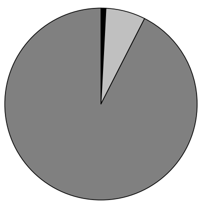
[22] GAs with state dependent functions are analyzed in more detail in [89].

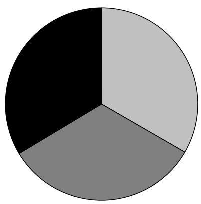
[23] Riechmann [356] also points out that the term "strategy" is used differently in game theory and in the theory of economic GA learning. Whereas in game theory, "strategy" refers to a set of contingent actions, it might refer simply to any myopic action in GA learning.

constant C to all raw fitness values $\Phi(i)$. This, however, causes a problem since simple fitness proportionate selection is not scaling invariant, i.e., adding an identical offset C tends to equalize the selection probabilities.

Table 4.1. The property of scaling invariance for fitness proportionate selection means that adding an identical offset—in this case 100 in column four—to the raw fitness values $\Phi(i)$ of three genetic individuals almost equalizes their selection probabilities $\Pi_{RW}(i)$.

Roulette Wheel 1			Roulette Wheel 2	
ind.	$\Phi(i)$	$\Pi_{RW}(i)$ (%)	$\Phi(i) + 100$	$\Pi_{RW}(i)$ (%)
1	0.01	0.93	100.01	33.22
2	0.07	6.54	100.07	33.24
3	0.99	92.52	100.99	33.54





The problem of scaling invariance for fitness proportionate selection is illustrated in table 4.1. For the second roulette wheel in columns four and five, an offset of 100 is added to the raw fitness values given in column two. When using the raw fitness values, the worst genetic individual has a very small selection probability of less than 1 %, while the best individual is chosen almost 93 % of the time. When an offset of 100 is added, the selection probabilities are 33.22 % and 33.54 %, i.e., the worst and the best genetic individual have almost an equal probability of being selected. Various approaches to dealing with scaling invariance are discussed in [154, 308]. Sigma scaling, for instance, remedies the problem by also taking into account the mean and standard deviation of fitness values in a population.

In contrast to fitness proportionate selection, the selection probabilities in ranking-based selection [163] depend only on the rank ordering of the individuals in the current population. Ranking eliminates the need for fitness scaling, since the selection intensity is maintained even if the fitness values converge to a very narrow range, as is often the case as the population evolves. Other popular selection procedures are tournament selection [38], and Boltzmann selection [271]. Additional selection mechanism are discussed in [131]. Most of these selection schemes are analyzed and compared in [39, 167].

Recombination (Crossover)

Since crossover involves the exchange of genetic material between two parents, it is often called a sexual genetic operator. By selecting two chromosomes from the mating pool with high fitness values as parents, genetic material from both chromosomes is exchanged to create two new offspring. In the canonical GA, the offspring replace their parents, and simple *one point crossover* is used.

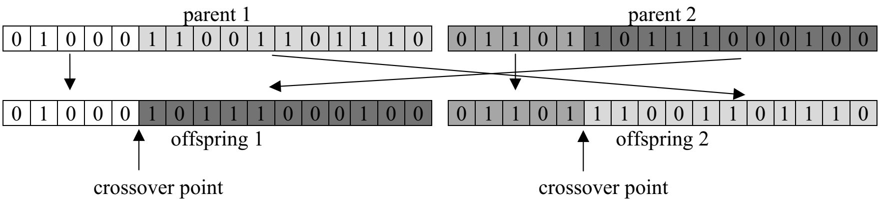


Fig. 4.2. Example of one-point crossover with a random break point after the fifth bit.

In one point crossover, a break point between 1 and $L-1$ is randomly chosen. Contiguous bit segments that begin or end at the break point are exchanged between the two parents. One-point crossover, however, is rarely used in practice [47].

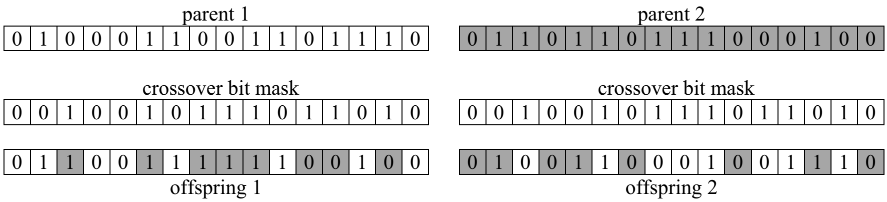


Fig. 4.3. Uniform crossover with a randomly generated crossover mask.

The class of k-point crossover operators uses several break points while the *uniform crossover* operator sets k at its maximum $L-1$.

Uniform crossover has been shown to be superior in most cases to one-point and two-point crossover [415]. In a first step, a binary mask of length L is randomly created. Whenever the value in this crossover mask is 0, the corresponding bit from parent one is copied to the first offspring, and the second offspring receives the corresponding bit from parent two. For value 1 in the crossover mask, the bit exchange is reversed. Other crossover operators are, for instance, *shuffle crossover*, and *punctuated crossover*. For the travelling salesman problem, special crossover operators such as PMX (partially mapped crossover) or OMX (ordinal mapped crossover) have been designed. These and other crossover operators are discussed in [182, 154, 47].

In a social learning context, recombination is often interpreted as communication or information exchange between agents. If modeled as individual learning, crossover might be seen as individuals experimenting by combining some of their successful strategies.

Mutation

Mutation refers to the creation of a new solution from only one parent through spontaneous allele changes. It is an important element of any evolutionary algorithm to prevent premature convergence. Unlike crossover, mutation is able to introduce non-existing alleles at previously homogeneous bit positions. If, for instance, the most significant bit is set in none of the trading rules, yet the optimal solution would require it to be set, a GA without mutation could never reach the optimum. Mutation aims at maintaining a sufficiently diverse population since without it, the final outcome of a GA is a uniform population with no diversity at all.

In mutation, each parent bit is flipped with mutation probability Π_M. Mutation probabilities are usually chosen to be very low, usually in the magnitude of $10^{-2} - 10^{-3}$. Originally intended only as a background operator [154], recent empirical and theoretical research demonstrates improvements in the GA when strengthening the role of mutation as a search operator [17]. Fogarty [130] empirically finds that GAs with initially bigger mutation rates which exponentially decreases over time do better than fixed mutation probabilities, a finding that is theoretically confirmed by Bäck [16].

As with crossover, a mutated offspring replaces its parent in the canonical GA. In both social and individual learning contexts, mutation is usually interpreted as a metaphor for either experimentation or discovery through unintended mistakes.

The Schema Theorem and the Building Block Hypothesis

The Schema Theorem and the Building Block Hypothesis developed by Holland [182] are attempts to theoretically explain why GAs work in practice. For a binary representation, a schema ξ is a ternary template consisting of the symbols 0, 1, and #, where # is known as the don't care symbol.[24] The *defining positions* of a schema are all non-# signs, and its *order* $o(\xi)$ is the number of all its defining positions. The distance between the first and last defining position of schema ξ is called its *defining length* $\delta(\xi)$, while members of a schema are called *instances*. For instance, the schema $\xi = \{1\#\#010\}$ is of order four and has a defining length of five. Its instances are $\{100010, 101010, 110010, 111010\}$.

The Schema Theorem, sometimes also called the Fundamental Theorem of GAs [154], derives a lower bound for the number of instances of a schema in the next generation, given the number of instances in the current population $N_\xi(t)$, fitness proportionate selection, and complete generational replacement:

$$N_\xi(t+1)_{|N_\xi(t)} \geq N_\xi(t)\frac{\widehat{f_\xi}(t)}{\overline{f}(t)}[1 - D_c(\xi)][1 - D_m(\xi)]. \qquad (4.8)$$

The *observed fitness* $\widehat{f_\xi}(t)$ is given by the mean fitness of all instances belonging to schema ξ in generation t, while $\overline{f}$ is the mean fitness of the whole population at t. $D_c(\xi)$ and $D_m(\xi)$ denote upper bounds for the disruptive effects on schema membership by crossover and mutation.[25] These disruptive effects have to be analyzed specifically for each operator. For instance, one-point crossover with crossover probability Π_c can disrupt the schema membership of an offspring only if its cross point falls within the defining region of a schema. For chromosomes of length L, the upper bound for the disruptive effect of one-point crossover on schema membership is then given by $\Pi_c\delta(\xi)/(L-1)$. Similarly, for mutation in which each gene is altered with mutation probability Π_m, schema membership is only disrupted if the mutation operator alters at least one of the defining positions of the chromosome. Since, by definition, the probability that none of the defining positions of a schema is altered is $(1 - \Pi_m)^{o(\xi)}$, the membership disruption $D_m(\xi)$ caused by

[24] A more general analysis for alphabets with higher cardinality can be found in [347].

[25] Note that if we neglect these disruptive terms, i.e., if we have a GA without crossover and mutation, the evolution of the GA population is completely driven by the dynamics of the selection operator. The inequality equation 4.8 then reduces to the already known replicator dynamics equation 4.3 on page 35.

mutation is $1 - (1 - \Pi_m)^{o(\xi)}$. For $\Pi_m \ll 1$, this is approximated by $\Pi_m o(\xi)$. That is, for one point crossover and point mutation, equation 4.8 can be written as

$$N_\xi(t+1)_{|N_\xi(t)} \geq N_\xi(t) \frac{\widehat{f_\xi}(t)}{\overline{f}(t)} \left[1 - \Pi_c \frac{\delta(\xi)}{L-1}\right] \left[1 - \Pi_m o(\xi)\right]. \qquad (4.9)$$

The Schema Theorem thus shows how low order schemata with small defining length and above average fitness will quickly propagate through a population. Holland and Goldberg [182, 154] call these schemata building blocks, which, through recombination, are likely to form more complex solutions with potentially higher fitness. They refer to this conjecture as the Building Block Hypothesis. However, while the Schema Theorem is easy to prove, Radcliffe doubts that the Building Block Hypothesis can ever be proved [347].

Overall, the significance of the Schema Theorem and of the Building Block Hypothesis remains highly debated [352, 92]. One of the shortcomings of the schema theorem is, for instance, that it only considers the disruptive effects of crossover and mutation [89]. In addition, the notion of *implicit parallelism* [182], i.e., the idea that GAs can process many different schemata at once, and its relation to the cardinality of the representation remains highly disputed. Implicit parallelism and the supposed optimality of binary representation is also discussed in [154, 347]. An alternative set of principles for understanding the working of a GA—the *evolutionary progress principle*, the *genetic repair hypothesis*, and the *mutation-induced speciation by recombination principle*—is suggested by Beyer [31].

4.4.3 Learning Through Classifier Systems

Classifier systems are machine learning systems that derive their name from their ability to actively classify their environment into recurrent patterns. Classifier systems in the tradition of the Michigan-approach were developed by John Holland and Judy Reitman at the University of Michigan [187], while those in the tradition of the Pitt-approach were developed by Smith at the University of Pittsburgh [403].[26] Both are a synthesis of expert systems as described in [301, 434] and Holland's genetic algorithms. In contrast to expert systems, though, classifier

[26] For a comparison of both styles, see [402]. The following description is based on Michigan-style classifier systems since they are more popular and have undergone more development.

systems do not need to be initialized with a fixed rule for every possible situation. Instead, classifier systems are flexible in that a GA adapts the rule sets to a changing environment. Today, it is common to speak of learning classifier systems (LCS) instead of classifier systems. The general structure of LCS and their interactions with the environment are shown in figure 4.4.

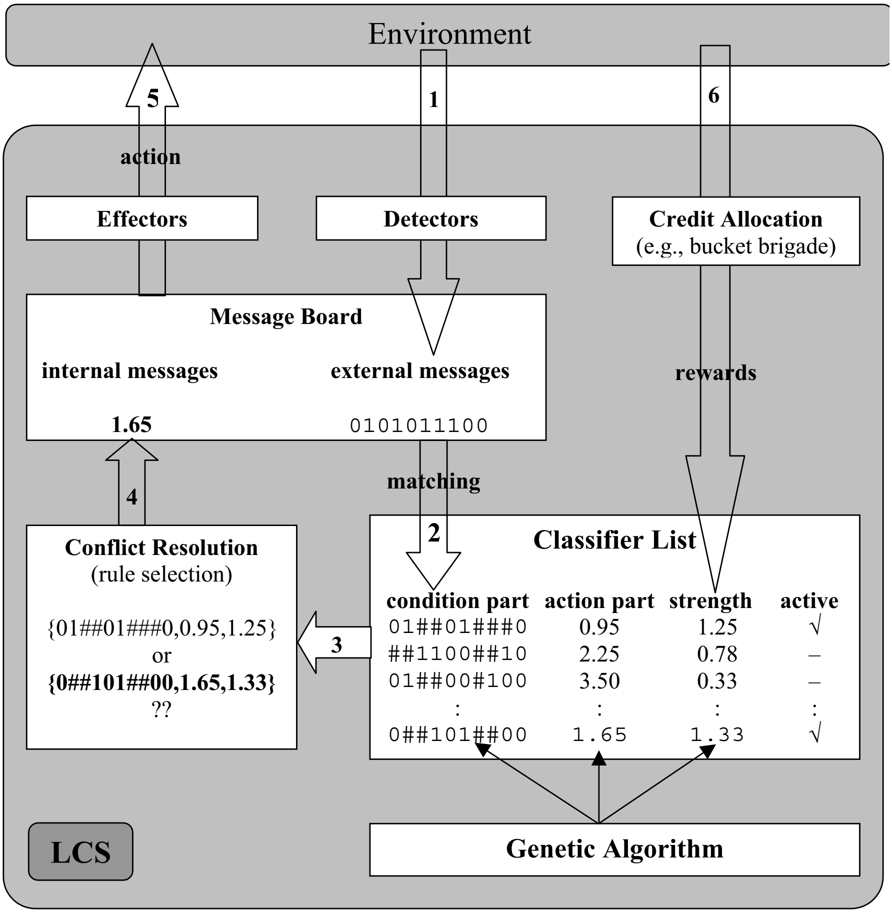


Fig. 4.4. Structure of a Learning Classifier System (LCS).

The state of the environment is sent to the LCS with a set of *detectors*. These detectors post standardized *messages* about the environment to a message board with which a list of *classifiers* are matched. Classifiers are condition/action rules of the form if (condition(s) fulfilled), then (take action).

Because of this particular structure, LCS are also called "condition-action" classifier systems [185].

The environment is often described with a binary representation, yet higher cardinality alphabets are also possible [402]. For instance, the first bit of the message $\{01100\}$ could mean that an asset price is below its long-term average while the second bit indicates that last period's trading volume exceeded a certain threshold level. Similarly, the last three bits are Boolean descriptors of other pre-defined environmental conditions.

For such a binary representation of the environment, the condition part of a classifier is a ternary list $\{0, 1, \#\}$, with # being a wildcard sign which indicates that the particular condition encoded by the bit on that position is ignored. Matching occurs by a comparison of each position between the condition part and the external messages issued by the detectors. For instance, the condition parts $\{\#11\#0\}$ and $\{0\#\#00\}$ match the above message, but not the condition parts $\{1\#\#\#0\}$ or $\{\#\#0\#1\}$. Detailed rules, i.e., those with only a few wildcard signs in their condition part, are called specific. General rules with many wildcard signs, on the other hand, match many different environmental conditions. By favoring more specific classifiers over general ones, LCS can implement a default hierarchy in which specific exemptions override more general situations.[27]

While the rule base of an expert system is formulated such that each possible situation is matched by only one rule, in LCS, several rules usually match and are marked as active. Activated classifiers compete with each other for execution based on their *strength*, which is a function of their past performance. Classifiers that have been proven to be highly successful, for instance, by being profitable, have a greater chance of being activated than less successful classifiers. The chosen classifier posts its message, for instance, a numeric value of how much of a commodity to produce, sell, or buy, to the message board which will then be executed by the *effectors*. *Credit allocation* is an environmental feedback mechanism that rewards profitable classifiers and punishes failing ones by increasing or decreasing their strength.

While credit allocation is a learning mechanism that identifies and rewards good rules from the set of existing classifiers, a GA is an integral part of any LCS to generate new classifiers. The GA is invoked on a

[27] It is also possible, however, to favor general rules over specific rules by punishing each bit other than # with some associated cost. Instead of default hierarchies, one could then model an agent's complexity aversion.

regular basis to create new offspring through selection, crossover, and mutation. In contrast to the canonical GA, which performs a complete generational replacement of the parent population, the offspring in LCS replace only some of the worst classifiers. The best solutions are thus, usually preserved in LCS.

The credit allocation process as described above is not particulary difficult since there are only single classifiers that are directly responsible for past actions. The feedback provided in terms of reward or punishment directly increases or decreases a classifier's strength. Most descriptions of LCS, however, refer to the *bucket brigade algorithm* when explaining the apportionment of credit.

In order to understand the bucket brigade algorithm, it is necessary to note that rules can be *coupled* to act sequentially. A rule is said to be coupled (or chained) to another if its action part generates a message that will potentially activate another rule in the next period. Coupled rule sequences are crucial for handling complex situations, for developing plans and strategies, and for modeling causal relations in the environment. Obviously, stage setting rules that contributed to profitable future actions of other rules should be similarly rewarded than the final rule which receives its reward directly from the environment.

The bucket brigade algorithm does exactly this. It derives its name from an analogy to an old-fashioned fire brigade whose members are handing buckets of water to each other. In the bucket brigade algorithm, every suppliant classifier that activates another classifier is directly paid by the latter. The consuming classifier, i.e., an activated and chosen classifier, pays its suppliant classifiers by 'investing', i.e., distributing its strength to all of its suppliant classifiers. If the action turns out to be successful, its strength value will be replenished by the environmental reward, otherwise not. In this way, the bucket brigade algorithm strengthens verified classifier chains and weakens less successful rule sequences.

More detailed descriptions of LCS and the bucket brigade can be found in [185, 184].

5

Replicating the Stylized Facts of Financial Markets

... to be able to imitate reality is a form of understanding ....
Benoit B. Mandelbrot and Richard L. Hudson [280, p. 19]

5.1 Efficient Markets and the Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) has been the cornerstone of finance for more than thirty years [392]. In 1978, Jensen claimed that no other hypothesis in economics had more empirical support [201]. He also acknowledged, though, that as econometric sophistication increased, more and more inconsistencies arose that could no longer be ignored. The years following this statement saw the ascent of behavioral economics, a new approach to financial economics that further amassed anomalies in apparent contradiction to the EMH. The interpretation of the empirical results, however, is often impeded by the fact that the EMH may be expressed in a number of ways, some of which are not equivalent and are differentiated only by subtle technicalities [245, 84].

5.1.1 Definitions

The EMH is often stated in Fama's words, namely that a "*market in which prices always 'fully reflect' available information is called efficient*" [115, p. 383]. Since only news is not yet available and, by definition, not predictable, the EMH holds that prices should randomly fluctuate in the same way as the news finally materializes. By defining "available information", Fama distinguishes three levels of market efficiency. For the *weak form* version of the EMH, the information set consists exclusively of past price and return information. The *semi-strong* version additionally includes all publicly available information at time t, e.g., balance sheets information, income statements, or announcements of dividend changes. Finally, the *strong form* version of the EMH asserts that all public and private information is fully reflected in market prices. In strong form efficient markets, even insiders would be unable to make an excess profit from their private knowledge.[1]

Fama's original definition, however, neglects the costs of information acquisition. If information acquisition is costly, perfect informational market efficiency is impossible to achieve, as Grossman and Stiglitz show [164]. If markets were perfectly efficient, informed traders could not earn a return on their investment in information acquisition, and markets would eventually collapse.[2] In the sequel to the original article, Fama [116] adopts a weaker, but more realistic definition of market efficiency suggested by Jensen [201]. A market is efficient with respect to information set Φ if it is impossible to make economic profits by trading on the basis of that information set. Economic profits are risk-adjusted returns net of all costs. That is, prices reflect information only to the point where the marginal benefits of acquiring information do not exceed the marginal costs of doing so. Informational efficiency is a benchmark which is impossible to achieve under the presence of information costs. Markets should, rather, be thought of as possessing an equilibrium degree of inefficiency.

When discussing market efficiency, financial economists usually refer to informational efficiency. Yet this begs the question whether informational efficiency implies a Pareto-efficient resource allocation in the economy. Cuthbertson claims that if stock prices always fully reflect fundamental values, then the market will optimally allocate funds among competing firms [84]. Stiglitz [411], on the other hand, seriously challenges this view with a variety of arguments. He shows that the view of efficiency in the finance literature is neither necessary nor sufficient for Pareto optimality in the economy. The recent stock market bubble

[1] In his 1991 article [116], however, Fama relabels the taxonomy of information sets which he developed in 1970. The first category is now called "tests for return predictability". Semi-strong tests of market efficiency are relabeled as event studies, while for strong-form tests, he suggests the term "tests for private information".

[2] A critique and solution to Grossman and Stiglitz's information paradox is given by Helliwg [176]. In an informationally efficient market, investors who do not engage in costly information acquisition derive information from the market price even before any transaction has taken place. In a dynamic setting, however, investors who actively acquire information are able to use it before it is revealed through prices.

in the 1990's with excessive investments in "new economy" firms, followed by significant underinvestment thereafter, is just another indication that the stock market does not always allocate resources efficiently. Since the EMH is perfectly suited to rationalize even "irrationally exuberant" asset prices, widespread belief in it may even obstruct the resource allocation function of capital markets.

A recent refinement of the notion of market efficiency deserves mention at this point though. Paul Samuelson [366] proposed a dictum for the stock market in which he argued that markets exhibit considerable micro efficiency while being macro inefficient as a whole. He considers the EMH to be more applicable for individual stocks than for stock market indices. Jung and Shiller [210] tested this hypothesis by analyzing dividend-price ratios for all 49 U.S. firms that existed in the period from 1926–2001. They do so by running a regression of future multi-year dividend changes on current dividend-price ratios and testing whether the dividend-price ratio predicts these changes. Efficient market theory predicts a negative slope of minus one when plotting dividend growth over dividend-price ratios. Jung and Shiller find a negative coefficient for their regression analysis which supports Samuelson's idea of efficiency on the level of individual stocks. Firms with lower dividend-price ratios did indeed have higher subsequent dividend growth. When constructing a stock market index from the 49 stock and doing the same regression, the regression coefficient becomes positive, which indicates that the EMH is less appropriate for the stock market as a whole.[3]

5.1.2 Random Walks or Martingales?

The idea that security prices fluctuate randomly was formulated as early as 1900 by Louis Bachelier in his doctoral thesis "*Théorie de la spéculation*" [15].[4] In it, he developed the idea that the prices of French government bonds resembled a random walk.[5] A random walk price series could be described by

[3] Samuelson's notion of micro efficiency and macro inefficiency of the stock market adds another twist to the discussion on aggregation issues. Often, the argument is put forth that irrationalities on the micro-level would cancel each other out, thus, causing aggregate behavior that is consistent with the assumption of a fully rational representative agent. Interestingly, the direction in Samuelson's argument is reversed, i.e., the amount of rationality and efficiency is higher on the micro-level.

[4] An English excerpt of Bachelier's dissertation can be found in [82]. An introduction to his life and work is given in [416].

[5] Bachelier, however, did not use the term "random walk". Probably the first time the notion of a random walk emerged was in a written conversation between Karl

$$p_{t+1} = p_t + \epsilon_t \tag{5.1}$$

with $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$, i.e., a sequence of independently and identically distributed random disturbances with zero mean and equal variances.[6] Bachelier's results did not receive any attention until an increasing number of studies in the 1950's, some of which are reprinted in [82], rediscovered that the behavior of various asset prices can be approximated by the random walk model.[7]

The first rigorous arguments explaining why security prices fluctuate randomly were given by Paul Samuelson [365] and Benoit Mandelbrot [277]. They showed that in competitive markets with rational risk-neutral investors, asset prices follow martingales. A stochastic discounted price process p_t would satisfy the martingale property if

$$E_t\left[\tilde{p}_{t+1}|\varPhi_t\right] = p_t \tag{5.2}$$

holds, $\varPhi_t$ being a given information set.[8] The tilde sign ˜ is used to indicate random variables which are not yet realized in period t. According to 5.2, the best predictor of next period's asset price is simply today's price. The martingale model implies that $(\tilde{p}_{t+1} - p_t)$ is a "fair game", i.e.,

$$E_t\left[(\tilde{p}_{t+1} - p_t)|\varPhi_t\right] = 0. \tag{5.3}$$

Equation 5.3 says that increments in value, i.e., price changes adjusted for dividend income, are unpredictable, given the information set $\varPhi_t$. This means that all available information is fully reflected in current

---

Pearson and Lord Rayleigh in 1905 in the journal *Nature* [333, 349]. These three letters can be downloaded from http://www.e-m-h.org/Pear05.pdf.

[6] Fama [115] notes that the terminology in finance is loose at this point. First of all, expected price changes can be non-zero such that we have a "random walk with drift". Secondly, if one period returns are iid, prices will not follow a random walk since the distribution of price changes is dependent on the price level. Third, instead of simple white noise $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$ disturbances, Gaussian white noise $\epsilon_t \sim iidN(0, \sigma_\epsilon^2)$ is often assumed [59]. Besides these variations, three versions of the random walk hypothesis are identified in the finance literature: The "independently and identically distributed returns" version, the "independent returns" version, and the "uncorrelated returns" version [65].

[7] For a more detailed account of the early history of efficient markets, see [115, 245, 100].

[8] It is common in the EMH literature to speak of prices as following martingales. We will adopt this somewhat imprecise but convenient usage by specifying that prices should be understood to include reinvested dividends. Since prices can easily be converted to returns and vice versa, equation 5.2 could also be written as $E_t\left[\tilde{p}_{t+1}|\varPhi_t\right] = p_t\left[1 + E(\tilde{r}_{t+1}|\varPhi_t)\right]$, with $\tilde{r}_{t+1} = (\tilde{p}_{t+1} - p_t)/p_t$ being next period's return of the asset.

prices, and that it is impossible to make excess profit by trading on the basis of Φ_t.

The random walk model of speculative asset prices is only a special case of the more general class of martingale models. Contrary to the random walk case, martingale models do not require the stochastic disturbances ϵ_t to be iid, and are, thus, less restrictive.[9] Many of the stylized facts of financial asset prices such as volatility clustering, fat tails, and other deviations in the return distribution from a normal distribution are, thus, again compatible with the notion of an efficient market.[10] By assuming the less restrictive martingale model, these stylized facts can be seen as a mirror image of the statistical properties of the unobservable news arrival process.

5.1.3 Tests for Market Efficiency

Market efficiency is often equated with the assumption of rational expectations and the martingale property. A litmus test for market efficiency is whether it can be shown that consistently higher risk-adjusted returns than a simple buy-and-hold of the market portfolio can be earned [362]. Roll [359], however, observed that it is extremely difficult to profit from even the most severe violations of market efficiency.

Dependent on the size of the information set, several types of tests for market efficiency have been developed. The first weak-form efficiency tests primarily used past return data [216, 114] before expanding their scope to other publicly available information such as dividend yields, earning-price ratios, or financial accounting data [118]. The semi-strong form of the EMH is often tested with event studies which measure the adjustment speed of asset prices to new information [117]. Event studies average the cumulative performance of stocks

[9] Equation 5.2 can equivalently be rewritten as $p_{t+1} = p_t + \epsilon_t$, which looks identical to the random walk model in equation 5.1. The only difference lies in the more restrictive assumptions for the random shocks ϵ_t in the random walk case. While the random walk model implies that the first four moments of the price or return distributions—mean, variance, skewness , and kurtosis—are independent of Φ_t, the martingale property only states that the mean of the price distribution is independent of the information set.

[10] Violations of the Gaussian price increments assumption discredit the modern asset pricing models or Black and Scholes' options pricing formula [37] more than the EMH, since the former all rely on the standard deviation as a well defined measure of risk. If asset returns were to follow stable Paretian distributions with no well-defined risk measure [276], the risk-return calculations in modern asset pricing models would become obsolete. Even Mandelbrot and Hudson [280] acknowledge that the martingale property is usually not contradicted by empirical evidence.

over a time windows covering a specified number of periods before and after an event. The cumulative abnormal stock returns are obtained by adjusting for market-wide movements in security prices. Ball finds, for instance, that the risk-adjusted abnormal returns after the announcement of an earning announcement are systematically non-zero which would be inconsistent with market efficiency [21].

In one of the first tests for the profitability of private information, Jaffe rejects strong-form market efficiency [196]. When surveying more recent contributions in this field, Fama [116] concludes that by now violations of strong market efficiency are well established, a finding that has been confirmed for the German stock market [380]. Two classic overviews over the different tests on market efficiency are the contributions by Fama [115, 116]. A textbook on informationally efficient capital markets is due to Sapusek [369] who also provides empirical data on market efficiency for the German and Austrian stock market.

Even though the area of weak and semi-strong market efficiency is highly contentious, the basic belief in the EMH is not shaken [100]. This is demonstrated by the fact that deviations from it are still referred to as “anomalies”.

5.2 Stylized Facts of Financial Markets

Much of our knowledge about the stylized facts of financial markets stems from permanent attempts to either prove or disprove the EMH or the classical asset pricing models that build on the EMH. Since it would be hard to define which empirical phenomena qualify as a stylized facts, an exhaustive description is impossible at this point. Instead, priority will be given to those empirical regularities that can be addressed by means of agent-based models of financial markets.

5.2.1 Non-Normal Return Distributions

At the time when the tenets of the EMH were formulated, logarithmic asset returns $r_{\triangle t}(t) = \ln p(t + \triangle t) - \ln p(t)$ were usually thought to be normally distributed. At the same time, however, Mandelbrot [276] presented convincing evidence that this assumption cannot be upheld. In his long-term analysis of cotton prices, he showed that for various choices of $\triangle t$, ranging from one day up to one month, extreme events occur much more often than implied under the Gaussian hypothesis.

Subsequent empirical studies unanimously confirmed this finding of leptokurtosis for virtually all financial asset return distributions, i.e.,

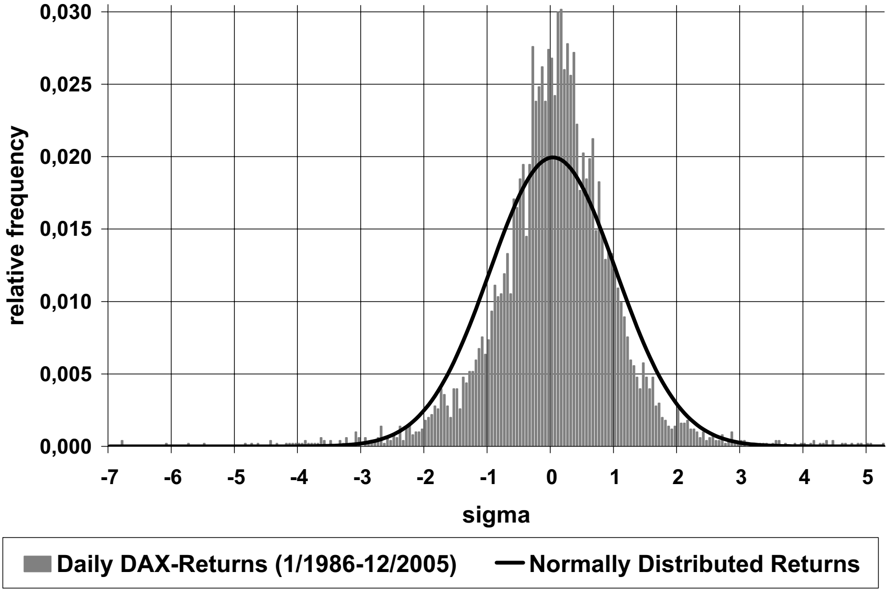


Fig. 5.1. Comparison of daily DAX returns for the period 01/03/1986 – 12/30/2005 with normally distributed returns. The negative 9.47σ-event on October 16, 1989 was cut off. The daily mean return at $\sigma = 0$ equals 0.027%, the standard deviation is 0.014 (implying an annual variance of 23.0%), the skewness is negative with -0.463, and the excess kurtosis is determined as 5.84.

the kurtosis[11] as a distribution's fourth moment

$$\kappa = \frac{1}{N} \sum_{j=1}^{N} \frac{(r_j - \bar{r})^4}{\sigma^4} \tag{5.4}$$

is always larger than three. A comparison of daily DAX returns with normally distributed returns in figure 5.1 illustrates that real financial data have more probability mass in the tails and in the center. In other words, they are fat-tailed and high-peaked. The kurtosis for the daily DAX returns in the analyzed period is 8.7, significantly more than the expected kurtosis of three for a normal distribution. Within the $N = 5,040$ daily observations of DAX returns, three $+5\sigma$, two -5σ,

[11] Kurtosis derives from the Greek word *kyrtos* for "curved" and is a measure of how tall or squat a curve is. A normal distribution is mesokurtotic with a kurtosis of three. Thin-tailed distributions with a kurtosis less than three are called platykurtotic. Excess kurtosis is defined as $\kappa - 3$.

three -6σ, and one -9σ events occurred. Under the Gaussian assumption of asset returns, this would have been extremely unlikely. The occurrence of extreme events in figure 5.1 is also in line with Mandelbrot's observation that the positive tails contain systematically fewer events than the negative tails. This particular feature of empirical return distributions is captured by a distribution's third moment skewness

$$S = \frac{1}{N} \sum_{j=1}^{N} \frac{(r_j - \bar{r})^3}{\sigma^3}. \tag{5.5}$$

A skewness smaller than zero indicates that the negative tail of the distribution has more probability mass than the positive tail.

As a joint test of sample skewness S and sample kurtosis κ, the Jarque-Bera test statistic [197, 30]

$$JB = N \left(\frac{S^2}{6} + \frac{\kappa - 3}{24} \right) \sim \chi^2(2) \tag{5.6}$$

of 7313.4 with a p-value of 0.000 clearly rejects the null hypothesis of normality.[12]

The leptokurtosis of return data is explained by the mixture of distributions hypothesis [74] which asserts that the returns are sampled from a mixture of distributions with different conditional variances. The stable Paretian hypothesis advanced by Mandelbrot, on the other hand, claims that returns distributions are best characterized by a different class of distributions. In addition to being non-Gaussian, Mandelbrot found that the shape of daily, weekly, and monthly return distributions of cotton prices appeared to be similar under different time scales [276]. This invariance is called scaling and it is characterized by power law behavior in the tails of a distribution, i.e., extreme events are distributed according to $f(r) \sim |r|^{-(1+\alpha)}$.[13] The stable Paretian hypothesis states

[12] Other popular tests for normality are, for instance, the Kolmogorov-Smirnov test, the Shapiro-Wilk W-test, and the Anderson-Darling test. For a review of normality tests, see [86].

[13] Power-law scaling was first discovered by Vilfredo Pareto (1848 – 1923) in the distribution of wealth. Pareto found that in a population, the proportion of individuals with wealth W above a certain threshold is determined according to a power law $f(W) \sim CW^{-(1+\alpha)}$. The power law exponent α was estimated by Pareto to be around $-3/2$, $f(W)$ being the probability density function, and C a positive constant. Scaling and power law distributions are also very familiar to physicists because they appear near critical phase transitions from ordered to disordered states. In fact, much of the work on scaling behavior in finance has been done by physicists in the econophysics-movement. For an introduction into econophysics, see [283].

that return distributions belong to a general class of stable Paretian distributions[14] with a characteristic exponent α strictly smaller than two. Paretian distributions generally have no closed analytical form, but can be expressed with their characteristic function $f(q)$, which is the Fourier-transform of the probability distribution, q being the Fourier transformed variable:

$$\ln f(q) = \begin{cases} i\delta q - \gamma|q|^\alpha \left[1 + i\beta \frac{q}{|q|} \tan\left(\frac{\pi}{2}\alpha\right)\right] & \text{if } \alpha \neq 0 \\ i\delta q - \gamma|q| \left[1 + i\beta \frac{2}{\pi} \frac{q}{|q|} \ln|q|\right] & \text{if } \alpha = 1 \end{cases} \tag{5.7}$$

Stable Paretian distributions are described by four parameters: i) a location parameter δ for the mean, ii) a scale parameter $\gamma > 0$, iii) a parameter for asymmetry (skewness) β, and iv) the characteristic exponent α. Special cases of Paretian distributions for which closed-form solutions are known are the Gaussian ($\alpha = 2, \beta = 0$) and the Cauchy ($\alpha = 1, \beta = 0$) distribution.[15] In his 1963 study, Mandelbrot [276] estimated the characteristic exponent α for cotton prices to be around 1.7, similar to what Fama reported for the Dow Jones Industrial Average [113].

While data availability and computational limitations allowed Mandelbrot only to investigate small data sets, the recent analysis of huge high frequency data sets suggests a more complicated scaling behavior of return distributions.[16] Mantegna and Stanley, for instance, noted that stable Lévy-type behavior does not apply to the most extreme parts of return distributions and suggested instead a truncated Lévy process whose tails exponentially decay after some threshold value [282]. Gopikrishnan et al. [158] reported that return distributions slowly converge to Gaussian behavior as $\triangle t$ gets larger. They also note that the scaling properties for small $\triangle t$ break down when they reshuffle their data set, thereby destroying all possible time dependencies in the

[14] Stable Paretian distributions are also called fractal, L-stable, or Lévy distributions. This general class of distributions was discovered by Paul Lévy. They are by construction stable, i.e., invariant under addition. When adding two Lévy random variables, the functional form, and hence, the shape of the distribution, is maintained (stable). For more information on stable Paretian distributions see, for instance, [276, 113, 114, 265, 346, 381].

[15] An implication of the stable Paretian hypothesis is that the second and higher moments do not asymptotically converge to a finite value, i.e., they do not exist [265].

[16] When analyzing rare events, sample size matters a lot. While Mandelbrot analyzed approximately 2,000 data points, modern high frequency data sets consist sometimes of up to $10^6 - 10^7$ data points [282, 158, 159].

returns. Therefore, they conclude that scaling behavior is caused by time dependencies in the returns.

Traditional parametric and non-parametric estimation methods usually provide a poor fit to the extreme tails of return distributions [147]. Instead of trying to fit the complete return distribution, many studies now adopt the methods from extreme value theory and exclusively focus on extreme values rather than the whole data set when estimating the power law exponent. Empirical studies using this methodology now find the characteristic exponent α to be outside the stable Lévy-regime. For the German stock market, Lux [265] finds α in the interval (2,4). Gopikrishnan et al., too, estimate $2 < \alpha < 4$ for the S&P-500 index for high frequency data [158]. On longer time scales, the characteristic exponent for the Hang-Seng and the NIKKEI index was even greater. Thus, while scaling behavior is usually confirmed, most studies agree that the return distributions scale outside the stable Lévy-regime [2, 259, 165].

The exact distribution of financial returns and its causes remains an open question. Various alternative return distributions are suggested, some of which are mentioned in [267]. Since knowledge about the correct probability distribution is crucial for many financial applications such as portfolio selection, derivative pricing, or value-at-risk estimations, research into the appropriate distributional model will continue.

5.2.2 Volatility Clustering of Returns

The first stylized fact of empirical return distributions stated that the frequency of extreme events is larger than under the Gaussian hypothesis. Volatility clustering as the second stylized fact notes that these large returns often tend to emerge in clusters, i.e., large returns tend to be followed by large returns of either sign.[17] This is also known as volatility persistence. The volatility clusters for the daily DAX returns for the period $01/1986 - 12/2005$ are displayed in figure 5.2.

Volatility persistence means that volatility is serially correlated and, therefore, partially predictable. Another way to visualize this is to plot

[17] Fat tails and clustered volatility are closely connected and not unique to financial time series. In a study of hydrology, Mandelbrot coined the term “Noah-effect” for the first phenomenon, referring to extreme precipitation in the biblical account of Noah [281]. The phenomenon of clustered volatility, i.e., persisting high or low levels in rivers, was labeled by them as the “Joseph-effect”, drawn from the familiar biblical story of Joseph interpreting the Pharaoh’s dream to mean seven years of feast followed by seven years of famine.

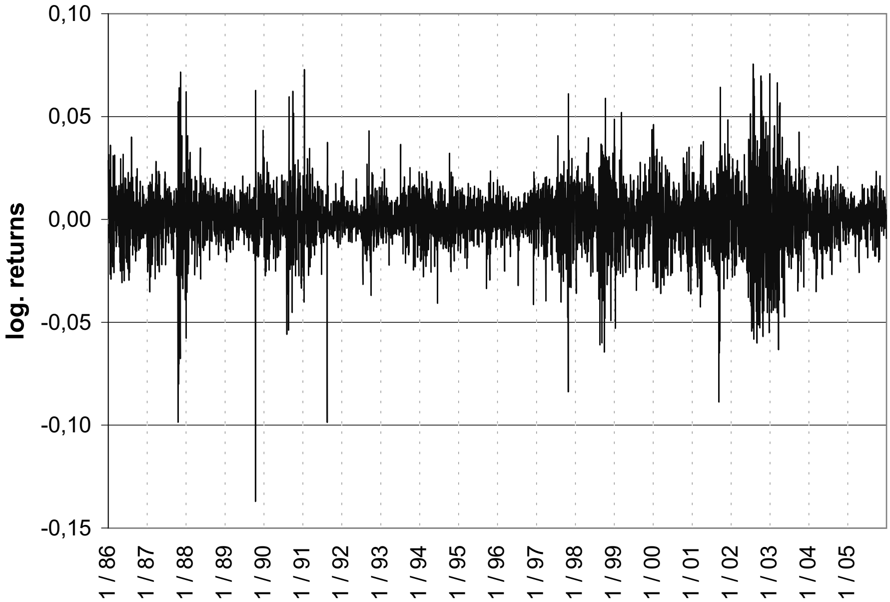


Fig. 5.2. Daily DAX-returns for the period 01/03/1986 – 12/30/2005.

the autocorrelation of absolute and squared returns, which both measure volatility. While the autocorrelation coefficients of raw returns fluctuate around zero, those for absolute and squared returns are usually much higher and indicate long-memory processes.[18] The autocorrelation coefficients for the first two hundred lags of raw, absolute, and squared daily DAX returns are plotted in figure 5.3 and are representative of most other financial return series.

Theoretical models that explain the possible causes of volatility persistence are rare. One of the few exceptions is the preference-based asset pricing model by McQueen and Vorkink [293] which generates low-frequency conditional volatility on the basis of agents' state-dependent risk-aversion and sensitivity to news. There are, however, many structural models of return generating processes that abstract from specific economic relationships. The existence of clustered volatility suggests abandoning the assumption of linear return generating processes such as Autoregressive (AR), Moving Average (MA), or ARMA processes and focusing on non-linear processes.

[18] While most older studies agreed on insignificant autocorrelation for raw returns [114], Lo [255] provides evidence of time-varying small, but significant, serial price correlation.

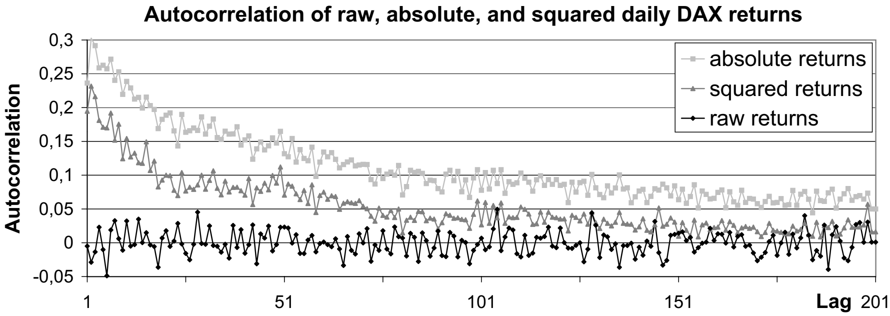


Fig. 5.3. The first 200 autocorrelations for the raw, squared, and absolute daily DAX returns (01/03/1986 – 12/30/2005).

Threshold Autoregressive (TAR) models [424] are additive non-linear return generating processes since they jump between several linear stochastic processes based on certain threshold levels. This class of models is thus locally linear, yet non-linear over the whole state space. Another more popular approach to model serial correlation of volatility was proposed by Engle [110]. He introduced the Autoregressive Conditionally Heteroskedastic ARCH(q) model

$$r_t = \mu_t + \sigma_t \epsilon_t \tag{5.8}$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \epsilon_{t-i}^2, \tag{5.9}$$

in which the variance of an error ϵ_t is conditional on lagged variances. The error terms ϵ_t are normally distributed $\epsilon_t \sim N(0,1)$, μ_t is the expected return in period t, and as parameter restrictions $a_0 > 0$ and $a_i \geq 0$ for all $i = 1, \ldots, q$ apply. For a stationary return processes like above, $\sum_{i=1}^{q} a_i < 1$ must hold. In this case, the unconditional variance $\overline{\sigma}^2$ exists and is a constant

$$\overline{\sigma}^2 = \frac{a_0}{1 - \sum_{i=1}^{q} a_i}. \tag{5.10}$$

Using equations 5.9 and 5.10, the conditional variance for an ARCH(1)-process

$$\sigma_t^2 = \overline{\sigma}^2 + a_1 \big(\epsilon_{t-1}^2 - \overline{\sigma}^2\big) \tag{5.11}$$

is mean-reverting around the unconditional variance $\overline{\sigma}^2$. Equation 5.11 also illustrates that unexpected large changes will result in a higher conditional volatility, i.e., clustered volatility. Simple ARCH(1) models are, however, unable to fully capture volatility persistence, and higher order ARCH models are computationally hard to estimate. Bollerslev [45] thus suggested the Generalized ARCH GARCH(p,q) model in which equation 5.9 for the conditional variance is replaced by

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2. \tag{5.12}$$

For σ_t^2 to be non-negative, the restrictions $a_0 > 0,\ a_i, b_j \geq 0$ for all $i, j = 1, \ldots, q, p$ must apply, and for stationary GARCH(p,q) processes

$$\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j < 1 \tag{5.13}$$

must hold. In contrast to ARCH(1) models, low-order models in the GARCH class often describe return distributions with fat-tail behavior and volatility persistence reasonably well. Fitting a GARCH(1,1) model for the daily DAX -returns yields

$$\sigma_t^2 = 5.33 \cdot 10^{-6} + 0.119\epsilon_{t-1} + 0.858\sigma_{t-1}^2$$

for the conditional variance.[19] Pagan, however, points out that the residuals from estimated GARCH models are still non-Gaussian and leptokurtotic [328]. In addition, Mikosch and Stărică [302] discuss some theoretical properties of GARCH(p,q) processes that are not in line with real financial data. For instance, stochastic long-term memory processes exhibit hyperbolic decays in their autocorrelation functions (ACF), yet the decline in the theoretical ACF for all GARCH(p,q) processes is exponential, thus, indicating only short-memory processes.[20] Mikosch and Stărică have also recently questioned the validity of GARCH models to describe stochastic volatility and argued that long-range dependence effects might be only spurious. They detected that structural changes, for instance, due to business cycles, cause a violation of the stationarity assumption by shifting unconditional variances [304, 305].

Various other refinements for the conditional variance have been suggested. In asymmetric GARCH models such as the **E**xponential

---

[19] All parameter estimates are highly significant and the stationarity condition is fulfilled.

[20] Note, however, that Fractionally Integrated ARCH (FIGARCH) models [18] are able to model long-run dependence in stochastic volatility. For a technical distinction between short- and long-term dependence, see [303].

GARCH [318], volatility depends not only on the size, but also on the sign of past shocks. Asymmetric GARCH models thus allow incorporating the *leverage effect*, i.e., the empirically observed asymmetric reactions of market volatility to positive and negative shocks. For reviews on GARCH time series models, see [46, 251]. An empirical comparison of different volatility forecast models can be found in [169, 168].

The continuing dependence on Gaussian innovations with time-varying variances by GARCH type models is challenged by Mandelbrot and Hudson [280]. As an alternative, they suggest the Multifractal Model of Asset Returns (MMAR) [279]. The MMAR combines fractional Brownian motion with the concept of multifractal trading time, i.e., a random distortion of clock time.[21] Rather than being an autoregressive process, the MMAR is a cascade model which generates price series by randomly splitting up a complete interval in smaller sub-intervals, thereby increasingly roughening the price series by successive interpolations. In the limit, the generated price series reproduce the main features of financial prices: fat tails in the return distributions, scale consistency, and long memory in volatility. In addition, the MMAR can generate either martingales or long memory behavior in log prices. First empirical investigations of the MMAR hypothesis show promising results [63, 314, 44], yet the assumption of multifractal time in an economic context is still being questioned [307].

5.2.3 High and Persistent Trading Volume

In the world of a representative agent, fully rational investors who share the same information and expectations would obviously never trade with each other. More surprising is the no-trade theorem by Milgrom and Stokey who showed that information asymmetry alone cannot induce trading [306]. Given the same prior beliefs and an ex-ante Pareto-efficient allocation, the possibility of asymmetric information arrival would make a risk-averse rational trader believe that a potential trading partner is better informed than himself. Under these circumstances, no one would ever engage in any trading activity.

To reconcile the spate of trading activity in real financial markets with no-trade theorems, several solutions have been suggested. For Black [36], the missing ingredient is noise trading, i.e., he believes that investors often trade on noise as if it were information, for instance, by following certain technical trading rules.[22] As long as these trading

[21] More on the concept of multifractal time can be found in [278]. A similar concept of time relativity in financial markets is espoused by Montassér [311].

[22] The noise-trader approach to finance is also advanced in [96, 393, 98].

strategies are uncorrelated, high trading volumes imply that irrational trades cancel each other out and asset prices are close to fundamental values. However, Shleifer [392] remarks that the latter hinges on the assumption of uncorrelated trading strategies, something that is at times violated in a world with mass media.[23] Similar to noise traders, liquidity traders or life cycle investors may act without regard to expected profits. Other theoretical explanations for ongoing trading activities are sought in dropping the assumption of common prior beliefs [213, 428], or in dropping the assumption of efficient and complete markets by integrating various insurance motives [360].

Besides being positive, transaction volume exhibits many other interesting features. Lobato and Velasco, for instance, have shown it to be a persistent long memory process [257]. Furthermore, empirical data show that trading volume is characterized by a positive autocorrelation coefficient which only slowly decays, and that a positive cross-correlation between volatility of returns and trading volume exists [417, 136, 328]. This raises the question about the informational content of volume data and whether it may be exploited for stock price prediction.

According to the mixture of distributions hypothesis [74], return volatility and trading volume are jointly dependent on a latent driving process. Theoretical support and empirical evidence for Clark's thesis that the conditional variance of log prices is a function of transaction volume is provided in [111]. Other studies investigating the relationship between transaction volume and asset prices are [417, 143, 57, 311]. A survey of the relationship between price changes and trading volume can be found in [214].

5.2.4 Existence of Technical Trading

The idea that asset prices and trading volumes tend to form certain geometric patterns that can be profitably exploited by forecasting future price movements is known as technical trading. Charting, as it is sometimes called, had been popular even before the disclosure of financial balance sheet information sparked the development of fundamental analysis as its main competitor. Its origins can be found in the works of Charles H. Dow (1851–1902) and William P. Hamilton (1867–1929), who published a series of 252 editorials in the Wall Street Journal at

[23] Coordination on similar trading strategies because of media-induced herding behavior would constitute a strategic complementarity [125] which can lead to large deviations from the outcome predicted by rational models.

the turn of the century. These editorials contained the basic tenets of what became later known as the "Dow Theory" [320]. Another school of technical analysis, Elliott-wave theory, claims that market prices follow several superimposed market trends of different degrees, all of which are rooted in cyclical investor psychology [107, 108, 342].

Since technical analysis is based on the assumption that markets exhibit not even the lowest degree of market efficiency, it is anathema to many academics. Malkiel [274], for instance, compares the scientific merits of chart reading to those of alchemy.[24] Jegadeesh [199], too, points out that academic concerns against technical analysis run deeper than simple "linguistic barriers" [256]. He identifies its weak foundations, i.e., the missing plausible explanations why one should expect market patterns to repeat, as the main culprit.

The Use of Technical Trading in Financial Markets

The academic studies that deal with the extent to which technical trading is used in real markets mainly investigate the foreign exchange market. Frankel and Froot, for instance, attribute the speculative bubble in the period from 1981–85 to a shift from fundamental trading strategies towards technical ones [138]. Questionnaire evidence from forex traders about their use of technical analysis was gathered by Allen and Taylor [7, 419]. Both studies found that the relative importance of charting strategies diminishes as the investment horizons increase. For horizons ranging from intra-day to one week, up to 90% of the respondents relied more on technical than fundamental analysis when forming exchange rate expectations, a finding that is reversed for time horizons exceeding one year. It is also reported that around 2% of traders appeared to never use fundamental analysis, independent of their investment horizons. Menkhoff confirms these results for a different place (Germany) and a different time and concludes "*that technical*

[24] According to Popper's demarcation criterion citePopper1959, however, technical analysis principally follows a scientific methodology. It produces falsifiable propositions that can be refuted by empirical observations. Given the questionable track record of many technical trading systems, their survival surprises many EMH adherents. One possible answer will be given at the end of chapter 9. Another plausible answer to this mystery is immunization strategies. A website dedicated to the Dow Theory, for instance, states the following: "*To define Dow Theory one must be open to the concept that there are specific rules that must be accepted without question, while at the same time being ever aware that the theory is dynamic in its evolution*" (http://www.dowtheoryproject.com/define.php). It is hard too conceive of a more blatantly unscientific statement than this.

currency analysis should be interpreted not as either a marginal phenomenon or representing secondary information or a self-eliminating or second-best strategy, but possibly as a kind of self-fulfilling prophecy" [298, p. 307].[25] Even though academic studies investigating the importance of technical trading in the stock market do not abound, anecdotal evidence and the mere existence of countless books and newsletters on technical stock market trading point to a widespread use, too. Questionnaire evidence gathered by Shiller following the stock market crash in October 1987 confirms that many traders were influenced by technical analysis considerations [390].

The Profitability of Technical Trading Rules

The main body of academic research on technical analysis focuses on its profitability. While the sixties and seventies were marked by a widespread consensus among academics that technical trading net of transaction costs did not generate systematic returns greater than a simple buy-and-hold strategy [115, 202], the "intellectual dominance" of the EMH had become less universal by the beginning of the new millennium [275].

Aside from reported weak form inefficiencies for obvious candidates such as Latin American or Asian emerging stock markets[26] [348, 309, 387, 441], return predictability has been found even for Western European equity markets, for instance, in [300] for the Italian stock market, and in [132] for the Spanish stock market. For the US, Brock et al. [56] tested two popular technical trading strategies, the moving-average oscillator and the trading-range break, on daily Dow Jones Industrial Average data from 1897 to 1986. The generated buy (sell) signals were followed by returns that were generally higher (lower) than "normal" returns, indicating that these technical rules have predictive power. Kwon and Kish extended the analysis of Brock et al. by including volume data and supported their basic conclusion [229, 228]. They did find, however, a gradual weakening in profit potential, implying that the market became more efficient for the period under considera-

[25] Other empirical studies based on questionnaire evidence in the foreign exchange market are [263, 299, 71, 72, 325, 1].

[26] The literature on technical trading profitability in the foreign exchange market is even more voluminous than for stock markets. Many studies have found profitable trading rules around central bank interventions [237, 317, 363]. Because of these interventions, foreign exchange markets are likely to be governed by other rules than equity markets and will thus be ignored in the discussion of profitable technical trading rules.

tion. Ready [350], however, considers the apparent profitability of the technical trading rules used by Brock et al. to be a spurious result of data snooping.

Instead of using ad hoc specifications of technical trading rules, other authors ask whether it is possible to derive "optimal" trading rules by means of artificial intelligence techniques. Allen and Karjalainen used genetic programming to derive new technical trading rules from building blocks—certain moving averages and maxima and minima of past prices—and evaluated them against a buy-and-hold strategy [5]. Using daily S&P 500 index data from 1928–1995, they found little evidence of economically significant technical trading rules. Even though the trading rules were able to detect periods with overall positive or negative returns, the economic profits minus the transaction costs were not consistently higher than a simple buy-and-hold strategy. Allen and Karjalainen attributed the limited forecasting ability to positive low-order serial correlation in stock index returns, since the introduction of a one-day delay to trading signals caused the positive returns to vanish.

Jasic and Wood [198] derive neural network predictions for daily returns of the S&P 500, DAX, TOPIX and FTSE stock market indices. Unlike Allen and Karjalainen, they find strong evidence of high and consistent predictability even in the presence of plausible transaction cost assumptions. Other studies that use neural networks on stock market data are [127, 443]. Another genetic programming approach is by Kaboudan [211]. In summarizing the current discussion about the profitability of technical trading rules, one can only conclude that the debate is far from being settled.

The Effect of Technical Trading on Asset Prices

More theoretical in nature are studies which focus on the effect that the use of technical trading rules may have on the aggregate market outcome. The surge of articles in the field of behavioral finance may be seen partly as a response to the failure of explaining asset price based on pure fundamentals.

Price-to-price feedback models, as discussed in [391], are closely related to trend-following trading strategies. As asset prices go up, the success of some investors attracts the attention of other investors, thereby fueling the expectations of future price increases and investor demand. As long as the feedback mechanism is uninterrupted, unreasonably high asset prices can be sustained over extended periods of time, yet eventually, the bubble has to burst. However, the causal direction, whether pre-existing technical trading strategies initiate trends as precursors of boom and crash periods or vice versa, is not always clear. The same can be said for momentum or over- and under-reaction theories of the stock market [200, 24, 88, 192]. Overreaction, i.e., the tendency of individuals to overweight recent information and to underweight prior data, implies that extreme stock price movements will be followed by subsequent adjustments. DeBondt and Thaler [95] remark that this hypothesis implies a violation of weak-form market efficiency. Certain technical trading rules should, therefore, be able to exploit the observed patterns as was shown in [200].

Multi-agent systems often address the effect of technical trading rules on asset prices by modeling the population dynamics of different trader groups, usually depicted as fundamentalists and chartists. Instead of two groups, agents may endogenously switch between these two groups, thus, influencing the importance of certain trading strategies. Lux [266], for instance, developed a multi-agent model of the foreign exchange market in which fundamentalists and chartists interact. Chartists have either a "bullish" or "bearish" attitude and observe both the behavior of their competitors as well as price movements. The transition probabilities between the two chartist subgroups reflect mimetic contagion and actual price movements, while the transition probabilities from and to the fundamentalist group are a function of the profitability of the existing trading strategies. As the ratio of chartists to fundamentalists constantly changes, the simulated model exhibits optimistic and pessimistic waves. More interestingly, though, are the statistical properties of the simulated price series. They reproduce the stylized fact of leptokurtotic return distributions which become less pronounced when returns are calculated using increasing time intervals.

A similar population dynamic is adopted by Goldbaum [152]. He varies an intensity of choice parameter with which agents switch between a fundamental and a technical trading rule as a result of perceived differences in expected profits. Goldbaum finds that high values of this parameter lead to large population shifts. Asset price movements reflect the market impact of technical traders rather than fundamental values, and an overall pattern of repeating price bubbles and collapses emerges. Similar multi-agent systems that try to explain asset price behavior as a result of the population dynamics are described in [264, 151, 269, 73, 121].

5.3 Alternative Market Hypotheses

Many of the stylized facts of financial markets discussed are in stark contradiction to the random walk version of the EMH. As a result, various other market hypotheses have been postulated, some of which will be briefly discussed in the following section.

5.3.1 The Fractal Market Hypothesis

An alternative market hypothesis to the EMH was proposed by Peters [335, 336]. The key ingredient of his Fractal Market Hypothesis (FMH) is the assumption that markets consist of traders with different investment horizons. Depending on the specific proportion of these investor groups, a market can either be in a stable or an unstable state. The FMH asserts that the market is stable when investors cover a large number of investment horizons. Long-term investors stabilize the market by offering liquidity to short term investors. If an event renders the validity of long-term fundamental information questionable, long-term investors either stop participating in the market or become short term investors themselves. As the overall investment horizon of the market becomes more uniform, the market approaches its unstable state. The assumption of different investment horizons implies that, contrary to the EMH, prices do not always reflect all available information. Short term investors, for instance, may value information differently than long-term investors. In times with more uniform investment horizons, information may thus only be partially reflected in asset prices.

The FMH derives its name from the prevalent fractal structure when the market is in its stable state. Because of the differing investment horizons, a market does not have a characteristic time scale. That is, when plotting the distributions of daily, weekly, or monthly returns, they all look similar. This feature of self-similarity on different (time) scales is one of the defining properties of simple fractals.[27]

The observed fat tails in the return distributions are caused by extreme price variations in the market's unstable state. Since high peaks and fat tails are not compatible with the assumption of Gaussian return distributions, Mandelbrot proposed the more general class of stable Paretian distributions for stock returns [276]. Within this class, the normal distribution is only a special case, i.e., when the characteristic

[27] Financial market charts are, however, better characterized as being self-affine than self-similar. Note that a random walk is also self-similar, but it is not fractal since its fractal dimension is an integer (and not fractional). For geometric fractals, self-similarity and scaling are spatial and not temporal properties.

exponent α, a measure of the degree of tail fatness, equals two. The stable Paretian hypothesis of stock returns asserts that α is strictly less than two. Empirically, the characteristic exponent α was, indeed, almost always found to be in the interval (1, 2). Besides fat-tailedness, this implies the existence of a well-defined mean, but second or higher moments do not exist [265].

Peters points out that the inverse of the characteristic exponent α is equal to the Hurst exponent H, which is a measure of trend persistence [336]. For Hurst exponents $0.5 < H < 1$, which correspond to $\alpha \in (1, 2)$, time series are said to be persistent or trend-reinforcing. Furthermore, by applying rescaled range (R/S) analysis [87], Peters detects periodic or nonperiodic components (cycles) in the returns of the Dow Jones Industrial Average.

5.3.2 The Coherent Market Hypothesis

A second market hypothesis with connection to chaos theory was proposed by Vaga [426]. Like the FMH, the Coherent Market Hypothesis (CMH) is based on the premise that markets can shift between stable and unstable regimes. The CMH applies the well-known Ising model of ferromagnetism to the stock market. Instead of modeling the formation of clusters with the same magnetic orientation, the Ising model in a social context is reinterpreted to describe group formation and group behavior. The two most important parameters are one for market sentiment or crowd behavior, and another for fundamental bias. Based on the particular combination of these two parameters, the market could be in one of four possible states. To be in the stable random walk state—the EMH case—, the crowding parameter must be below a certain threshold level, and aggregate fundamental views should not have a bullish or bearish bias. As the crowd parameter increases above a critical threshold, the model becomes instable, and the probability distribution of stock market fluctuations becomes bimodal. If the fundamental bias is not very pronounced, the stock market is in its chaotic state where large swings can be triggered by small news. In periods with considerable crowd behavior and strong fundamental bias, the stock markets are in either the coherent bull or coherent bear state. Vaga defines coherent behavior as a state of macroscopic order in a complex system, i.e., when a large number of freely interacting parts temporarily align with each other.[28]

[28] Obviously, this perception of coherent market behavior is close to emergent market behavior. The latter, however, is more general since it does not exclusively refer to correlated behavior.

Vaga claims that predictive powers are greatest in coherent market periods. If investors miss being fully invested in these periods, they will underperform the market in the long run. Unlike the FMH, Vaga constructed the CMH to assist in investment decisions through better market timing. He describes technical signals which presumably indicate the transition between different market states, yet his own investment record compared to a buy-and-hold strategy remains ambiguous. This, however, does not disprove the CMH since the transition signals might have been inadequate. Steiner and Wittkemper [410], for instance, implement the CMH in a portfolio optimization model and generate the transition signals through artificial neural networks. A simulation with out-of-sample data resulted in consistently higher positive annual returns compared to the market portfolio. The EMH allows for consistently higher profits only as a fair compensation for risk-taking. Yet the outperformance of Steiner and Wittkemper's model was achieved with only 41% of the risk of a buy-and-hold strategy of the market portfolio.

5.3.3 The Adaptive Market Hypothesis

The Adaptive Market Hypothesis (AMH) is a new behavioral market hypothesis that was recently suggested by Lo [255]. He considers the AMH as a new version of the classical EMH, yet derived from evolutionary principles. Many of the deviations from economic rationality that have been pointed out by the behavioral finance literature—overconfidence, overreaction, loss aversion, herding, hyperbolic discounting, and other behavioral biases—are in Lo's view consistent with an evolutionary perspective on individuals adapting to a changing environment via simple heuristics. From an evolutionary perspective, a financial market can be seen as a co-evolving ecology of trading strategies. The strategies would correspond to biological species, and the total capital commanded by a strategy is analogous to the population of that species. New strategies are constantly created, thereby changing the profitability of pre-existing strategies, in some cases even replacing them or driving them to become extinct [122].

In such a dynamically changing world, it should come as no surprise that the heuristics pertaining to an outdated environment may be ill-advised in the context of a new environment. Those heuristics might seem irrational, but labeling them as maladapted is more appropriate. Lo views the classical EMH as a steady state situation with fixed environmental conditions, whereas the AMH describes constantly adapting and co-evolving groups of market participants or strategies. At the present stage, the AMH is more of a qualitative concept that still needs to be operationalized. Nonetheless, Lo derives a number of concrete implications.

First, the relationship between risk and returns is unlikely to be stable over time. As a result, the equity risk premium and aggregate risk preference are not immutable constants, but time-varying and possibly path-dependent. Natural selection affects the relative importance of different participating investor types who may have differing degrees of risk aversion. The recent burst of the technology bubble probably led to a vastly different population of active investors whose preferences might have changed, too.

A second implication of the AMH is the occasional existence of profitable arbitrage opportunities. As they are exploited, they disappear. An example provided by Farmer [120] demonstrates how slowly the achievement of market efficiency may come about.[29] Figure 5.4 shows the gradual decline in correlation between a model generated signal and the future market movement. During the 23-year period that is shown, the quality of the signal has declined, but not vanished.

Fig. 5.4. Decline in magnitude of a market inefficiency. Correlation between a trading signal and the following forward return of a specific US stock market segment over time. The signal was created by a model used by Prediction Company for actual trading purposes. Source: [120, p. 65].

Lo also points out that a changing environment leads to the constant creation of new profit opportunities. Figure 5.5, which again is taken

[29] Farmer is a co-founder of Prediction Company, a Santa Fe-based firm that develops and markets financial forecasting systems. A lively and entertaining account of Prediction Company and its two founders, John Doyne Farmer and Norman Packard, is given in [26]. Farmer is a trained physicist whose views of finance were shaped by conversations with traders and academics. He claimed that Prediction Company made highly significant profits which should have been impossible according to the EMH [119].

from [120], may be interpreted as an example of a new inefficiency on the market.

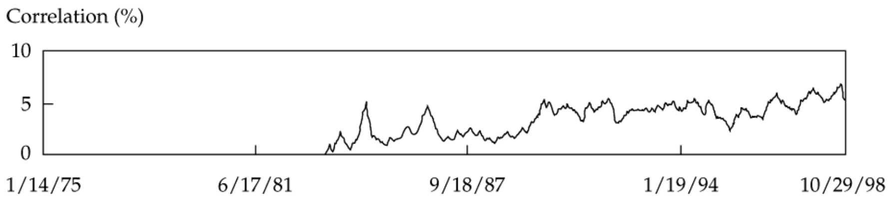


Fig. 5.5. Emergence of a new market inefficiency. Correlation between a trading signal and the following forward return of a specific US stock market segment over time. The signal was created by a model used by Prediction Company for actual trading purposes. Source: [120, p. 65].

Over a longer time horizon, the emergence and disappearance of arbitrage opportunities should result in a cyclical behavior of market efficiency. Lo illustrates this conclusion by plotting the first-order autocorrelation coefficients for monthly returns of the S&P Composite Index over a 132-year period. They are at time zero, but fluctuate wildly with peaks up to 50%. According to the random walk hypothesis, these serial correlations should have been zero all the time. Finally, Lo concludes that the key to survival in a changing environment is innovation. And while profit or utility maximization may be important aspects in financial markets, survival is the main driving force in market evolution, the only objective that really matters to market participants.

While the label "Adaptive Market Hypothesis" seems new, the ideas expressed by it have been around for a while. Chen and Yeh [69], for instance, explore the possibility whether the EMH and the Rational Expectation Hypothesis can be explained as emergent properties in an artificial stock market consisting of co-evolving heterogeneous agents. The importance of financial innovation in a changing environment is also emphasized by Markose et al. who invoked the Red Queen principle[30] to explain market efficiency as an emergent behavior of co-evolving agents [287]. The Red Queen Principle is known in evolutionary biology and describes the idea that a species needs continuing development

[30] The Red Queen Principle is named after a famous passage in Lewis Carol's book "Alice Through the Looking Glass". *"'Well, in our country' said Alice, still panting a little 'you'd generally get to somewhere else if you ran very fast for a long time as we've been doing.' 'A slow kind of country!' said the Red Queen. 'Now here, you see, it takes all the running you can do, to be in the same place'".*

just in order to maintain its fitness relative to other species with which it is co-evolving [427]. It is often referred to as an evolutionary arms race between competing species. One of the first to study Red Queen effects in economics was Robson [357], and Markose further discusses Red Queen examples in an economic context [286].

5.3.4 The Interacting-Agent Hypothesis

The EMH essentially assumes that the return distribution simply mirrors the non-observable news arrival process. As researchers acknowledge more and more that stock returns were not normally distributed, the random walk hypothesis was replaced with the less demanding martingale model. The random shocks no longer needed to be independently distributed, thus, allowing for deviations from normality in the higher moments of the return distribution and for volatility clustering.

The Interacting Agent Hypothesis (IAH) by Lux and Marchesi [268, 269], on the other hand, assumes that the news arrival process is again Gaussian distributed. In their agent-based simulation of a stock market, however, Lux and Marchesi find that the time series of returns exhibits both fat tails and volatility dependence.[31] Thus, they conclude that these statistical properties appear as emergent phenomena as a consequence of the market interactions of heterogeneous agents. It is the trading process itself that transforms and magnifies exogenous news, modeled as white noise, into realistic fat-tailed return distributions with clustered volatility. This feature has been verified by Chen at al. [67]. Many other agent-based simulations of financial markets exhibit the same emergent characteristics of fat-tailed return distributions with clustered volatility, some of which will be briefly introduced in the following sections. The IAH is thus a serious competitor to the EMH in explaining how the observed return characteristics come about, yet an empirical test to discriminate between the two hypotheses is hard to conceive of. The Heterogeneous Market Hypothesis (HMH) proposed by Hommes [189] rests on a argument similar to the IAH.

5.4 Agent-Based Computational Models of Financial Markets

Even though the field of heterogeneous agent models in economics and finance is still relatively young, categorizing the different models has already become quite difficult. The first problem arises in deciding which

[31] Their artificial stock market model is presented in section 5.4.3 on page 83.

type of models to include or not. A pragmatic approach is to distinguish between those models that can be analytically handled and those that cannot and need to be studied by means of computer simulations. The first group of analytical heterogeneous agent models are surveyed in Hommes [190], while LeBaron [243] focuses on agent-based computational models.[32] I will follow this distinction and concentrate on computational models.

The second problem arises in defining the criteria according to how these models should be categorized. One could easily arrange them according to which stylized facts of financial markets they replicate, to what kind of learning or market clearing mechanism they employ, or to the degree of agent-heterogeneity.[33] As with probably any categorization scheme, there will be overlap and possibly models that do not fit in the proposed scheme.

The guiding criterion in the following selection of artificial financial markets is the principal of minimal rationality. I shall assume that agents with random demands exhibit lower degrees of rationality than those who either learn their optimal demands over the course of time or derive them by optimizing a given utility function.

5.4.1 Allocative Efficiency with Zero-Intelligence Traders

A natural starting point when discussing various artificial markets from the viewpoint of minimal rationality are the "zero-intelligence" (ZI) traders by Gode and Sunder [150].[34] In order to determine how much of the performance in terms of allocative efficiency can be attributed to market structure and how much to human intelligence, Gode and Sunder designed an experiment in which the results of a control group of human traders were compared with two types types of ZI machine traders.

[32] Other surveys on agent-based computational models in finance can be found in [236, 238, 248].

[33] LeBaron [239] discusses all the different design issues when building agent-based financial markets. These building blocks, for instance, agents, trading and market clearing mechanism, and the kind of available securities, could all certainly serve as classification criteria.

[34] To my knowledge, Gode and Sunder are currently the only ones to use the term "minimal rationality economics" as keywords and when referring to their research interests. Minimally rational traders are also studied by Farmer et al [123]. Related to the zero-intelligence approach is Schweitzer's concept of Brownian agents [384, 386, 385]. Brownian agents possess intermediate complexity and are minimalistic in the sense that they act on the simplest set of rules without deliberative actions.

An arbitrary commodity was traded in a double auction market consisting of six buyers and six sellers. If a buyer's bid (quantity and price) was matched or crossed by a seller's offer, a trade was executed between the two. In the case of crossing bids and offers, the transaction price p_i for unit i of the commodity equaled the earlier of the two. For each buyer, individual demand functions were implemented by defining a sequence of redemption values v_i, $i = 1, 2, \ldots, n$ for each unit i of the commodity. Since the redemption values were private information, the aggregate demand function was unknown to buyers and sellers. Individual supply functions for sellers were similarly defined by imposing a private sequence of production costs c_i. Hence, profits for buyers equaled $v_i - p_i$, and for sellers $p_i - c_i$.

The ZI traders were simple computer programs that randomly generated either bids or offers. Since the ZI traders had no utility functions, there was no need to memorize or learn anything. It seems appropriate then to label them as having zero intelligence. The ZI traders come in two variations. Budget-constrained ZI-C buyers were not allowed to bid above their redemption value while ZI-C sellers were not allowed to sell below their production costs c_i. Hence, the random bids where independently, identically, and uniformly distributed over the interval $[1,\ v_i]$, while the random offers were restricted to the interval $[c_i,\ 200]$. ZI-U traders were not required to trade within a budget constraint. The bids and offers for these unconstrained ZI-U traders were then uniformly distributed across the entire range of possible trading prices [1, 200].

Gode and Sunder designed the experiment in such a way as to separate the performance differences that are due to market discipline from those that are a result of human intelligence and profit maximization.

It can be seen from figure 5.6 that the transaction prices for ZI-C traders are much less volatile than for ZI-U traders and converge to the equilibrium price (depicted by a horizontal line). Since both trader types possess no intelligence, this difference is attributable to market discipline imposed by the budget constraint on ZI-C traders. The price series in markets with human traders rapidly converges to the equilibrium price and exhibits almost no volatility. The difference in markets with ZI-C traders reflects the contribution of human rationality and profit maximization on the outcome.

The allocative efficiency of each market is assessed by dividing the total profits earned by the maximum attainable profit, i.e., the sum of consumer and producer surplus [404]. Compared to humans and budget-constrained ZI traders, the efficiency of ZI-U markets is always

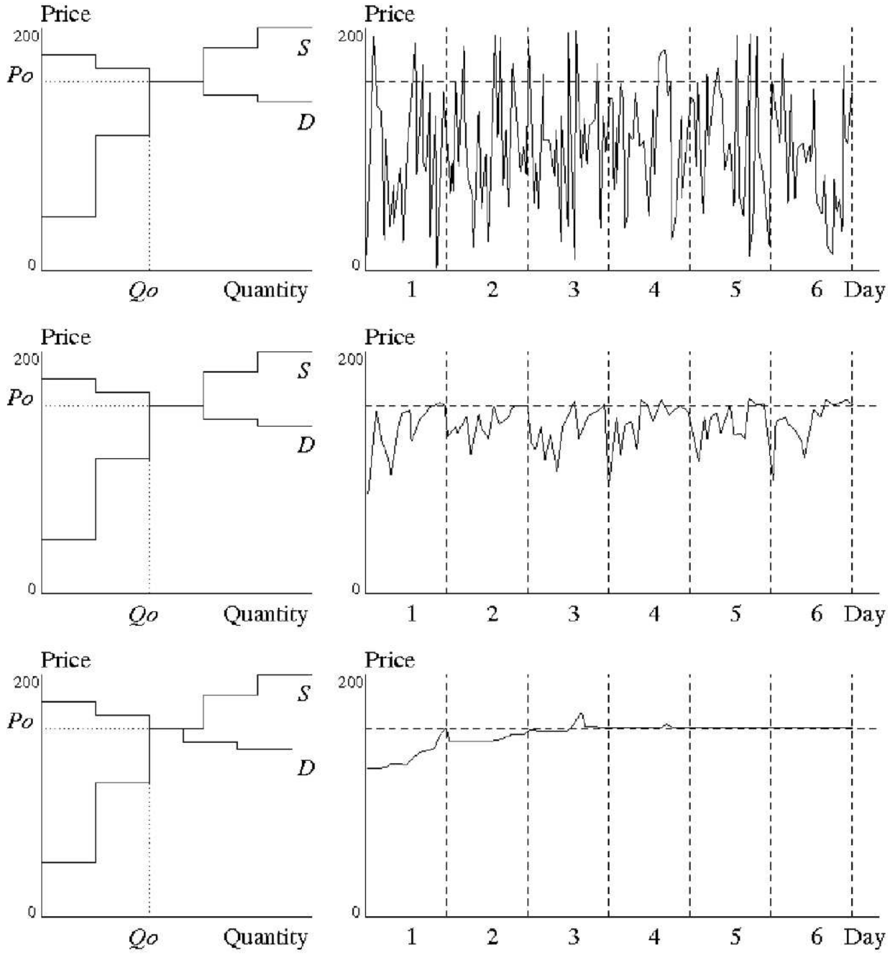


Fig. 5.6. Typical results from one of Gode and Sunder's experiments. Top: ZI-U traders. Middle: ZI-C traders. Bottom: Human traders. Supply and demand schedules for this experiment are shown on the left. Source: [76, p. 6], redrawn from [150, p 127].

lowest. Extra-marginal units, i.e., those beyond the equilibrium point and therefore difficult to trade by budget-constrained traders, were also traded, albeit at a loss, thus accounting for the lower efficiency. It varies, depending on the demand and supply schedules, between 50% and 90%. The allocative efficiency in human markets averaged 97.9%, and that for ZI-C markets was about 98.7%. Even though Gode and Sunder did not perform any statistical tests, this difference in efficiency hardly seems statistically significant. Gode and Sunder concluded that imposing market discipline on random-unintelligent behavior is enough to improve the efficiency from the baseline level to that attained by human traders.

While Cliff and Bruten [75] do not question that the structure of a double auction market is responsible for achieving high levels of allocative efficiency, they show that Gode and Sunder's claim that the convergence of transaction prices to the theoretical equilibrium price is a consequence of market discipline is incorrect. By using probability density functions, they show that the mean transaction price is close to the equilibrium price only when the gradients of linear supply and demand curves are of the same magnitude and of opposite signs. For all other cases, the expected difference between transaction prices in ZI-C markets and the equilibrium price can be determined in advance by a probabilistic analysis. Simulation results supported these findings. Hence, Cliff and Bruten conclude that ZI-Q traders do not have sufficient rationality to exhibit the equilibrating tendencies of human traders in double auction markets.

In a related paper, Cliff and Bruten [76] developed "zero-intelligence-plus" (ZIP) traders that use a first-order adaptive mechanism to adjust their bids and offers. By increasing or decreasing their profit margin based on the latest entered bid or offer, Cliff and Bruten implicitly establish some kind of utility function which agents do not maximize, but improve upon. Simulations with different types of supply and demand schedules showed that these improved ZIP traders were able to slowly converge to the equilibrium price. A more extensive discussion of Gode and Sunder's zero-intelligence traders can be found in [103]. A new model applying the zero-intelligence approach to financial markets has recently been presented by Farmer et al. [123].

5.4.2 Models with a Random Communication Structure

A series of models with little rationality on the agents' part was developed by researchers from the field of econophysics [81, 194, 195, 408, 407]. These models have in common that the decision processes leading to the individual demands are not explicitly modeled, and individual demands are assumed to be a random variable. Cont and Bouchaud [81] argue that it is highly unlikely that the individual demands in real markets are independent. Therefore, they add a random communication structure, allowing agents to share information and coordinate their actions. Random communication structures, even though differently specified in most models, are a distinctive feature shared in this class of models.

The Cont-Bouchaud Percolation Stock Market Model

In Cont and Bouchaud's model [81], an agent i is linked to any other agent j with probability $p_{ij} = p = c/N$. The links could be considered as communication channels over which agents exchange ideas and coordinate their actions. The average number of agents to whom a trader is connected to is given by $(N-1)p$, N being the number of agents. This linking structure and the linking probabilities are similar to those of percolation models which are used in physics to describe how liquids percolate through porous objects.[35] Percolation systems in physics are characterized by a critical percolation threshold p_c below which no percolation can occur. If the probability with which the constituting elements are linked is in the vicinity of the threshold level p_c, the system is near its critical point at which its state variables can be described through power laws. This feature makes percolation theory attractive to financial modeling since return distributions and other variables exhibit power law behavior. Percolation models in economics may help in understanding how local influences "percolate" through an entire economy, causing macroeconomic variables to fluctuate, even though the links are randomly and uniformly distributed [129].

All agents that are linked to each other form a coalition that acts in unison. That is, they all either buy or sell one unit of stock, or stay inactive. Aggregate excess demands are transformed into asset price changes via a standard linear price adaptation rule. The numerical value of the parameter c is crucial for the model behavior since it controls the willingness of agents to form clusters which could be interpreted as mutual funds or investors following specific trading rules in real markets. Whether a cluster acts as a seller, buyer, or stays inactive is also randomly determined and independent of cluster size. At the percolation threshold $c = 1$, the probability density for the cluster size distribution decreases asymptotically as a power law, while for values for c slightly smaller than one, the cluster size distribution is cut off by an exponential tail.

Given the random demand and supply of stock, a naive market model would most likely give rise to normally distributed asset re-

[35] Cont and Bouchaud depart from the general percolation structure of nearest-neighbor percolation on lattices by assuming that all agents have a positive probability of being connected with each other. This creates infinitely long interactions instead of the usual local interactions [408].

turns. The endogenous formation of trading clusters slightly below the percolation threshold of $c = 1$, however, mimics non-sequential herding processes through communication processes.[36] Since the analytically derived asset price distribution is characterized by fat tails and excess kurtosis as in real asset returns, Cont and Bouchaud's model suggests that herding behavior in financial markets is responsible for this stylized fact. Since their model is exogenously forced slightly below the critical point, it is not surprising to have such power law behavior in the state variables. Cont and Bouchaud thus point out that it would be interesting to know whether a modified model would endogenously converge to the critical region, as the concept of self-organized criticality by Bak et al. [20] suggests.

Ising-Models of Stock Markets

Other interesting models in this model class reproduce the positive cross-correlation between volatility and trading volume [194] and volatility clustering [195]. Traders are represented by the nodes of an $L \times L$ square lattice where the links depict their connectivity, i.e., each trader is connected to its four nearest neighbors. Individual demands are again determined without recourse to a utility function. Agents do, however, take a budget constraint into account. Before reaching a final decision, a trader i repeatedly inquires about the temporary decisions $S_{i,j} \in \{-1, 0, +1\}$ of his four nearest neighbors and calculates an activation signal

$$Y_i = \sum_{\langle i,j \rangle} J_{i,j} S_j + A\nu_i, \tag{5.14}$$

where ν_i is an idiosyncratic noise term and $J_{i,j}$ a measure of mutual influence.[37] Agent i arrives at his final decision

---

[36] Cont and Bouchaud and most other authors in the field of econophysics tacitly bypass the subtleties contained in the concept of herding. In this model class, herding is simply considered as a clustering process. Cluster formation could also be the result of "spurious herding", i.e., when groups face similar decision problems and information sets and end up acting similarly [35]. In economics, however, it has become standard to view herding as a process that can lead to systematic sub-optimal decisions by entire populations [99]. For instance, investors may reverse their initial investment decision after having observed other investors and intentionally herd by imitating their decision.

[37] The specification of the $J_{i,j}$ terms determines the system's behavior. $J_{i,j} = 1$ yields the well-known Ising-model which is used in physics to describe magnetization processes. In Iori's model, low levels of idiosyncratic noise would cause agents to give up their "degrees of freedom" and align themselves to large "magnetic"

$$S_i(t) = \begin{cases} -1 & : \quad Y_i(t) \leq \xi_i(t) \\ \ \ 0 & : \quad -\xi_i(t) < Y_i(t) < \xi_i(t) \\ +1 & : \quad Y_i(t) \geq \xi_i(t), \end{cases} \tag{5.15}$$

where $\xi_i(t)$ denotes individual threshold activation levels. This mechanism could be interpreted as a trade friction such as a transaction cost, which causes some agents to be inactive. A market maker determines

$$\text{aggregate demand } D(t) = \sum_{i:S_i(t)>0} S_i(t)$$

and supply

$$Z(t) = \sum_{i:S_i(t)<0} S_i(t),$$

and announces a new stock price

$$P(t+1) = P(t) \left(\frac{D(t)}{Z(t)}\right)^{a\left(\frac{D(t)+Z(t)}{L^2}\right)}. \tag{5.16}$$

While Cont and Bouchaud's market maker has a symmetric reaction function to order imbalances, Iori's price adjustment rule is asymmetric. Finally, the next period's individual threshold levels are adjusted according to

$$\xi_i(t+1) = \xi_i(t) \frac{P(t)}{P(t-1)}. \tag{5.17}$$

Numerical simulations were run for $L = 100$ agents and for different parameter values. Neither linking probabilities of zero, the independent agent scenario, nor probabilities close to the critical point, the Cont-Bouchaud percolation case, yielded satisfactory price and volume data. Only for the Ising case, i.e., when agents are linked to their nearest neighbors with a probability of one, volatility clustering of returns and a positive cross-correlation between price volatility and trading volume emerged. These characteristics arose purely from communication and imitation among traders, even in the absence of an aggregate exogenous shock, thus supporting the Interacting Agent Hypothesis by Lux and Marchesi [268]. The asymmetric price adjustment rule and the endogenous adjustment of individual threshold activation levels ξ_i, however, seem critical in achieving these interesting results.

---

clusters. They would choose identical buying/selling decisions and thus, would cause large stock price fluctuations. Note that setting $J_{i,j}$ to one with probability p, and to zero with probability $1-p$, leads to Cont and Bouchaud's percolation model.

Other examples in the indexpercolation model class of percolation and Ising-type models of stock markets are [66, 409] and Raberto et al. [344, 345] who developed the Genoa artificial stock market.

While the assumptions of random individual demands seem to require only minimally sophisticated agents, Bhamra studies rational profit maximizing agents in an Ising-like model [32]. In it, he finds that profit-maximization implies that agents should, to some extent, imitate each other's behavior. Thus, the use of the Ising-type models to describe agent-agent interactions can be justified on economic grounds.

5.4.3 Models of Chartist-Fundamentalist Interactions

Models investigating the dynamics from two [98] or three investors types [85, 97] are not new to the finance literature. While relaxing the assumption of investor homogeneity, the restriction to only a few strategy types, often a chartist and a fundamental trading rule, allowed those models to be analytically solved. There are, however, several models of chartist-fundamentalist interaction that require simulation techniques in order to fully describe their dynamics, some of which will be discussed below.[38]

Instead of postulating the simultaneous existence of two different trader types who stick to their initial trading rules, computational models of chartist-fundamental interaction often model the endogenous switching between trading strategies. An example in this line of research is the artificial stock market by Lux [266], which is further analyzed in [268, 269, 67].

The market is populated by chartists and fundamental traders. Fundamentalists are supposed to know the asset's fundamental value p_f with certainty. Assuming that the asset price p will sooner or later revert to its fundamental value, fundamental traders simply buy (sell) the asset when its price is below (above) its fundamental value.[39] Chartists, who are either optimistic or pessimistic about the market's development in the near future, derive their price expectations from past asset price movements.

Agents meet randomly, compare the profitability of their strategies, and may switch to another trading group if they perceive the other

[38] Some early computational models belonging to this group are discussed in [243].

[39] The fundamental value in Lux' model is simply a constant. In [268, 67], an exogenous news arrival process is introduced and the log changes of the fundamental value are assumed to be Gaussian random variables.

strategy to be more promising in the short run. Thus, even fundamental traders may be induced to become chartists since the market price p may stay above or below the fundamental value p_f for several periods.[40] The switching behavior within the chartist group between optimistic and pessimistic attitudes is modeled through mimetic contagion. A positive opinion index $x \in [-1, +1]$, i.e., when there are more optimistic chartists than pessimists, increases the probability that the latter jump on the bandwagon and become optimists themselves or vice versa. Switching between these different trader groups is formalized through six dynamic transition probabilities. Obviously, time varying fractions of optimistic or pessimistic chartists and fundamentalists effect the aggregate supply and demand for the risky asset in this financial market. The asset price dynamics are finally determined through a market maker whose price adjustment rule reacts sluggishly to order imbalances. Note that this specification implies that the market maker has to temporarily absorb additional excess demands.

Lux and Marchesi [269] first analyze the population dynamics in their artificial market. If z denotes the fraction of chartists, the market possesses the following stationary solutions:

1) $x^* = 0$ and $z^* = 1$ with arbitrary p,
2) $z^* = 0$ and $p^* = p_f$ with arbitrary x,
3) $x^* = 0$ and $p^* = p_f$ with arbitrary z.

Lux and Marchesi show that if the opinion index x is biased $(x \neq 0)$ and if the asset price does not equal its fundamental value, stable states do not exist. The two absorbing states 1) and 2) at which neither chartists nor fundamentalists exist are excluded from the numerical simulations by additional assumptions. The stability of the equilibria in 3), i.e., when the market price equals the fundamental value and when the chartist group as a whole is neither pessimistic nor optimistic, depends crucially on the fraction of noise traders z in the market. In the vicinity of a critical point, the artificial market is characterized by permanent fluctuations of the portion of both chartists and fundamentalists. Temporary increases in volatility occur when the number of chartists is relatively large. The market behavior, however, tends to stabilize itself

[40] That is, arbitrage profits for fundamentalist tends to be realized over the course of several trading periods. Capital gains and losses, however, accrue immediately. This asymmetry might explain why myopic fundamentalists might choose not to use their information about the asset's fundamental value.

through the tendency of agents to become fundamentalists when there are large deviations between market price and fundamental value.[41]

Numerical simulations with 500 agents confirmed the theoretical results in that extended periods of quiet were followed by sudden bursts of clustered volatility in returns. The generated return distribution is thus heteroskedastic and leptokurtotic with its tails following a power law with a tail index between 2 and 4. Similar to empirical data, the autocorrelations of squared and absolute returns are positive and decay hyperbolically while the raw returns are almost uncorrelated. Given the Gaussian news arrival process and the deviations of the return distributions from normality, Lux and Marchesi [268, 269] and Chen et al. [67] conclude that market interactions of agents magnify and transform exogenous noise (news) into fat-tailed returns with clustered volatility, a phenomenon they labeled the Interacting Agent Hypothesis (see also section 5.3.4).

Similar microscopic simulation models of chartist-fundamentalist interactions have been proposed in [19, 121, 151, 152, 153]. Analytical models of chartist-fundamentalist interactions are surveyed in [190].

5.4.4 Many-Strategy Models with Learning

The models discussed so far were, in terms of model structure and agent rationality, quite simple. None of the models explicitly assumed utility maximizing agents. Rather, demands were often random and/or binary, i.e., agents either bought or sold exactly one unit of stock. Furthermore, the choices that agents faced were very limited in that they were restricted to buying or selling one unit of stock or following a chartist or fundamentalist trading strategy.

Many-strategy models, on the other hand, usually abandon many of these limitations and focus on the question of emergence. They are interested in which types of trading strategies will appear and survive in a dynamically changing market environment. The question whether the learning agents with their constantly co-evolving mix of trading strategies will give rise to a rational expectations equilibrium is another commonality in this model class. Often, these models assume a higher level of agent rationality by using agents who derive their optimal demands by maximizing utility functions.

[41] In the physics literature, this dynamic behavior is known as on-off intermittency, i.e., an attracting state becomes temporarily unstable due to a local bifurcation. Endogenous forces, though, drive the physical system back to its equilibrium state.

A simple example is the artificial market by Lettau [246]. Although the market structure is still very simple, the model takes us one step further with respect to agent rationality. Instead of having only two or three decision alternatives, Lettau's agents are confronted with a continuum of choices of how many shares x to hold of a risky stock that pays a stochastic dividend d with mean $\bar{d}$ and variance σ_d^2. Even though agents possess a constant absolute risk aversion utility function with coefficient λ

$$U(w) = -e^{-\lambda w}, \tag{5.18}$$

$w = x(\bar{d} - p)$ being the net payoff, they do not determine an optimal amount $\widehat{x}$. Lettau is rather interested in how close adaptive agents equipped with an evolutionary GA-learning algorithm could come to the optimal solution

$$\widehat{x} = \widehat{\alpha}\left(\bar{d} - p\right) = \frac{1}{\lambda\sigma_d^2}\left(\bar{d} - p\right) \tag{5.19}$$

of the maximization problem.[42] In his computer model, agents are thus depicted as binary solution strings[43] of length L, where each string $\omega_t = (\omega_{1,t}, \ldots, \omega_{L,t})$ is an instance of α_t in the interval $[\min, \max]$. The binary coding of an agent is thus simply

$$\alpha_t = \min + (\max - \min)\frac{\sum_{j=1}^{L} \omega_{j,t} 2^{j-1}}{2^{L-1}} \tag{5.20}$$

(compare also section 4.4.2). At the beginning of the simulation, all bit positions of an agent are randomly initialized to either zero or one. Each period t is divided into S subintervals in which agents order stock according to the current value of α. At the end of each period, a new generation of agents is generated. The probability of agent i being copied into the next generation is positively correlated with its cumulative utility in t

[42] In order to focus on learning behavior, Lettau assumes that the stock price is exogenously determined and not influenced by the adaptive agents.

[43] Note that this specification is a single-population GA where each agent represents a solution. In multi-population GAs, each agent holds a variety of different candidates and the GA is actually run inside each agent. Single-population GAs are sometimes equated with social learning since the genetic operators are applied across different individuals. Multi-population GAs then refer to individual learning since no exchange of genetic material between individuals occurs. Vriend showed that there are different aggregate outcomes depending on whether learning is modeled as social or individual learning [431].

$$U_{cum} = \sum_{j=1}^{S} U_i(\omega_{i,j}). \tag{5.21}$$

S determines the sample length over which a rule's fitness is evaluated. After copying has taken place, agents are modified by mutation and crossover.

The length of the rule evaluation period S turns out to be crucial for the effectiveness of the learning algorithm. For smaller S, agents hold considerably more of the risky asset than the optimal amount. For large S, Lettau's agents come very close to the optimal portfolio weight $\hat{\alpha}$, yet the bias to hold more risky stock remains. The intuition behind this bias is quite simple. For small sample sizes, those agents that took risks and did well because of favorable realizations of the random dividends have a higher reproductive probability than conservative agents. As the rule evaluation period S increases and agents learn more about the true dividend distribution, former lucky agents will experience some rare negative events and the fraction of conservative agents tends to grow. As $S \to \infty$, the bias vanishes.

Lettau's basic model is interesting in itself because of the details it reveals about genetic algorithm learning. He then continues to explain the flows in and out of mutual funds through a slightly modified model version. Standard financial theory such as the CAPM [288, 289, 389] cannot explain why the flows into mutual funds are positively correlated with returns and why investors act more sensitively to negative returns than to positive ones. These empirical mutual fund flows, however, are surprisingly well replicated by Lettau's market when setting S at 1 and replacing 3 agents in each period with new randomly initialized ones.

Because Lettau's artificial market used an exogenous price process which was not influenced by the agent's buying and selling decisions, it is not well suited to replicate the aforementioned stylized facts of real financial markets. Another market with many co-evolving trading strategies and agent learning was suggested by Chen and Yeh [68]. Their basic model framework is similar to Grossman and Stiglitz [164], i.e., agents have a constant absolute risk aversion utility function which they maximize subject to a budget constraint. When deriving their optimal demands, they need to form expectations about next period's price and dividend.[44] They do so by acquiring specific forecasting models from a business school where faculty members engage in creating and publish-

[44] The SFI-ASM [244] has the same basic structure and will be introduced in more detail in the next chapter.

ing new forecasting models.[45] The creation of new trading strategies is implemented by means of genetic programming. Even though Chen and Yeh do not intentionally fine-tune their model parameters for the purpose of replicating the stylized facts of financial time series, the endogenous price and return series still turn out to be non-normally distributed. Besides fat-tailedness, return series are also independently and identically distributed (iid), which supports the EMH. Chen and Yeh emphasize that this feature is especially surprising since martingale traders, i.e., those whose set of trading strategies reflect a belief in the EMH, become extinct in the course of a typical simulation run. In a newer paper, Chen and Yeh thus interpret the validity of the EMH on the macro-level as an emergent property since it should not be expected from the behavior of individual investors [69].[46] In addition, they test the Rational Expectations Hypothesis by examining the series of aggregate forecasting errors. The mean forecasting errors are found to be only insignificantly different from zero, thus, on the aggregate level, they do not make systematic errors in the mean. Furthermore, by looking for linear and non-linear patterns in the error series, the null of iid-nes cannot be rejected for most subperiods. Thus, Chen and Yeh conclude that the Rational Expectations Hypothesis is another emergent property in their artificial stock market.

[45] This learning structure addresses some methodological criticism concerning the spreading of trading strategies within populations. It is argued that in single-population GAs/GPs, only a sequence of actions may be observable, but not the strategies that lead to these actions. Thus, imitation cannot explain the spreading of strategies within the population. Since Chen and Yeh consider the multi-population GA/GP approach to be an unsatisfactory response to this criticism, they resolve this methodological issue through a "business school" in which faculty members are forced to publish their results.

[46] Contrary to their older paper [68], the share of martingale believers is not zero, but with a share of approximately 1%, very small.

Part II

The Santa Fe Institute Artificial Stock Market Model Revisited

6

The Original Santa Fe Institute Artificial Stock Market

Science is a long history of learning how not to fool ourselves.
Richard P. Feynman[1]

6.1 Introduction

Similar to the models discussed in the previous section, the Santa Fe Institute Artificial Stock Market (SFI-ASM) is a model with many trading strategies which are improved over time. The departures from the neoclassical framework are only minimal. Agents derive their optimal demands by maximizing a utility function and, therefore, behave quite rationally in the standard economic sense. Only because they are assumed to be heterogeneous with respect to their expectations, can they not be modeled as fully rational.

By focusing on the dynamics of learning, the SFI-ASM identifies a single parameter, i.e., the learning speed of agents, which is able to shift the model to either a regime that is close to the homogeneous rational expectations equilibrium, or to a more complex regime that better fits the empirical facts. The complex regime emerges for fast learning rates and is characterized by more complicated price time series and by substantial levels of technical trading.

After a detailed description of the SFI-ASM's basic model structure and its results, it will be shown in this chapter that the emergence of the complex regime could be an artifact of design assumption. A closer investigation of the genetic algorithm that updates the trading rules of agents reveals that the SFI mutation operator causes a systematic upward bias in the level of set bits in the condition part of

[1] Quoted in Cole [77, p. 10].

trading rules. Faster learning speeds with more mutations per time period imply increased levels of technical (and fundamental) trading. Thus, postulating the emergence of technical trading by only looking at the aggregate level of technical trading bits may be too premature.

6.2 The Marimon-Sargent Hypothesis and the SFI-ASM

According to Waldrop [433, p. 270], the development of the SFI-ASM was a direct result of a discussion that took place at the Santa Fe Institute in March 1989. Ramon Marimon and Thomas Sargent contended that adaptive agents in an artificial stock market model would quickly discover the rational expectations equilibrium solution. In spite of a few random fluctuations up and down, prices would converge to the fundamental values of traded stocks as predicted by neoclassical theory. Marimon and Sargent bolstered their claim by their own research [285] in which they had adaptive classifier agents assigned to solve Wicksell's triangle.[2] Their artificially intelligent agents were always able to learn the neoclassical solution previously found by Kiyotaki and Wright [226], i.e., the good with the lowest storage cost emerged as a generally accepted medium of exchange.[3]

John Holland and Brian W. Arthur, on the other hand, could not conceive of adaptive agents being able to find the neoclassical equilibrium solution. In their view, an artificial stock market would be far more complicated than the simple and well-defined problems that had been proven solvable by adaptive agents. The rational expectations equilibrium would require that heterogeneous agents infer their own expectations by an infinite regress of other agents' expectations, a task that Holland and Arthur considered to be too difficult to be solved by realistically modeled artificial agents. The complications arising from

[2] Wicksell's triangle refers to a situation in which three types of agents would like to consume the goods possessed by one other type of agent, yet they have nothing to offer that has any consumption value to these agents. In other words, there are no double coincidences of wants in the economy. Assuming that the reallocation of consumption goods has to be quid pro quo, a Pareto efficient allocation can never be achieved as long as agents are not willing to give up their good in exchange for something they do not value for their own consumption.

[3] Sargent's general view of the use of genetic algorithms in economics becomes clearer in his book on bounded rationality in macroeconomics [372]. In it, he mainly uses genetic algorithms as a substitute to justify rational expectations in macroeconomic general equilibrium models. This approach, however, is criticized by Moss and Edmonds who see the role of genetic algorithms unduly reduced to bring about the rational expectations equilibrium [312].

the interactions of many interacting heterogeneous agents would lead to a complex stock market behavior quite different from a homogeneous rational expectations regime.

6.3 An Overview of SFI-ASM Versions

Since the debate could not be settled without having an actual version of an artificial stock market, Arthur and Holland decided to program their own stock market simulation which became known as the Santa Fe Institute Artificial Stock Market (SFI-ASM). The first version was operable by the end of 1989 and was described in [331]. The studies by Arthur et al. [11] and LeBaron et al. [244] were based on a modified Objective-C version which used a market clearing mechanism instead of an excess demand price adjustment mechanism. It was later revised by Brandon Weber and Paul Johnson to run with the Swarm libraries, a well-known toolkit for agent-based simulations for Objective-C and Java. This ongoing effort is currently hosted by Paul Johnson at http://ArtStkMkt.sourceforge.net [204]. The SFI-ASM has inspired various modelers to do their own research. Joshi et al. [206, 207] slightly adapted the original Objective-C version to analyze wealth levels. Tay and Linn [418] extended the original model by using fuzzy logic for expectation formation. Wilpert [437], who used his own Borland C++ implementation, tested the model with different modifications. He used, for instance, the generated profits of trading rules instead of forecast accuracies as a fitness criterion. Gulyás [166] programmed a participatory market model in which real humans were placing market orders alongside the artificial adaptive agents. They were thus able to study the effects on actual human decision making on the market behavior and vice versa.

For this book, the Objective-C version 7.1.2. was ported to Java (Java 2 SDK, standard edition, version 1.4.0).[4] In order to distinguish between the original SFI-ASM and the current Java-version with which the simulation results were derived, I will refer to the latter as either the Java-version of the SFI-ASM, or, when referring to the suggested modification of the mutation operator, as the modified SFI-ASM. The Java-version of the SFI-ASM was programmed by using the Repast (*Recursive Porous Agent Simulation Toolkit*) library [78].[5] Agent-based

[4] The source code is available upon request.

[5] Repast is being developed at the University of Chicago and is freely available under the BSD license agreement. The Repast website is located at

simulation toolkits such as Swarm and Repast save agent-based modelers from building their computer models from scratch. These libraries provide non-content specific routines such as simulation interfaces or data output procedures that are designed and thoroughly tested by professional developers. Social scientists are thus able to concentrate on their model specific structures and interactions. The apparent standardization that comes with the use of a commonly accepted agent-based toolkit also allows for a more efficient exchange of program code. This enhances the understanding and readability of the by now many agent-based computer simulations.

6.4 The Basic Structure of the SFI-ASM

The SFI-ASM borrows much of its structure from Grossman and Stiglitz [164]. It is inhabited by N traders, who are all initially endowed with one unit of risky stock and $20,000$ units of cash. During each period, traders have to decide how much to invest in risky stock and how much to keep in cash which yields a risk-free rate of return r_f.

The stock pays a stochastic dividend per period which is generated by a mean-reverting autoregressive Ornstein-Uhlenbeck process

$$d_{t+1} = \bar{d} + \rho(d_t - \bar{d}) + \epsilon_{t+1}. \tag{6.1}$$

$\bar{d}$ denotes the dividend mean, ρ is the speed of mean reversion, and ϵ refers to stochastic shocks which are normally distributed with mean zero and variance σ_ϵ^2. Traders are homogeneous with with respect to their utility function, i.e., they all have the same constant absolute risk aversion (CARA) expected utility function

$$U(W_{i,t+1}) = -e^{-\lambda W_{i,t+1}}, \tag{6.2}$$

with λ being the degree of risk aversion and $W_{i,t+1}$ being agent's i expected wealth level in the next period. In determining their optimal

http://repast.sourceforge.net from which the latest distribution can be downloaded. While the reprogramming of the SFI-ASM used Repast 1.4, the Repast version at the time of writing is 3.1. In a recent comparison of free JAVA-libraries for agent-based simulation in the social sciences [423], Repast has been found to be the clear winner. Among the several rating criteria were official program documentation, statements by developers and users, and experiences and impressions by the evaluation team. Another review of agent-modeling toolkits is in [148]. A good overview of current agent-based simulation toolkits is maintained by Leigh Tesfatsion at the University of Iowa and can be found at http://econ.iastate.edu/tesfatsi/acecode.htm.

demand for the risky stock, agents are perfectly myopic in that they only consider next period's expected returns. Agents maximize their expected utility subject to the budget constraint

$$W_{i,t+1} = x_{i,t}(p_{t+1} + d_{t+1}) + (1 + r_f)(W_{i,t} - p_t x_{i,t}), \tag{6.3}$$

where $x_{i,t}$ is the amount of stock agent i holds in period t. Under the assumption of normally distributed stock returns, the optimal amount of stock $\widehat{x_{i,t}}$ that agents desire to hold is then determined as

$$\widehat{x_{i,t}} = \frac{E_{i,t}[p_{t+1} + d_{t+1}] - p_t(1 + r)}{\lambda \sigma^2_{t,p+d}}, \tag{6.4}$$

where $E_{i,t}[p_{t+1} + d_{t+1}]$ is i's expectation in t about next period's realization of the stock price and dividend, and $\sigma^2_{t,p+d}$ the empirically observed variance of the stock's combined price plus dividend time series. LeBaron et al. point out that the normality assumption of stock returns holds in the homogeneous rational expectations equilibrium (hree), but outside the hree-regime, it is not clear whether stock returns will be normally distributed [244].

The effective demand of an agent is the difference of his actual and desired stock holdings. Once agents have determined their effective demands, they submit them as well as their partial derivatives with respect to the price to a specialist, who tries to balance the effective demands by setting a market clearing price. If the specialist is not able to find a market clearing price in the first place, an iterative process is started in which new trial prices are announced and agents update their effective demands and partial derivatives accordingly. If complete market clearing is not reached within a specified number of trials, one side of the market will be rationed.

6.4.1 Trading Rules and Expectation Formation

While traders are homogeneous with respect to their utility functions and degrees of risk aversion, they differ when deriving their expectations about future prices and dividends $E_{i,t}[p_{t+1} + d_{t+1}]$. In a way, they differ in processing an identical information set. Price and dividend forecasts are generated by using individual trading rules of the form

if (condition fulfilled), then (derive forecast).

These "condition-forecast" rules are a modified version of the condition-action classifier system by John Holland [185]. While the latter maps directly from condition into action, the "condition-forecast" rules do so only indirectly.[6] First, a forecast is produced which, by using equation 6.4, then will be converted into an action, i.e., an agent's bid or offer for the risky stock.

Each of the $j = 1 \dots 100$ trading rules that every agent possesses consists of a condition part, a forecast part (predictor), a numerical value $\Phi_{t,i,j}$ for its fitness, and a value for its forecast accuracy $\nu^2_{t,i,j}$, i.e.,

$$\text{rule}_{i,j} = \{(\text{condition part}); (\text{predictor}); \text{fitness, forecast accuracy}\}$$

The condition parts are checked against a Boolean market descriptor D_t which holds current and past price and dividend information. For example, a particular market state could be that the price of the stock is greater than n-times its fundamental value, while at the same time, the 25-period moving average of the stock price is greater than the current price. When a particular predefined condition is met, the corresponding descriptor bit is set to 1, and otherwise to 0.

A rule's condition part, on the other hand, is coded as a ternary string holding either 1 or 0, depending on whether the corresponding bit in the market descriptor has to be matched or not, or holding # if the rule ignores that particular descriptor bit.[7]

An example of how rules are checked against a market descriptor with seven conditions $A – H$ is shown in table 6.1. Rule 2 and rule 100 are not activated, since in both cases, two conditions are not fulfilled. On the other hand, rule 1 and rule 99 match the market descriptor and hence, are activated. As one can see, rule 1 is rather general. Since it contains numerous #-signs, it is quite insensitive to changes in the environment and will probably be activated quite often, as it is the case in this period. All other rules are more detailed and describe more specific market states.

The bits of a trading rule may be characterized as either technical or fundamental. Technical bits check only price information, for instance, whether the stock price has gone up or down during the last periods,

---

[6] The first SFI-ASM version mapped directly from states of the world into actions. All later versions use the described condition-forecast rules.

[7] Technically, the units in the ternary strings should be called *trits*. A trit is the smallest unit that can hold three values. However, as is usually done in the literature, I will refer to them as bits. Alternatively, when using the terminology known from genetic algorithms, I will refer to these bits as genes. The condition parts correspond to chromosomes. The bit positions at which genes reside in the chromosome are called gene loci, and the three possible values that a bit (gene) can take on are called alleles. The cardinality of the allele-alphabet in the SFI-ASM is three.

Table 6.1. Comparison of an agent's rule set against a hypothetical Boolean market descriptor. Rules with matching condition parts are marked as active.

condition	A	B	C	D	E	F	G	H	match ?
market descriptor D_t	0	0	0	1	0	1	0	1	-
rule 1	0	#	#	#	#	#	#	#	✓
rule 2	#	1	#	0	#	1	#	#	-
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
rule 99	0	#	#	1	#	1	#	1	✓
rule 100	0	1	#	#	#	0	0	#	-

or whether certain moving averages of prices are bigger or smaller than other moving averages. Fundamental bits, on the other hand, relate the price of a stock to its fundamental value by using dividend information. For instance, prices are checked against a stock's fundamental value by comparing for each ratio in the brackets whether

$$\text{price x interest rate/dividend} > \left\{ \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{95}{100}, 1, \frac{9}{8} \right\} \tag{6.5}$$

is fulfilled. The conditions used by the SFI-ASM versions differ from those in the current Java SFI-ASM version.[8] Increasing the number of checked trading conditions might have an effect on what Axelrod [13] calls distributional equivalence, i.e., different statistical properties of time series, but the relational equivalence should remain unaffected.

From the set of 100 individual trading rules, normally several rules match the market descriptor. From the set of activated rules, agents now have to choose one for their forecast production. LeBaron et al. use the forecast accuracies of activated rules as a tie breaker, i.e., the best rule is always selected.[9] Finally, agent i determines his forecast of

[8] To state how many conditions are actually checked in "the original SFI-ASM" is quite difficult to say. LeBaron et al. [244] document only 12 conditions (including 2 dummy bits), while the Objective-C version 7.1.2 used by Joshi et al. [206], which served as the blueprint for the Java SFI-ASM, had a total of 61 conditions with three dummy bits. The reprogrammed Java version checks a total of 64 conditions, 32 fundamental and 32 technical trading bits. The differences in checked trading conditions between the three SFI-ASM versions are documented and discussed in appendix 11.2.

[9] Instead of always selecting the most accurate trading rule, Joshi et al. [206, 207] use the roulette wheel mechanism. This selection algorithm assigns to each rule a wedge on a roulette wheel which is proportional in size to its relative fitness. Thus, rules with higher fitness values are more likely to be chosen than those with

next period's price and dividend according to the linear equation

$$E_{t,i}[p_{t+1} + d_{t+1}] = a_{t,i,j}(p_t + d_t) + b_{t,i,j}, \tag{6.6}$$

where $a_{t,i,j}$ and $b_{t,i,j}$ are real-valued parameters constituting the predictor part of the chosen trading rule j. Only when no rules match the market descriptor, parameters a and b are determined as a fitness-weighted average of all trading rules in his rule set.

One period later, the accuracy of all activated rules is checked by comparing their predictions $E[p_{t+1} + d_{t+1}]$ with the actual realization of $(p_{t+1} + d_{t+1})$. A rule's forecast accuracy is determined as

$$\nu^2_{t,i,j} = \left(1 - \frac{1}{\theta}\right) \nu^2_{t-1,i,j} + \frac{1}{\theta}\Big[(p_t + d_t) - [a_{t,i,j}(p_{t-1} + d_{t-1}) + b_{t,i,j}]\Big]^2 . \tag{6.7}$$

This forecast accuracy is measured as a weighted average of previous and current squared forecasting errors. The parameter θ determines the size of the time window that agents take into account when estimating a rule's accuracy. As LeBaron et al. have pointed out, the value of θ is a crucial design question since it strongly affects the speed of accuracy adjustment and learning in the artificial stock market. If $\theta = 1$, trading rules would be judged only on last period's performance, and forecast accuracy would be strongly prone to noise. As in LeBaron et al., a value of 75 is chosen for θ.

The forecast accuracy $\nu^2_{t,i,j}$ is used as a rule's variance estimate $\sigma^2_{t,(p+d)}$, which is used in equation 6.4 to determine the optimal stock holdings of agents. Furthermore, it is the main determinant of a rule's fitness

$$\Phi_{t,i,j} = C - \left(\nu^2_{t,i,j} + \text{bit cost} \times \text{specificity}\right), \tag{6.8}$$

with specificity being the number of conditions in a rule that are not ignored, with bit cost as an associated cost for each non-ignored condition, and C as a positive constant to ensure positive fitness.[10] Non-zero-bit costs per set trading bit (0 or 1) could be interpreted as the cost of acquiring and evaluating new information. It could also be seen as a complexity aversion, since simple rules are favored over more specific ones. Most importantly, however, LeBaron et al. emphasize that positive bit costs would ensures that non-# bits in a trading rule serve

low fitness. The specific problems that arise from roulette wheel selection were already discussed in section 4.4.2 on page 41.

10 The maximum squared forecast error of a trading rule has an arbitrary ceiling of 100, hence, the numerical value for C is set at this value.

a useful purpose, i.e., that they have some informational content.[11] Since bit costs bias the resulting bit distribution towards the all-# rule, LeBaron et al. claim that non-# trading bits will only survive if they have some predictive value.

6.4.2 Learning and Rule Evolution

So far, agents have been equipped with a static rule set. Feedback learning in the stock market has taken place by identifying and using the rules that produced better forecasts than others. The quality and the speed of this type of learning was strongly dependent on the parameter θ which determined over how many past periods agents averaged the forecast accuracy of their trading rules. However, if agents started with a rule set that contained only bad performing rules, in the absence of any other learning mechanism, they would not be able to find better ones.

Agents therefore use an additional genetic algorithm (GA) learning procedure that allows them to alter their rule set by replacing poorly performing rules with new, possibly better ones.[12] Exploratory GA learning usually happens on a slower evolutionary time scale than the feedback learning and examines the search space in a random, yet not directionless fashion. For each agent, the GA is, on average, asynchronously invoked every K periods and replaces the 20 worst rules of the rule set. The GA-invocation interval K alters the learning speed of agents and turns out to be the most crucial model parameter.

New trading rules are created either through *mutation* (with predictor mutation probability $\Pi = 0.9$) or through *crossover* (with probability $1 - \Pi$). Crossover is a sexual genetic operator that needs two parents to work, both of whom are chosen by tournament selection.[13] Even though there are a variety of different crossover operators available, the SFI-ASM uses exclusively uniform crossover for the condition

---

[11] "*The purpose of this bit cost is to make sure that* ***each*** *bit is actually serving a useful purpose in terms of a forecasting rule*" (p. 1497). Emphasis added by author.

[12] See, also, section 4.4.2 on GA learning.

[13] For tournament selection in the SFI-ASM, two genetic individuals, i.e., trading rules, are randomly selected from the gene pool, and the fitter of both is chosen. Goldberg [154] points to another popular tournament selection algorithm by Wetzel. Instead of randomly selecting two candidates, Wetzel-ranking picks the two candidate solutions by using roulette wheel selection. It is not clear whether simple tournament or Wetzel-ranking is superior or whether it greatly affects results, yet Wetzel-ranking should return, on average, fitter parents.

parts. Here, an offspring's bit is chosen with equal probability from the corresponding bit positions of either parent.

Table 6.2. Example for uniform crossover on the condition part.

parent 1	#	0	1	#	#	#	#	1	1
parent 2	#	0	1	1	1	0	0	0	0
offspring	#	0	1	#	1	#	0	1	0

Note that the fraction of set trading bits in an offspring is an unweighted average of the two parents' bit fractions. There is no systematic influence on the average specificity through the working of the crossover operator.

As for the real-valued forecasting parameters a and b, one of three methods is randomly chosen with equal probability. First, both parameters are exclusively taken from one randomly selected parent. Second, each individual parameter is selected from either one of the two parents. Third, new parameter values are created by determining a weighted average of the two parents' values, with $1/\sigma^2_{j,p+d}$ as the weight for each parent. The weights are normalized to sum up to one.[14]

It is apparent that the crossover operator is incapable of introducing either zero, one, or #-bits to a certain bit position if they are not yet contained in the parent gene pool for that particular bit position. Crossover is also unable to generate real-valued parameters for a and b that are outside the interval created by the minimum and maximum values in the parent gene pool. These limitations are overcome by mutation, which is an important part for any evolutionary algorithm. It helps to maintain a diverse population by forcing new solutions into the population and thus, avoids premature convergence of the search algorithm. For mutation, one parent is chosen by using tournament selection. First, a genetically identical offspring of that parent is created. The real-valued parameters of the predictor parts are mutated in one of three possible ways. With probability 0.2, they are uniformly changed to a value within the permissible ranges of the parameters which is $[0.7, 1.2]$ for the a parameter and $[-10.0, 19.0]$ for the b parameter. With probability 0.2, the current parameter value is uniformly dis-

[14] LeBaron et al. pointed out that there is little experience in the GA community on how to perform real-value crossover. By now, it has become more common and the early distinction between genetic algorithms and evolutionary strategies has become fuzzy.

tributed within ± 5% of its current value, and it is left unchanged for the remaining cases.

The individual bits in the condition parts are mutated with a small bit mutation probability of $\pi = 0.03$. Once a bit in the condition part has been chosen for mutation, it will be changed according to the following bit transition probabilities:

$$\mathrm{P} = \begin{array}{c} \\ 0 \\ 1 \\ \# \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & \# \end{array} \\ \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \end{array} . \tag{6.9}$$

This matrix of transition probabilities specifies that a 0-bit is changed with a probability of one third to a 1-bit, and with probability of two thirds to a #-bit. Similarly, a 1-bit changes with one third probability to 0, and with two thirds to #. Don't care signs # change with equal probability of one third to either 1 or 0, or remain unchanged. LeBaron et al. assert that these transition probabilities would, on average, maintain the specificity, i.e., the fraction of #'s in a rule.

For crossover, the offspring inherits the average forecast accuracy of its two parents. For mutation, the offspring's forecast accuracy is set at the median forecast error over all rules in the agent's rule set.

6.4.3 Other Programming Details and Initialization of Model Parameters

Trading rules in the SFI-ASM contain bit sequences that belong together. Within such a sequence, certain bit combinations are invalid and constitute an illogical trading rule.

Table 6.3. Example of an illogical trading rule.

$\frac{\text{price x interest Rate}}{\text{dividend}} >$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{3}{2}$
rule 1	#	#	0	1	#	0	#

An example of an illogical trading rule is shown in figure 6.3. It describes a situation in which $p_t r_f / d_t$ is smaller than $7/8$, yet at the same time, that ratio is supposed to be greater than 1. Obviously, rules of this type will never be matched, but they will get inserted into an agent's rule set since the genetic algorithm does not check the trading rules it creates for logical consistency.

The SFI-ASM deals with this problem in two ways: through a large rule set, and with a generalization procedure. Once a trading rule has not been matched for more than 4,000 periods, it is generalized by converting one fourth of its 0 and 1 trading bits to #. The fitness of a generalized rule is set at the median value.

Agents also face some trading restrictions in that their orders are constrained to lie within their budget constraint and that they cannot go short more than five shares. In addition, the stock price is capped by the specialist at 200 and is bounded from below at 0.01. These constraints, however, seem to be binding only in the beginning of the simulation when the randomly initialized trading rules of untrained agents result in causing the stock price to fluctuate wildly. The same is true for a constraint that limits the squared forecast error to 100. This allows us to choose the constant C in equation 6.8, on page 98, such that fitness values are always positive.

A graphical depiction of the timing sequence of major activities in the SFI-ASM can be found in appendix 11.1

6.5 The Homogeneous Rational Expectations Equilibrium

Since agents are assumed homogeneous with respect to their preferences, the known structure of the dividend process makes it possible to determine the properties of the homogeneous rational expectations equilibrium (hree). Since the hre-equilibrium is in the set of possible outcomes in the SFI-ASM, it will be interesting to see whether agents will be able to converge to this hree-solution.

A rational expectations equilibrium specialist would determine an equilibrium price such that the expectations of agents

$$E_t^{hree}(p_{t+1} + d_{t+1}) = a^{hree}(p_t + d_t) + b^{hree} \tag{6.10}$$

are, on average, fulfilled. In the case of a linear rational expectations equilibrium, we can assume a linear function mapping the current dividend into a price

$$p_t^{hree} = f d_t + g. \tag{6.11}$$

We seek the parameters f and g for the hree-specialist and the parameters a^{hree} and b^{hree} for the agents forecasts that are compatible with a hre-equilibrium. Since agents are homogeneous and hold exactly one unit of the risky stock in this benchmark scenario, the known structure of the dividend process allows us to determine the hree-parameters for the specialist as

$$f = \frac{\rho}{1 + r_f - \rho} \tag{6.12}$$

and

$$g = \frac{(1+f)(1-\rho)\bar{d} - \lambda\sigma^2_{p+d}}{r_f} \tag{6.13}$$

with

$$\sigma^2_{p+d} = (1+f)^2\sigma^2_\epsilon \tag{6.14}$$

being the variance of the combined price plus dividend time series. For the agents, their hree-forecast parameters compute to

$$a^{hree} = \rho \tag{6.15}$$

and

$$b^{hree} = (1-\rho)\left((1+f)\bar{d} + g\right). \tag{6.16}$$

The hree-specialist thus sets the hree-price according to the given dividend. This price then ensures that all agents want to hold exactly one unit of stock, given their hree-forecast of next period's price plus dividend. The normal model behavior for heterogeneous agents can now be assessed by comparing it with the properties of the homogeneous rational expectations equilibrium.

6.6 The Marimon-Sargent Hypothesis Refined

Having determined the properties of a hre-equilibrium within the SFI-ASM model structure, it is now possible to refine the Marimon-Sargent Hypothesis so that adaptive agents would quickly converge to this benchmark solution. Within the SFI-ASM framework, the Marimon-Sargent Hypothesis comprises three parts.

First, the hre-equilibrium is characterized by a no-trade situation since no agent will ever deviate from his optimal position of holding exactly one unit of the risky stock. Given that the ongoing stochastic GA learning always introduces some noise into the system, minor trading activities will be inevitable.

Second, equation 6.11 tells us that the statistical properties of the price process are simply a linear transformation of the known properties of the stochastic dividend process. In the hre-equilibrium, the price process should thus be linear, normally distributed, and non-skewed.

Finally, equation 6.10 reveals that agents only need to know last period's price and dividend to derive their forecasts about next period's price and dividend, given the mean-reverting dividend process. Therefore, all the information provided by the condition parts of the classifier system is unnecessary. Given that the use of technical and fundamental trading bits is punished with associated bit costs, the hree-solution should be characterized by complete negligence of technical and fundamental trading bits.

6.7 Simulation Results of the SFI-ASM

6.7.1 Time Series Behavior

The simulation results obtained by the original SFI-version are documented in LeBaron et al. [244]. Since the computer model used in this book was a reprogrammed Java-version, it was necessary to test whether the two model versions behave similarly for the same parameter settings. This was done by running the same statistical tests on the time series and comparing the results with those published in [244, p. 1499-1512]. All model parameters were set to the same values reported there.

In order to allow for sufficient learning to occur, price and dividend time series were recorded between the periods 250,000 and 260,000. The simulated data of 25 simulation runs for two learning speeds were then compared by LeBaron et al. to the hree-benchmark case. For the slow learning case, the GA was invoked by the agents every 1,000 periods, on average. In the fast learning regime, that parameter was set to 250 periods.

In the hree-mode, the dividend and market price should be a linear function of their first order lags. Therefore, they are regressed on a lag and a constant

$$p_{t+1} + d_{t+1} = a(p_t + d_t) + b + \epsilon_t, \tag{6.17}$$

and the estimated residual time series $\hat{\epsilon}_t$ is analyzed to see whether it satisfies being iid and N(0,4) distributed. The results are summarized in table 6.4.

First of all, one notices that for both learning speeds, the new Java-version of the SFI-ASM produces time series that are close to those of the original SFI-ASM, but generally a little bit further away from the hree-benchmark. The standard deviations in the residuals are slightly

Table 6.4. Comparison of simulated data of the reprogrammed SFI version with the reported data from the original Objective-C version. Means over 25 runs. Numbers in parentheses are standard errors estimated using the 25 runs. Numbers in brackets are the fraction of tests rejecting the no-ARCH or iid-hypothesis for the ARCH and BDS tests, respectively, at the 95% confidence level.

Description	GA 1,000		GA 250	
	Java-SFI	Obj.-C SFI	Java-SFI	Obj.-C SFI
Std. Dev.	2.145 (.010)	2.135 (.008)	2.225 (.015)	2.147 (.017)
Excess kurtosis	0.085 (.015)	0.072 (.012)	0.285 (.033)	0.320 (.020)
ρ_1	0.031 (.008)	0.036 (.002)	0.025 (.012)	0.007 (.004)
ARCH(1)	3.502 [0.60]	3.159 [0.44]	25.34 [1.00]	36.98 [1.00]
ρ_1^2	0.020 (.002)	0.017 (.002)	0.075 (.008)	0.064 (.004)
BDS	1.41 [0.32]	1.28 [0.24]	4.08 [0.92]	3.11 [0.84]
Excess return	2.92% (.03%)	2.89% (.03%)	3.25% (.08%)	3.06% (.05%)
Trading volume	0.364 (.025)	0.355 (.021)	0.854 (.065)	0.706 (.047)

higher, thus, indicating a small increase in price variability. Excess kurtosis, albeit small, is positive for both the fast and slow learning cases. Empirical return distributions are usually more fat-tailed and higher-peaked than the simulated return distributions. The autocorrelation in the residuals, as shown in the third row, demonstrates that there is little linear structure remaining except for the extreme case of updating the rule set in every period. As LeBaron et al. indicate, any artificial stock market should exhibit negligible autocorrelations since they are very low for real markets.

The next row reports the means of the test statistics for the ARCH test proposed by Engle [110]. The ARCH dependence in the residuals increases with learning speed and is slightly higher for the Java-version. The table gives the means of the test statistics, averaged over 25 simulation runs. The numbers in brackets tell the fraction of runs that reject the no-Arch hypothesis at the 95% confidence level. In row five, the first order autocorrelation of the squared residuals is another test for volatility persistence. Again, it increases for faster learning speeds, but is generally lower for the original model version.

The BDS test in row six is a test for nonlinear dependence developed by Brock et al. [58]. Its test statistic is asymptotically standard normally distributed under the null hypothesis of independence.[15] One can notice an increasing amount of nonlinearities for faster exploration rates, yet again, it is slightly higher for the new Java-version. Trading volume, which should be zero in the hree-case, increases significantly for faster learning speeds. This points to a greater degree of heterogeneity between the agents.

Overall, the conclusion that the learning speed affects the price series behavior is confirmed by the new model version. The observed changes in both model versions have the same direction, and the absolute differences between them are relatively small. The tendency of the new Java-version being further away from the hree-benchmark case could be a result of the substantial increase in checked trading conditions. The original version checked only 12 conditions, while the new Java-version increased this to 64 trading conditions.

6.7.2 Forecast Properties

When analyzing whether the agents made use of the trading information given in the condition parts of their trading rules, LeBaron et al. analyzed the fraction of set trading bits, i.e., the fraction of all non-# bits, averaged over all the rules of all agents. They reported that for both learning speeds, the equilibrium bit level settles to levels below 5%. More important, though, is in their view the elevated bit level for the fast learning case, which they define to be at a GA-invocation interval of 250, with an indication of a continued increase beyond the 250,000th period. One long-run test with 1,000,000 periods yielded that the bit level eventually stops increasing, but it continued to show large swings. When focusing only on the 4 moving average trading bits in the condition parts, they found the increase in the bit level to be even more pronounced than in the slow learning case.

Figure 6.1 shows that the new Java-SFI version does not exactly replicate the described bit behavior. At period 250,000, the aggregate bit level for the fast learning case is slightly below the bit level for the slow learning case. While this simulation evidence was at first disturbing, extending the simulation horizon revealed that the equilibrium bit

[15] There are two free parameters for this test. The distance r is measured as a fraction of the standard deviation and was set to a value of 0.5, while for the embedding dimension m, a value of two was chosen.

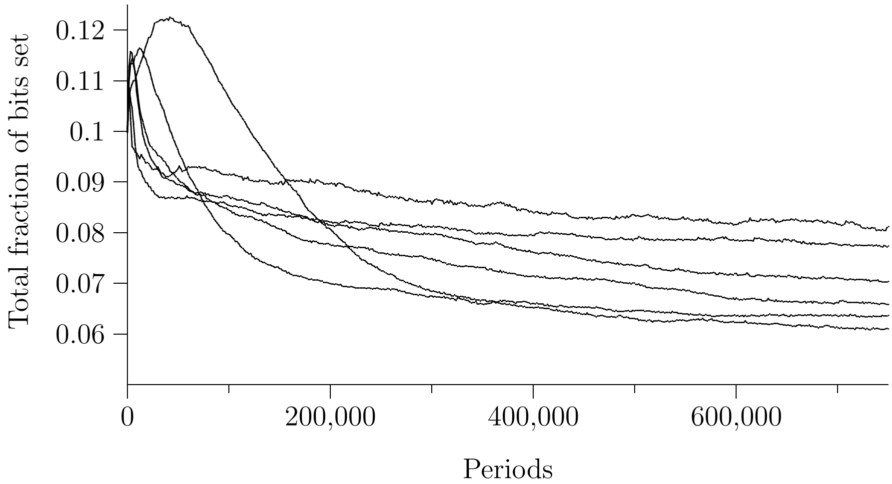


Fig. 6.1. Total fraction of bits set as a function of learning speed in the replicated Java version of the SFI-ASM. Data were averaged over 5 separate runs at different random seeds for a GA-invocation intervals of 1,000 (bottom at period 750,000), 250, 100, 75, 50, and 25 (top at period 750,000).

levels had not been reached at period 250,000. Even after 750,000 periods, some of the curves in figure 6.1 did not settle down to their equilibrium levels. However, the basic statement of the original Objective-C SFI-ASM can be confirmed. There is an increase in the aggregate bit level for faster learning speeds.

When trying to identify the reasons why the adjustment processes towards bit equilibrium obviously take longer than in the original version, one immediately thinks of the increase in checked trading conditions. The original version had a meager amount of 12 conditions (including two dummy bits), while the current Java-version checked a total of 64 trading conditions.[16] It seems logical that when agents have many more opportunities to combine trading conditions to create elaborate trading rules, it will take much longer to test them all. Chances

[16] The hypothesis that an increase in trading conditions leads to an increase in adjustment time was confirmed on two occasions. At first, there were only 57 conditions encoded in the Java-version. The two long integers that represent the condition parts, however, can hold up to 64 conditions. When making full use of this available space, a slight increase in adjustment time was observable. When realizing that the break points in the Objective-C version of Joshi et al. [206] were not optimally chosen, a reassignment of the break points much closer to the one's used by LeBaron et al. resulted in an even higher increase in adjustment time and a higher equilibrium bit level.

are also higher that they end up with a higher fraction of useful trading rules. This could explain the higher equilibrium levels that are attained by the new Java-version. There is, however, one caveat at this point. When starting with different initial bit probabilities, the final equilibrium bit levels are different, too. There seems to be a path dependency hidden in the model structure. The reasons for this path dependency will become clear in the final chapter.

In summing up the behavior under the two learning speeds, Arthur et al. [11] label the slow learning case as the rational expectations regime. The price closely tracks the value predicted by the homogeneous rational expectations equilibrium, and the trading volume is small. The low number of set trading bits implies that the information contained in the condition parts is largely neglected by agents. The fast learning case is characterized by Arthur et al. as the complex or rich psychological regime. The price series exhibit larger deviations from the hree-case such as a larger kurtosis and volatility persistence. Equally important, though, is the increase in the level of set trading bits. Agents appear to create more trading rules that can exploit useful information in the price series.

The difference in bit levels between the two GA-invocation intervals of 1,000 and 250 shown in figure 6.1 does not seem that impressive in the new Java implementation. The increases in bit levels for learning speeds faster than 250 are more apparent. Even though the boundary between the slow and fast learning regime is not a strict one, I prefer to speak of a complex regime for learning speeds equal to or smaller than a GA-invocation interval of 100.

6.8 A Potential Problem: A Biased Mutation Operator

For the creators of the SFI-ASM, the increase in technical trading bits for faster learning speeds pointed to emergent technical trading. Because of the cost they had attached to every non-# bit, they conjectured that emerging trading bits must have, on average, some fitness-based advantages by producing more accurate forecasts. Their intuition guided them to conclude that the classifier systems enabled agents to detect short-term trends in the price series upon which they started to act. At the same time, the price dynamics became more complicated, yet more realistic, and fast learning regimes were labeled as complex. The dependence of equilibrium bit levels on learning speed became one of the main results of the original SFI-ASM and has been replicated by subsequent studies, e.g., by Joshi et al. [206, 207]), and Wilpert [437].

Arthur et al. asked themselves to what extent the existence of the complex regime is an artifact of design assumptions in their model. They found "*by varying both the model's parameters and the expectational-learning mechanism, that the complex regime and the qualitative phenomena associated with it are robust. These are not an artifact of some deficiency in the model*" [11, p. 35].

However, there seems to be a decisive influence of another model parameter on the simulated outcome that has been neglected so far. When varying the rate of mutation in the SFI-ASM, one notices a strict dependence of the equilibrium bit level.

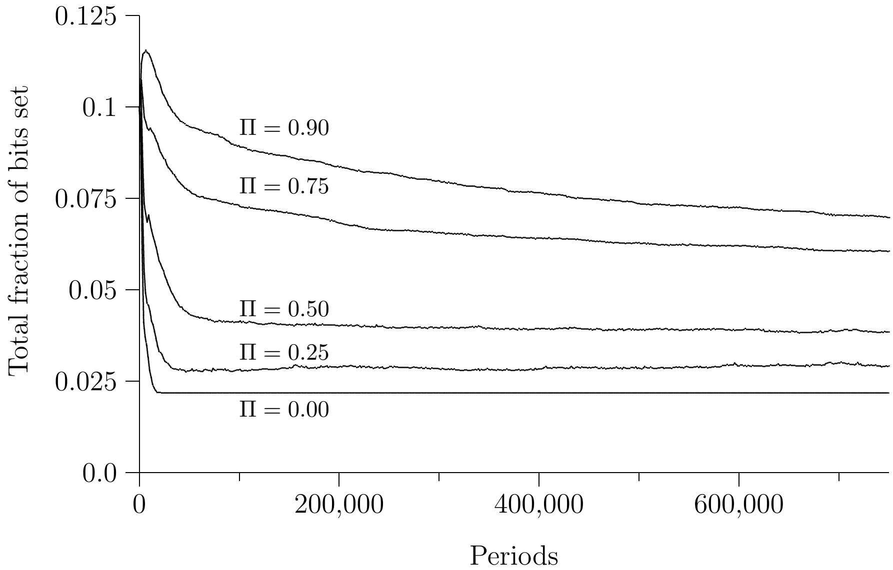


Fig. 6.2. Total fraction of bits set as a function of mutation probability Π in the Java-version of the SFI-ASM. Data were obtained from a cross-section of 5 separate runs at different random seeds for a GA-invocation interval of 100.

It is apparent from figure 6.2 that the level of bit usage in the model does not only depend on learning speed, but also on mutation probability.[17] It should thus be crucial to investigate why increasing

[17] In figure 6.2, the predictor mutation probability Π was adjusted. A similar effect could have been achieved by changing the bit mutation probability π and holding Π constant. Figure 6.2 nicely demonstrates that the GA converges quickly for no mutation at all. The more agents experiment, the longer it takes them to arrive their equilibrium bit levels.

the mutation probability, and hence, decreasing the rate of crossover, has a bit-increasing effect on the condition parts. Can we really be sure that the emergence of the complex regime is not an artifact of design assumption in the SFI-ASM? Postulating emergent technical trading due to faster mutation rates would, after all, be hard to justify. A technical explanation based on model design seems more appropriate to explain this behavior.

Increasing the rate of mutation in the SFI-ASM model means decreasing the rate of crossover. That is, the equilibrium bit level could also be affected by either the crossover operator or the mutation operator. Recall, however, that the uniform crossover operator chooses an offspring's bit with equal probability from the corresponding bit positions of either one or the other parent. Hence, the offspring's fraction of bits set is an unweighted average of the two parents' bit fractions, thus, no systematic influence on the resulting bit level through the working of the crossover operator can be stated.

The picture is quite different for the mutation operator though. In order to analyze its influence on the bit dynamics, we have to look again at the matrix of bit transition probabilities

$$P = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}. \tag{6.18}$$

LeBaron et al. have asserted that these transition probabilities would, on average, maintain the specificity, i.e., the fraction of #'s in a rule. However, a Markov chain analysis reveals that this statement is not true. If we denote the vector of probabilities of the three possible states in period t as $\mathbf{p}^t = \{p_0^t, p_1^t, p_\#^t\}$, in equilibrium $\mathbf{p}^t = \mathbf{p}^t P = \mathbf{p}^{t+1}$ must hold. By repeatedly invoking the mutation operator, the vector of probabilities will converge to its equilibrium distribution of $\mathbf{p}^* = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\}$, i.e., on average, a quarter of all bits will be zero, another quarter will be one, and the remaining fraction of one half will be the don't care sign #. That is, the pure undiluted effect of mutation is to drive the equilibrium bit level towards its fixed point of one half.[18]

[18] An alternative approach to derive the fixed point of one half is to start by denoting the initial fraction of bits set before mutation with $P \in [0, 1]$. The non-# bits are mutated to non-# bits with a probability of $P'_{|0,1\mapsto 1,0} = \frac{1}{3}P$, while the probability that a #-bit is mutated to either 1 or 0 equals $P'_{|\#\mapsto 0,1} = \frac{2}{3}(1 - P)$. Thus, for any given P, the fraction of bits set after mutation is determined by adding the two probabilities above, i.e., $P'_{|0,1,\#\mapsto 1,0} = \frac{1}{3}(2 - P)$. Since $P'_{|0,1,\#\mapsto 1,0} \in [0, 1]$ is a continuous function for all $P \in [0, 1]$, we know by a fixed point theorem that

This analysis shows that the mutation operator is not neutral to the initial level of non-# bits. Because the model usually functions well below this level, the mutation operator introduces an upward bias in the bit distribution. Consequently, when increasing the learning speed, the mutation operator is invoked more often per time period, and its upward bias results in a higher equilibrium bit level.

The influence of the mutation operator on the equilibrium bit level probably went undiscovered so far because the theoretical level of one half is usually far from being attained. The key parameter that exerts a downward influence on the bit level is the fractional rule replacement through the GA. Usually, only a fraction of trading rules, typically one fifth, is changed by the GA. Another downward pressure is brought about by the positive bit costs which causes the GA to favor rules with low specificity for reproduction.

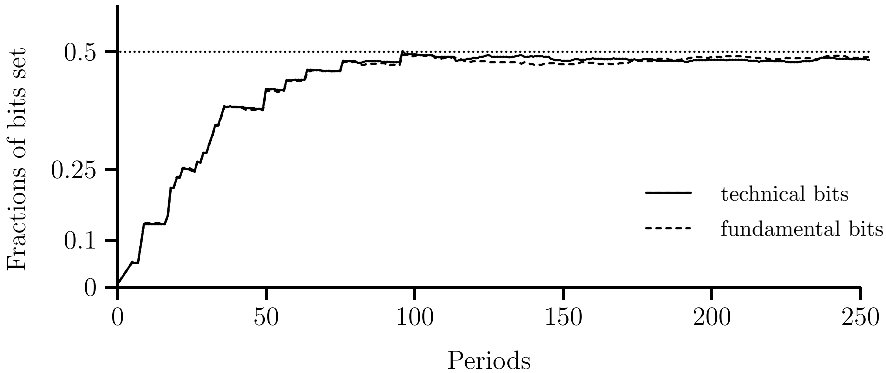


Fig. 6.3. Fixed point convergence of trading bits in the SFI-ASM for $\Pi = 1.0$ and $\pi = 1.0$. Further parameter settings were 25 agents, a GA-invocation interval of 10, no bit costs, and a complete generational replacement of all trading rules.

---

an equilibrium exists. By repeatedly invoking the mutation operator, $P'_{|0,1,\# \mapsto 1,0}$ converges to its equilibrium value of one half. This result also holds if every bit in the bitstring is mutated with a probability of less than one. In the SFI-ASM, this mutation probability π is set at 0.03. Deriving the fraction of non-# bits in the same manner as above yields $P'_{|0,1,\# \mapsto 1,0} = \frac{2}{3}\pi(1-P) + P\left(1 - \frac{2}{3}\pi\right)$. It is easy to check this formula by setting $\pi = 1$, which will yield equation (18), or by setting $\pi = 0$, which will yield $P'_{|0,1,\# \mapsto 1,0} = P$ since no mutation will ever be performed. The equilibrium level equals 1/2, even though convergence occurs a little bit slower than before.

When using a complete generational replacement of trading rules and setting bit costs to zero, figure 6.3 shows that the equilibrium bit level quickly converges to its theoretical level of one half. To speed up convergence, predictor mutation probability Π and bit mutation probability π are set at their maximum values of 1.0, and the GA-invocation interval has been set to a very high rate of 25. For the fastest possible learning rate of GA=1, the convergence to one half would occur immediately.[19] In the long run, the generalization procedure also introduces a slight downward pressure on the bit level by converting some of the 0 and 1-# bits in illogical rules to #-bits.

In light of the upwardly biased mutation operator, one has to be cautious about the claim by LeBaron et al. that positive bit costs would ensure that each surviving trading bit contains useful information. Because of the constant upward pressure on the bit level, the simple existence of trading bits does not necessarily mean that they are useful. LeBaron's argument would imply that the model could be forced into a zero-bit solution if bit costs were only high enough. It can be shown, however, that there will always be some trading bits in the rule sets of agents, no matter how big the associated bit costs are.

[19] Upon closer inspection, however, the mean bit levels seem to hover slightly below the value of half. This either points to an unidentified parameter that also affects the equilibrium level, or a minor plus/minus one programming problem when determining the bit fractions in the model. A careful analysis of the source code, however, leads me to believe that the latter is highly unlikely.

7

A Suggested Modification to the SFI-ASM

An expert is a man who has made all the mistakes which can be made in a very narrow field.

Niels Bohr

7.1 Introduction

Along with complex price series behavior, the endogenous appearance of technical trading bits for faster learning speeds was one of the most striking results of the SFI-ASM. However, the theoretical and experimental analysis in the previous chapter has shown that the interpretation of this increase in trading bits as emergent technical trading may be a design artifact caused by a biased mutation operator. Announcing emergent technical trading simply because the number of trading bits increases for faster learning is premature when a biased mutation operator injects them at increasing rates.

This chapter, therefore, suggests a modification to the SFI-ASM by developing an unbiased mutation operator. The following simulations show that the results are drastically altered with respect to bit levels. All agents now seem to discover the hree-solution of non-bit usage, suggesting that the classifier system does not provide any useful information in terms of improved forecast predictions. Since the finding of zero-bit usage in equilibrium is completely different from that of the original SFI-ASM, the remainder of this chapter then focuses on checking the reliability and stability of the zero-bit solutions. Finally, a rule consistency check as an additional improvement over the original SFI-ASM is introduced.

7.2 An Unbiased Mutation Operator

In order to derive valid conclusions about the bit usage in the model, one should take care in designing bit-neutral operators and procedures. Bit-neutral refers to the feature of leaving the fractions of set bits unaltered, unless an impact is explicitly desired. The bit decreasing effect of the bit cost parameter, for instance, is desirable as it is a fitness-based influence. The bit-increasing effect of the SFI mutation operator, on the other hand, seems problematic since it is completely technical and economically uninterpretable.

The suggested alternative bit-neutral mutation operator works with dynamically adjusting bit transition probabilities. In order to infer whether technical and fundamental bit usage differs in the stock market, this mutation operator works separately for fundamental and technical trading bits.[1] Therefore, it is necessary to distinguish between the initial fraction of fundamental bits set $F_{fund.}$, and the initial fraction of technical bits set $F_{techn.}$. The transition matrix for the fundamental bits is then given by

$$P_{fund.} = \begin{pmatrix} 0 & F_{fund.} & 1 - F_{fund.} \\ F_{fund.} & 0 & 1 - F_{fund.} \\ \frac{1}{2}F_{fund.} & \frac{1}{2}F_{fund.} & 1 - F_{fund.} \end{pmatrix}, \tag{7.1}$$

and, similarly, for the technical bits by

$$P_{techn.} = \begin{pmatrix} 0 & F_{techn.} & 1 - F_{techn.} \\ F_{techn.} & 0 & 1 - F_{techn.} \\ \frac{1}{2}F_{techn.} & \frac{1}{2}F_{techn.} & 1 - F_{techn.} \end{pmatrix}. \tag{7.2}$$

It is easy to verify that these transition matrices ensure that $F^t_{techn.} = F^{t+1}_{techn.}$ and $F^t_{fund.} = F^{t+1}_{fund.}$, i.e., the fractions of set bits remain, on average, unaltered. Unlike the original SFI mutation operator, there is no built-in attractor towards which the resulting bit distribution converges. Without an artificial upward bias, surviving trading bits should only emerge through competition and fitness considerations, implying that they indeed contain useful information.

While the predictor mutation probability Π and crossover probability $1-\Pi$ in the stock market model remain unchanged, the effective rate

[1] Theoretically, there is no clear distinction between fundamental and technical trading bits, neither of which have any use in a rational expectations equilibrium. One reason to treat them differently is motivated by the SFI-ASM, which explicitly distinguishes between these two types of trading bits. Note that the possibility of divergent bit behaviors does not create a bias for it.

with which an individual bit in a trading rule is mutated is constantly adapted such that the expected number of set trading bits before mutations equals the expected number of set trading bits after mutation. Self-adapting operator rates in a GA have recently been investigated by Gomez who, however, based the adaption mechanism on the performance achieved by the offspring [157]. If a genetic operator such as mutation produced fitter offspring than those of another operator such as crossover, its invocation rate was subsequently increased. Gomez found such self-adapting operator rates quite efficient for a variety of problems.[2]

When analyzing the properties of the new mutation operator, two points in the possible bit distribution deserve attention. Once the bit level has reached zero, the updated mutation operator will become inactive. A state j that may never be left again once it is entered is known in Markov chain analysis as an absorbing state. Its transition probability is $p_{jj} = 1$. At a bit level of one, a 1-bit will always be changed to zero and vice versa, but never to the don't care sign #. A state j at the one-bit level thus has an alternating counterpart, both of which are repeated after two mutation steps. Therefore, the zero and one bit levels are limiting distributions in the sense that once agents have arrived at either of them, they will be stuck there forever. However, these corner solutions are not attracting states in that the bit distribution is not torn towards either of them. Because of these two limiting distributions, one could conceive of imposing a minimum and maximum bit level when determining the bit transition matrices, e.g.,

$$F_{fund.} = \min\left(\max(\overline{F_{fund.}^{min}}, F_{fund.}), \overline{F_{fund.}^{max}}\right) \tag{7.3}$$

for an agent's fundamental bit level with $0 < \overline{F_{fund.}^{min}} < \overline{F_{fund.}^{max}} < 1$. To accentuate the following results, however, this idea was not implemented.

7.3 Simulation Results with the Modified SFI-ASM

7.3.1 Trading Bit Behavior

The main difference in the model behavior is shown in figure 7.1. No matter what GA-invocation interval is used, most agents choose to give

[2] Another similarity between Gomez' approach and the mutation operator defined by 7.1 and 7.2 is that they both work on the level of individual agents, i.e., they employ the same strategy of decentralized parameter control.

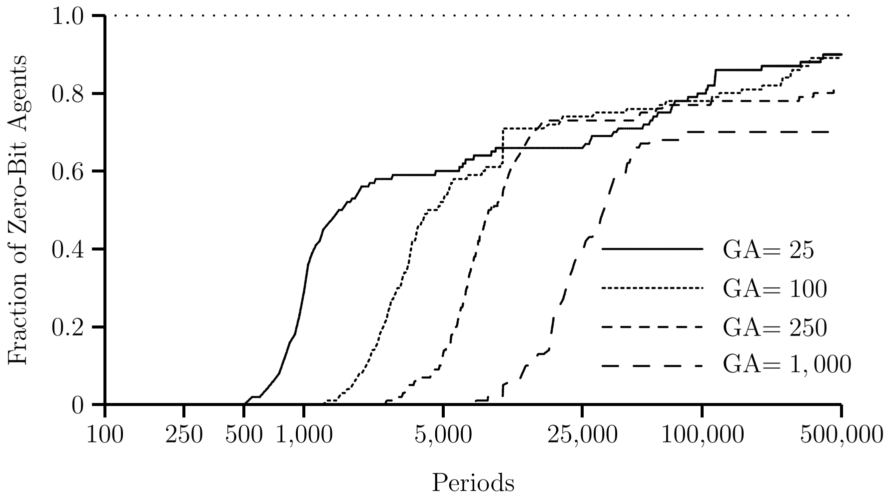


Fig. 7.1. Fraction of bit-neutral agents which discovered the correct hree non-bit usage solution (Zero-Bit Agents) recorded for different GA-invocation intervals and averaged over 10 simulation runs. (25 agents with $\lambda = 0.3$, bit cost = 0.01, and initial bit probability of 0.01.)

up the use of their classifier system and neglect any information provided by it. Since homogeneous agents in a rational expectations equilibrium would be characterized by a total neglect of trading information provided by the classifier system, the Marimon-Sargent Hypothesis is finally supported with respect to bit behavior. When replacing the original mutation operator with an updated operator that has the desirable property of being bit-neutral, most agents seem to realize that, under the given dividend process, all they need for their forecast production is the last period's price and dividend information. Given the biased mutation operator in the original SFI-ASM, emergent technical trading bits thus seem indeed to be a design artifact and not a reflection of ongoing technical trading in the market. Changing the mutation operator appears to be a relatively small change in the design of the artificial stock market. It leads, however, to radically different results with respect to bit behavior from the original model. I will henceforth refer to this version as the modified SFI-ASM and to the agents as bit-neutral agents.

Bit-neutral agents who sooner or later arrive at the zero-bit solution do so by using smaller and smaller mutation rates. Since Fogarty [130] has shown that decreasing mutation rates may improve the performance of a genetic algorithm, it is conceivable that the dynamically adjusting bit transition probabilities of bit-neutral agents assist them in zeroing in on the optimal hree-solution. However, since a zero-bit agent is characterized my no mutations at all, a baseline mutation rate greater than zero is highly recommended. Only if mutation remains an option for agents are they able to react to structural breaks such as a change in the dividend process. DeJong has suggested using 1/(number of bits in the population) as a minimum mutation rate [91].

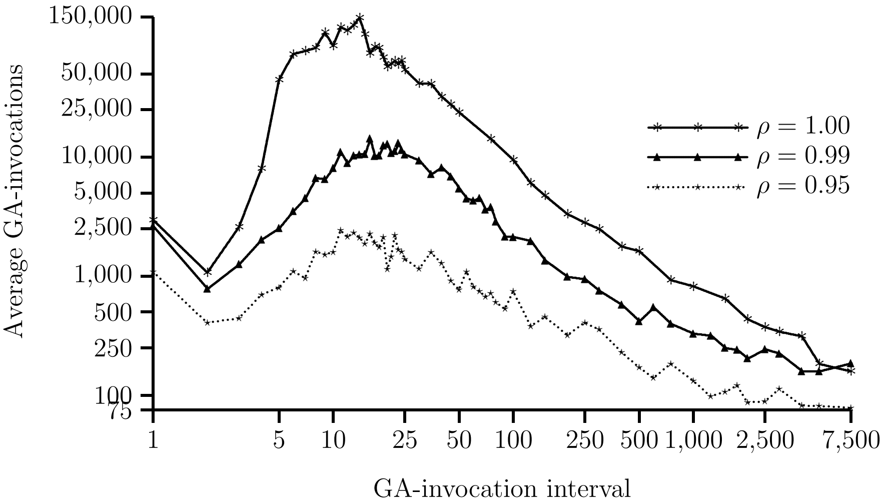


Fig. 7.2. Number of GA-invocations needed until 80% of bit-neutral agents have reached the zero-bit level for the fundamental bits, recorded for different speeds of mean reversion for the dividend process. It was averaged over 25 simulation runs with 30 agents, $\lambda = 0.3$, and bit cost = 0.01.

Depending on the characteristics of the dividend process, however, agents may experience problems in discovering the zero-bit solution. Figure 7.2 plots the average number of GA-invocations needed until the fundamental bits of 80% of the agents have reached the zero-bit level. This was done for several speeds of mean reversion of the dividend process.[3] First, one notices a pronounced hump in the curves

[3] The behavior of technical trading bits is qualitatively similar. However, they generally seem to arrive earlier at the zero-bit level. These results, however, were derived with an earlier Java-version of the SFI-ASM which encoded four technical bits less than fundamental bits, which could account for this difference. Because of the enormous computational costs—each data point in figure 7.2 was averaged

for GA-invocation intervals between 3 and approximately 250. These small invocation intervals seem to hamper the ability of the agents to discover the long-term randomness of what they first believe to be regular patterns. More interestingly, though, is that figure 7.2 clearly demonstrates that the closer the mean-reverting dividend process gets to a random walk, the more difficult it is for the agents to arrive at the zero-bit level.[4] It implies that pure random asset behavior in real financial markets is much harder to detect than more predictable asset behavior. Investors who keep testing the vast amount of possible trading rules may simply be too impatient to discard non-profitable strategies or lack the necessary memory to do so. In real financial markets, the constant departure of experienced traders and arrival of new traders might lead to constant out-of-equilibrium behavior of markets. Given the average lifetime of an investor, figure 7.2 suggests that any practical learning speed could put us in the region where it is extremely difficult to discover the randomness of asset prices.

7.3.2 Time Series Properties

A typical neoclassical rational equilibrium solution would not only be characterized by a total neglect of any additional information contained in condition bits, it would also satisfy the second part of the Marimon-Sargent Hypothesis by exhibiting "nice" price series properties. One should keep in mind that the proposed change to the GA only affects the condition part and not the real-valued forecast parameters of a trading rule. Thus, one would expect the two models to produce similar time series, that is, "well behaved" ones for the slow learning case and more complicated ones for faster GA-invocation intervals.

This hypothesis was tested by comparing the statistical properties of the price time series with those of the Java SFI-version from table 6.4 on page 105. Remember that in hree-mode, the dividend and market price are a linear function of their first order lags, i.e.,

$$p_{t+1} + d_{t+1} = a(p_t + d_t) + b + \epsilon_t. \tag{7.4}$$

---

aver 25 simulation runs which sometimes lasted several million periods—, the simulations were not repeated when the model was upgraded to utilize all possible 32 bit locations in the two condition parts.

[4] There is, however, no monotone relationship between speed of mean reversion and required time to find the zero-bit solution. For very small ρ, the dividend process is very close to a white noise process, and agents again experienced difficulties in detecting this randomness.

The estimated residual time series $\hat{\epsilon}_t$ are then analyzed to see whether they satisfy being i.i.d. and N(0,4) distributed. The results are summarized in table 7.1.

Table 7.1. Time series comparison of the modified SFI (M-SFI) and the replicated Java SFI-ASM with the original mutation operator (J-SFI). Means over 25 runs. Numbers in parentheses are standard errors estimated using the 25 runs. Numbers in brackets are the fraction of tests rejecting the no-ARCH or iid-hypothesis for the ARCH and BDS tests, respectively, at the 95% confidence level.

Description	GA 1,000		GA 250		GA 20	GA 1
	M-SFI	J-SFI	M-SFI	J-SFI		
Std. Dev.	2.084 (.009)	2.145 (.010)	2.141 (.013)	2.225 (.015)	2.229 (.013)	3.397 (.034)
Excess kurtosis	0.004 (.009)	0.085 (.015)	0.001 (.001)	0.285 (.033)	0.050 (.011)	9.046 (1.56)
ρ_1	0.011 (.002)	0.031 (.008)	0.014 (.002)	0.025 (.012)	0.029 (.001)	0.491 (.006)
ARCH(1)	2.610 [0.20]	3.502 [0.60]	2.754 [0.40]	25.34 [1.00]	5.722 [0.48]	1,871.9 [1.00]
ρ_1^2	0.013 (.002)	0.020 (.002)	0.015 (.004)	0.075 (.008)	0.020 (.003)	0.425 (.017)
BDS	1.06 [0.20]	1.41 [0.32]	1.10 [0.24]	34.08 [0.92]	1.44 [0.28]	38.63 [1.00]
Excess return	1.52% (.02%)	2.92% (.03%)	1.59% (.03%)	3.25% (.08%)	1.51% (.03%)	25.34% (3.41%)
Trading volume	0.244 (.008)	0.364 (.025)	0.271 (.007)	0.854 (.065)	0.876 (.009)	1.359 (.015)

First of all, one notices that the modified SFI-ASM produces time series that are usually closer to the hree-benchmark than those of the original SFI-ASM. The standard deviations in the residuals are generally smaller, thus indicating less price variability. Excess kurtosis is almost negligible for both the fast and slow learning cases, which does not line up very well with empirical fat tailed return distributions. Yet when further enhancing the learning speed, both the increase in standard deviation and excess kurtosis suggest that the modified SFI-model shifts into a more complex regime for faster learning rates than the original SFI-model suggests.[5] The autocorrelation in the residuals, as shown in the third row, demonstrates that there is little linear structure remaining except for the extreme case of updating the rule set in every period. As LeBaron et al. indicate, any artificial stock market should exhibit negligible autocorrelations since they are very low for real markets. The large autocorrelation coefficient for GA=1 indicates that the economic structure of the model might break down at this speed, i.e., equation (7.4) becomes misspecified.[6]

The next row reports the means of the test statistics for the ARCH test proposed by Engle [110]. There is considerably less ARCH dependence in the residuals for the modified SFI-version. It is interesting to note that even for very small GA-invocation intervals, some test runs are not able to reject the no-ARCH hypothesis. Only for a GA-invocation in every period can extreme ARCH-behavior for all test runs be observed.[7] In row five, the first order autocorrelation of the squared residuals is another test for volatility persistence. Again, it increases for faster learning speeds but is generally lower than for the SFI-case.

The BDS test in row six is a test for nonlinear dependence developed in [58]. Its test statistic is asymptotically standard normally distributed under the null hypothesis of independence.[8] One can notice an increasing amount of nonlinearities for faster exploration rates, yet again it is substantially lower for the modified SFI-version. Trading volume, which should be zero in the hree-case, increases significantly for faster learning speeds. This points to a greater degree of heterogeneity between the agents.

Overall, the original conclusion that the learning speed affects the price series behavior can still be confirmed after the proposed change. However, it is also apparent that for identical GA-invocation intervals, the modified SFI-results are generally closer to the hree-benchmark.

[5] While LeBaron et al. have reported the results only for the GA-intervals of 1,000 and 250, two additional learning speeds are included in table (7.1). The statistical tests were performed for even more GA-intervals, in particular, for 100, 50 ,25, 20, 10, 5, 2, and 1.

[6] The autocorrelation coefficient for a GA-interval of 2 with a value of 0.06 (standard deviation 0.0035) is considerably lower than for an invocation interval of one. This supports the hypothesis that there is a structural break in the model when using the fastest possible learning speed.

[7] Even for an invocation interval of two, the no-ARCH hypothesis cannot be rejected for 16% of the test runs.

[8] There are two free parameters for this test. The distance r is measured as a fraction of the standard deviation and has been set to a value of 0.5, while for the embedding dimension m, a value of two has been chosen.

Compatible with these results are the findings of Wilpert [437]. He reports less kurtosis in the residuals and less trading volume when agents have no access to their classifier system in the first place. In the modified SFI-ASM, agents endogenously arrive at neglecting the classifier system and converge to the hree-solution.

7.4 Robustness of the Zero-Bit Solution

7.4.1 Stochastic versus Periodic Dividends and the Classifier System

Since the finding of zero-bit usage by bit-neutral agents contradicts the findings of original SFI-ASM, various tests were designed to check the robustness of this result. First, in order to check the proper working of the classifier system, the model behavior was tested for classifier mode and non-classifier mode. For the latter, agents had no access to condition bits at all, and their trading rules consisted only of prediction parts with associated fitness information. Classifier and non-classifier agents were both confronted with a noiseless periodic dividend stream such as a sine wave or a square wave. Even though the simulated price series tracked the crude risk-neutral price astoundingly well in the non-classifier mode, the tracking behavior in the classifier mode was better for most GA intervals.[9] The bit usage of most classifier agents was characterized by bit positions in their trading rules that were either almost completely set or completely ignored within a relatively short amount of time.

This finding was revealed by figure 7.3. When looking at how often certain bit positions in the whole economy are set, one notices that they are often multiples of 100, which is the number of trading rules that agents possess. One can easily see that the fundamental bit no. 32 has been independently discovered by seven bit-neutral agents, while four agents found technical bit no. 31 to be useful. Many other bit positions were heavily used by either one or two agents. In spite of the periodic

[9] There are, however, some parameter restrictions that may make it tricky for agents to recognize the periodicity of the dividend process. First, the wavelength should be shorter than the parameter τ which controls the time window over which trading rules are evaluated. It should also be shorter than the GA-invocation interval. Furthermore, the forecast parameters in the prediction parts of trading rules can only be changed by the GA within very narrow intervals which were calculated on the basis of a mean-reverting AR(1) dividend process. For a completely different dividend process, the required parameter values may lie outside the allowable interval and cannot be reached by the GA.

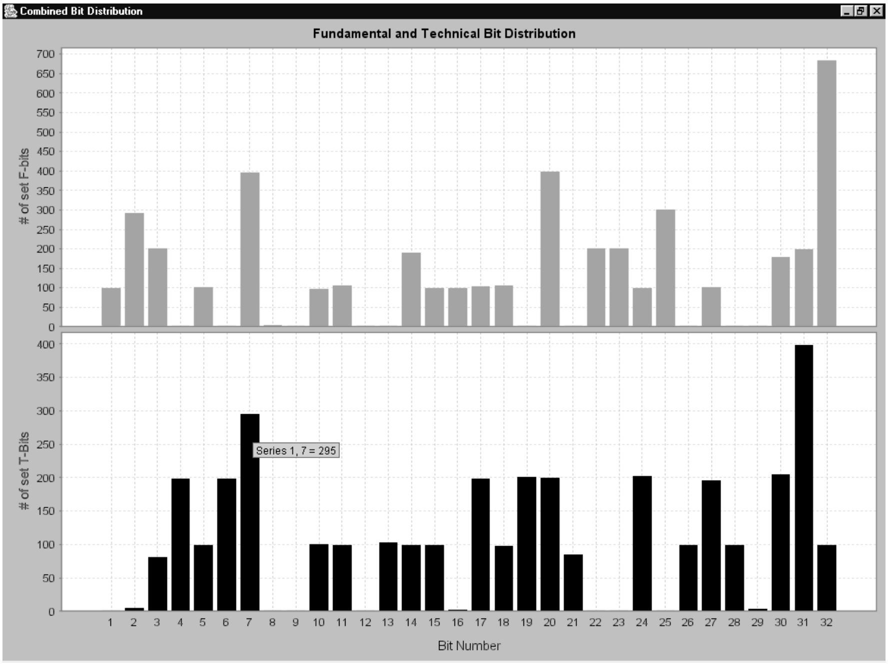


Fig. 7.3. Bit Distribution of 25 bit-neutral agents when confronted with periodic square wave dividends (wavelength = 50, Ga-interval = 250, θ=75).

nature of the dividend process without any added noise, agents became heterogeneous in that most of them decided to use a different technical trading rule.

In summing up one can state that the simulation results under periodic and stochastic dividends show that the classifier system works quite efficiently. When confronted with periodic dividends, it detects these patterns, yet when working with stochastic data, it also discovers the “right” solution of non-bit usage. Even though the mean-reverting dividend process is able to produce short term trends toward its mean, these are by no means regular. Therefore, in the long run, the stochastic nature of the dividend process dominates any (random) short term trends and patterns.

7.4.2 Dependence on Other Parameter Values

The zero-bit solution proved robust under a wide range of parameter combinations. For instance, changing the selection procedure for rule usage (select best or roulette wheel selection) did not alter the main finding of convergence towards the hree zero-bit solution. Increasing the bit cost increased the convergence speed, especially for "latecomer" agents who experience difficulties in finding the correct hree-solution.

Varying the crossover and mutation rates revealed some differences in convergence speed, but did not alter the main result of agents finding the zero-bit solution. In the crossover only case, some zero-bit agents emerged rather quickly. Other agents had some bits set in all their trading rules while the other bits tended to be completely unused. In the extreme case of mutation only with both mutation probabilities set to one, trading bits were uniformly distributed across all agents and their trading rules. Since the mutation only case could be seen as pure experimentation without making use of previously good solutions, it is not surprising that agents gave up their classifier system, too.

The zero-bit solution thus turns out to be robust under wide parameter combinations. The finding that the classifier system in this artificial stock market does not provide any advantage or useful information that agents could exploit is also supported by Wilpert who replaces the fitness evaluation based on forecasting ability with profit generating ability. He, too, finds that giving up the condition parts of the trading rules does not have drastic effects on the model's behavior, and he questions the usual interpretation of bit usage. He emphasizes that its importance should not be overestimated.

7.4.3 Generalization or Consistent Trading Rules?

In section 6.4.3 it was mentioned that the GA in the SFI-ASM regularly produced illogical trading rules that would never be activated. The solution to this problem was to have a large rule set and to generalize rules that have not been matched for more than 4,000 periods by converting some of their 0 and 1-bit to #.

Since the original SFI-mutation operator introduced an upward bias in the bit distribution, it was impossible to assess the degree of downward pressure exerted by the generalization procedure. Even though the zero-bit solution seemed largely unaffected by changes of the generalization parameters, i.e., the maximum number of inactive periods before a rule is generalized and the fraction of bits to be generalized, it cannot be excluded that it is the generalization procedure that is responsible for the zero-bit solution.

To exclude that possibility, a rule consistency had been developed which ensures that only logical trading rules are accepted into an agent's rule set.

Table 7.2. Creating logical trading rules through a consistency check.

$\frac{\text{price x interest Rate}}{\text{dividend}} >$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{3}{2}$
rule 1	#	#	0	1	#	0	#
rule 2	#	#	1	1	#	0	#
$\overleftarrow{\text{rule 1}} \longmapsto$ rule 3	#	#	#	1	#	0	#
$\overrightarrow{\text{rule 1}} \longmapsto$ rule 4	#	#	0	#	#	#	#

An illogical trading rule is rule 1 in figure 7.2. It describes a situation in which $p_t r_f / d_t$ is smaller than 7/8, yet at the same time, that ratio is supposed to be greater than 1. Obviously, rules of this type will never be matched but will get inserted into an agent's rule set. Rule 2, on the other hand, might be considered a corrected version. However, the third rule demonstrates that the same information can be coded by using less activated bits.

The consistency check in the modified SFI-ASM is thus designed not only to create logical trading rules, but also to express corrected bit sequences with the least amount of non-# bits, i.e., one or two. However, depending on whether the checking procedure starts analyzing the bit sequence from the lower bits upward or vice versa, different results will be returned by it. Rule 3 would be returned by the consistency check if it starts analyzing from the higher order bits. Otherwise, the fourth rule would be returned. The consistency check thus chooses both variants with equal probability.

Since the consistency check removes bits from illogical trading rules, one could easily argue that it aggravates a possible downward pressure. The conjecture of a downward bias through the consistency check, however, turns out to be wrong. Surprisingly, figure 7.4 shows that turning the consistency check on increases the level of set trading bits. The key for understanding this astonishing result lies in the way the bit fractions are calculated.

Consider the case without the consistency check. If we denote the number of set fundamental bits with y, the fundamental bit fraction in a trading rule is determined as $y/32$, 32 being the number of all fundamental conditions. With the consistency check, however, the bit fractions are calculated differently. Recall that 2 bit sequences of 7 bits are checked in the fundamental condition word of 32 bits. Within each bit sequence, at maximum, 2 bits can be set. The maximum bit fraction of 100 % would be achieved if all 18 unchecked bits are set, in addition to the 4 bits in the two checked sequences. Instead of 32 bits, the denominator for the calculation of bit fractions is then $18 + 4 = 22$

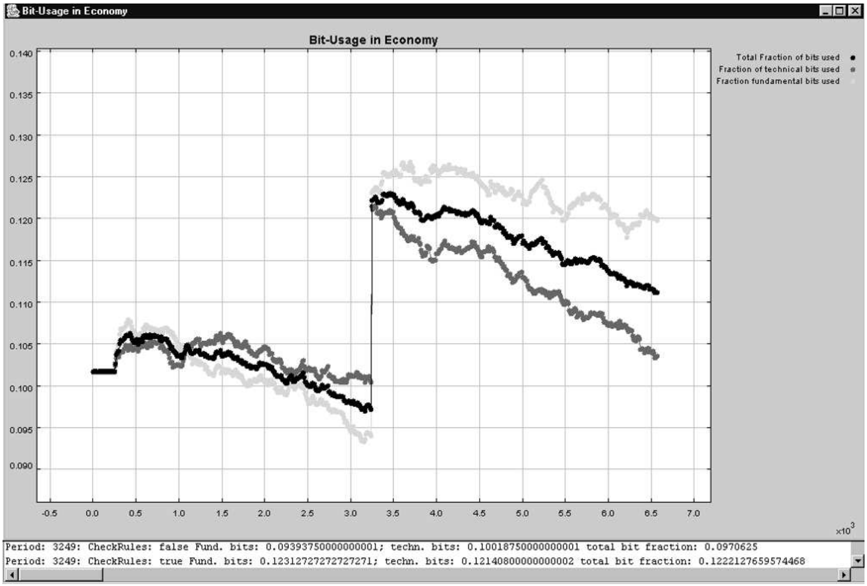


Fig. 7.4. Demonstration of the upward pressure on the bit distribution when switching the consistency check on in mid-simulation.

bits. Using a smaller denominator increases the bit fraction for a given number of set trading bits. The simple recalculation explains the jump of the bit fractions the moment the consistency check is turned on. But remember that these augmented bit fractions for technical and fundamental bits are then used to determine the dynamically adjusting bit transition probabilities for the updated mutation operator. Since the bit-neutral mutation operator applies these elevated bit transition probabilities on the whole bit string, turning the consistency check on results in a slight increase in the bit distribution in addition to the one time effect shown in figure 7.4.

The major result of agents finding the correct hree-solution are usually replicated with or without the consistency check. It is probably a useful extension of the SFI-ASM since it allows agents to focus on creating and testing only logical trading rules. From a computational perspective, it could even help save computational time since agents do not need a huge rule set to work with.

8

An Analysis of Wealth Levels

If you're so smart, why aren't you rich?

8.1 Introduction

The previous chapter has presented evidence that fundamental and technical trading bits do not improve the forecast accuracy of agents' trading rules when confronted with stochastic dividend data. As a result, agents abandon their classifier system in the long run. Unless a rule's forecast accuracy is a wrong proxy for its profitability, the convergence at the zero-bit solution should imply that classifier agents, i.e., those who use the information provided by their classifier system, cannot outperform agents who neglect the classifier system altogether. This is in contradiction to studies by Joshi et al. [206, 207]) who found that agents with access to technical trading bits acquire more wealth than fundamental traders.

This chapter first briefly introduces the two studies by Joshi et al. and presents the explanations given for the observed differences in wealth levels. However, a sensitivity analysis of wealth levels reveals that things are more complicated than previously thought. Wealth differences between different trader types—technical or fundamental traders, fast or slow learning agents, bit-neutral or unmodified SFI agents, classifier or non-classifier agents—generally arise. While the explanation of these wealth differences has been sought by Joshi et al. through elaborate economic reasoning, the following analysis identifies the number of active trading rules and the chosen selection mechanism as main determinants for varying terminal wealth. Additionally, two benchmark levels are developed with which the absolute wealth levels attained by traders can be compared. This chapter concludes by stating that the economic structure of the SFI-ASM makes it unsuitable to give an economic interpretation to wealth effects.

8.2 Wealth Levels in the SFI-ASM: An Economic(al) Explanation

Since agents in the modified SFI-ASM endogenously give up the use of their classifier system, it is strongly implied that it does not provide any profitable trading information. If technical trading rules were to generate excess profits, technical trading should flourish rather than vanish.[1] The conclusion of a useless classifier system is in direct contradiction to Joshi et al. [206], who find that agents with access to technical trading bits acquire more wealth than fundamental agents. The wealth differences that they report are not negligible and grow over time, as figure 8.1 vividly illustrates.

When trying to explain the source of these wealth differences, Joshi et al. [206] offer some vague economic rationalization. They argue that the use of technical trading rules creates a negative externality for other agents. If a single agent with access to technical trading bits detects a short term price trend, he might be able to exploit this pattern without it dissipating. If more and more agents detect the price trend, the particular technical trading rules are reinforced through positive feedback, thus making them self-fulfilling prophecies, which can cause bubbles and crashes. The latter are negative market externalities which worsen everyone else's strategies by lowering their forecast accuracy. As a result, stock prices are driven away from their fundamental values. This is a loss of market efficiency and Joshi et al. claim that this would translate into lower average returns and less accumulated wealth.[2]

---

[1] Strictly speaking, the trading rules are not evaluated according to their wealth generating ability, but on their forecast accuracy. While it is possible that a more accurate trading rule is less successful in terms of generated wealth, Arthur et al. [11] believed this scenario to be highly unlikely. Their intuition was confirmed by Wilpert [437], who evaluated trading rules according to their generated profits. His results did not differ much from those of the original SFI-ASM.

[2] Similar arguments on the emergence and profitability of technical trading can be found in [11, 205]. The papers by Arthur et al. [11], Palmer et al. [330], and LeBaron et al. [244] do not discuss any possible wealth effects. In [331], we find the rather vague claim that "*if a trained agents is extracted from the market and then reinserted much later, it tends to do rather poorly.*" "Doing poorly" probably means acquiring less wealth. Furthermore, they report that the wealth distribution evolves into a wide distribution, with some agents being very rich

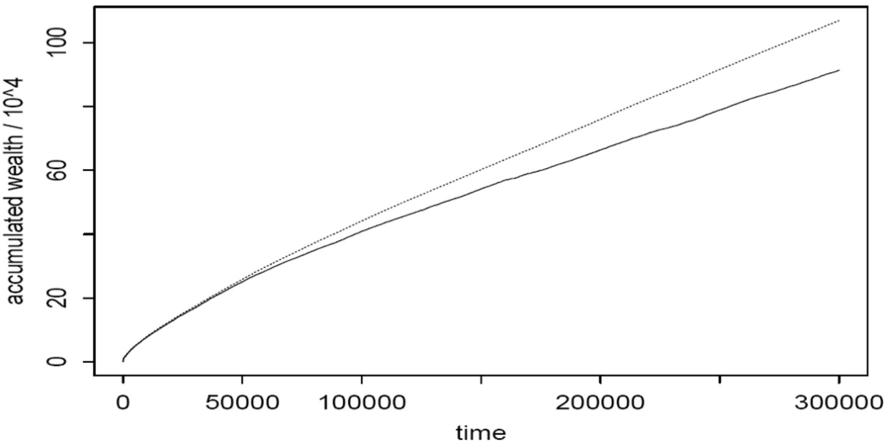


Fig. 8.1. Wealth levels for the situation when one agent includes technical rules while all others exclude them. Note that the singular agent using technical rules accumulates significantly more wealth than those agents using only fundamental rules almost all through the run, and that this difference grows over time. Source: [206, p. 11]

8.3 Previous Studies Based on Wealth Levels in the SFI-ASM

The observed differences in wealth accumulation between fundamental and technical traders were used by Joshi et al. for two game-theoretic analyses which will be briefly introduced in the following two sections.

8.3.1 Financial Markets Can Be at Sub-Optimal Equilibria

The learning rate of agents in the SFI-ASM was treated as an exogenous variable by LeBaron et al. For slow learning rates, the model closely resembled the homogeneous rational expectations equilibrium, but for faster learning rates, the model shifted into a more complex regime that better fits the empirical facts. However, the question of which learning rate would be chosen by agents if given the opportunity to do so was not addressed by them.

Joshi et al. [207] close this gap by combining their simulation results with a game-theoretic analysis. They assume that there are S possible

and some being poor. While the identity of winners and losers changes over time, the distribution stays rather constant.

learning rates that agents can chose from.[3] Given the assumption that $N - 1$ traders in the market have adopted a common, but unknown revision rate $\overline{K}$, a single trader has to find his optimal learning speed K^*.[4] Since all traders in the market attempt to find their optimal responses, this situation constitutes a symmetric simultaneous-move N-person game with a $S \times S$ decision matrix in which a single trader's dominant strategy would be a dominant strategy for all other agents.

The payoffs in the $S \times S$ decision matrix are determined by fixing the base revision rate $\overline{K}$ to one of the S different values, and then recording the terminal wealth levels that the remaining agent can acquire by varying his learning rate K. The information in the $S \times S$ decision matrix can be summarized in a reaction function, which assigns to every $\overline{K}$ the optimal response K^* at which the single trader maximizes his terminal wealth level.

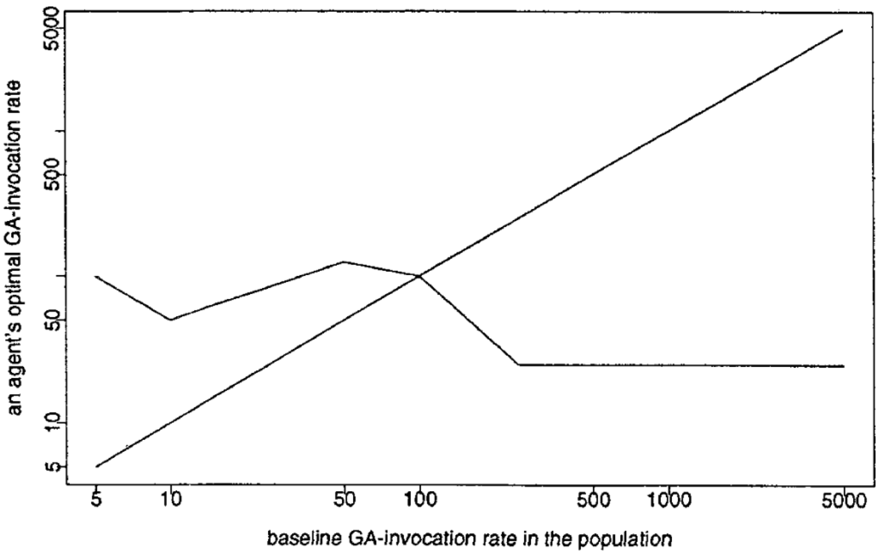


Fig. 8.2. A trader's reaction function, indicating his optimal GA-invocation interval K^*, given the baseline GA-invocation interval $\overline{K}$ used by all other traders. The symmetric Nash equilibrium is where $\overline{K} = K^* = 100$. Source: Joshi, Parker, and Bedau [207, p. 12]

[3] These rates are 5, 10, 50, 100, 250, and 5,000.

[4] For simplicity, Joshi et al. assume that all other agents chose the same single base revision rate $\overline{K}$. They claim, however, that their results do not change much when other agents are allowed to follow mixed strategies.

It is apparent from figure 8.2 that an agent's optimal response matches the base GA-invocation interval $\overline{K}$ only at one point, i.e., where $\overline{K} = K^* = 100$. The single trader cannot gain by deviating from his strategy, thus he keeps updating his rule set at the same frequency as everyone else does in the economy. Since all traders in the market face an identical situation, no one has an incentive to deviate from this learning rate either. $\overline{K} = K^* = 100$ thus constitutes a unique symmetric Nash equilibrium.

Joshi et al. emphasize that this market equilibrium is sub-optimal since it clearly falls into the complex regime as described earlier. Price series are complicated and exhibit a higher variance than the rational expectations regime, making investments riskier. Furthermore, the condition parts of trading rules contain many set fundamental and technical trading bits which requires more resources devoted to creating and evaluating them. Last, but not least, they claim that the wealth levels at the end of the simulation are lower for the complex regime than for the rational expectations regime. Everyone in this Nash equilibrium is worse off than if all agents would pursue more long-term oriented trading strategies. This is a typical prisoner's dilemma situation in which the rational behaviors of individual agents cause the optimal social outcome, the rational expectations regime, to be missed. Joshi et al. therefore conclude that financial markets can operate at such sub-optimal equilibria in which we have significant levels of technical trading.

8.3.2 Technical Trading as a Prisoner's Dilemma

In a similar and related study, Joshi et al. [206] vary their focus and ask whether agents will choose to include technical trading information in their trading rules. Throughout this study, a constant learning rate of $K = 100$ is used. This is done for two reasons: First, at this rate, all agents may use technical trading information since they are in the complex regime. Second, they have shown in their 2002 article that this would be the learning rate chosen by agents.

The framework is slightly modified from the study above in which agents had to choose their optimal learning rate. This time, a single agent faces a classic 2×2 decision problem of whether to include the technical trading information or not, given his assumption that all of the other traders either include them or not.[5] Again, this design resembles a multi-person simultaneous-move game. Thus, the only potential

[5] All traders have access to fundamental trading information though. The exclusive use of technical trading information is ruled out as unrealistic.

Nash equilibria are symmetric in the sense that all agents either include or exclude the technical trading bits.

Table 8.1. "The decision table for an agent contemplating whether to include technical trading rules to make her market forecasts, when she is uncertain whether the other traders in the market are doing so. The agent's payoff in each of the four situations A–D is her expected final wealth (divided by 10^4, to make more readable), derived by averaging the results of 45 simulation runs of each situation. Error bounds are calculated using standard deviations of each set of 45 simulations." Source: Joshi, Parker, and Bedau [206, p. 9]

		all other agents	
		include technical rules	exclude technical rules
single agent	includes technical rules	A: 113 ± 7	B: 154 ± 7
	excludes technical rules	C: 97 ± 7	D: 137 ± 5

The payoffs in table 8.1 are averaged terminal wealth levels.[6] It is obvious from the payoff table, that including technical trading rules in one's rule set is a dominant strategy for a single trader since $A > C$ and $B > D$. Because we can have only symmetric Nash equilibria, the Nash equilibrium that will be selected is the one in which all agents include technical trading information. All other states are unstable.

According to this game-theoretic analysis, it is rational for each agent to include technical trading information, even though it leads to an outcome in which all earn less wealth than as if they all agreed on ignoring technical trading bits. Thus, Joshi et al. [206] conclude that technical trading is a typical multi-person prisoner's dilemma and that financial markets can get locked into such an inefficient equilibrium.

[6] For each cell, Joshi, Parker, and Bedau averaged over 45 simulations, each of which was run for 300,000 periods.

8.4 Wealth Levels in the SFI-ASM: Alternative Explanations

The following sections, however, will provide other answers to the question of what determines the wealth accumulated by agents. Section 8.4.1 analyses the model design and finds that the economic explanations for absolute wealth levels given by Joshi et al. [206] turn out to be conjectures that are not supported by the SFI-ASM design. This theoretical analysis is supplemented by some simulation evidence in section 8.4.2. The question of relative wealth, i.e., why certain trader types become richer than others, is addressed in section 8.4.3.

8.4.1 Risk-Premium, Taxation, and Two Benchmark Wealth Levels

When observing differences in absolute wealth levels due to learning speed or access to certain trading information, a mind trained in economics easily spots some potential for optimization. For instance, Joshi et al. [206] asked themselves whether these wealth differentials could be helpful in finding an agent's optimal learning speed. Because of the specific design of the SFI-ASM, however, wealth levels are somewhat counterintuitive to interpret and optimization attempts might fail.

Agents in the SFI-ASM increase their wealth in two ways. First, the cash they own earns them a risk-free interest income and second, they collect the stochastic dividends on their stock holdings. To avoid a long-run explosion of wealth levels through compound interest effects, however, an agent's wealth is taxed at a rate equal to the exogenous interest rate. The cash and wealth positions of an agent i are computed according to

$$W_{i,t} = C_{i,t-1} + p_t x_{i,t} \tag{8.1}$$

$$C_{i,t} = C_{i,t-1} + r_f C_{i,t-1} + x_{i,t} d_t \tag{8.2}$$

$$C_{i,t} = C_{i,t-1} - r_f W_{i,t}, \tag{8.3}$$

$x_{i,t}$ being the amount of stock held by agent i. Equation 8.1 denotes the previous wealth before adjustment, and equation 8.2 takes interest and dividend earnings into account, while equation 8.3 lowers cash through tax payments at a tax rate equal to the risk-free interest rate. Equations 8.1 – 8.3 can be summarized as

$$C_{i,t} = C_{i,t-1} + x_{i,t} \left(d_t - r_f p_t \right), \tag{8.4}$$

which is the equation with which the cash positions of traders in the model are updated. Note that for risk aversion, the term in parentheses reflects a positive risk premium, and $C_{i,t} > C_{i,t-1}$ holds.

Gulyás et al. [166] point out that in the SFI-ASM, agents usually increase their wealth more or less independent of their actions. It is, therefore, possible to determine a benchmark wealth level under the assumption of inactivity. Agent i would hold onto his one unit of stock, i.e., $x_{i,t} = 1$ for all t, and would take the market price as given. In this case, equation 8.4 would simplify to

$$C_{i,t} = C_{i,t-1} + (d_t - r_f p_t). \tag{8.5}$$

Since wealth is defined as an agent's cash plus the value of his stock position, which is $x_{i,t} = 1$ in this case, the first benchmark wealth level is calculated as

$$W_{i,t} = C_{i,t-1} + d_t + p_t \left(1 - r_f\right) \tag{8.6}$$

and will be referred to as base wealth.

To determine an agent's base wealth, however, it is necessary to have an actual simulation run to have stock prices with which stock positions can be valued. One way to determine an absolute benchmark wealth level without the need for a simulation is to assume that the market is run in hree-mode, i.e., not only one, but all agents are inactive, holding onto their one unit of stock, collecting its dividend, and receiving interest on their cash. An approximation for hree-base wealth can be obtained by substituting the simulated prices and dividends in period t by their theoretical averages $\overline{d}$ and $\overline{p^{hree}}$.[7] Since these are constants, the recursive relationship of equation 8.5 can then be written as a function of initial cash endowment C_0, i.e.,

$$W_t^{hree} \approx C_t^{hree} = C_0 + t\left(\overline{d} - r_f \overline{p^{hree}}\right). \tag{8.7}$$

Since agents are identical, an agent subscript is not needed. In the long run, the value of one unit of stock becomes negligible in comparison to the cash value, hence, the theoretical hree-base wealth level W_t^{hree} in period t is reasonably well approximated by hree-base cash. Again, risk aversion implies that $\overline{p^{hree}} < \overline{d}/r_f$, hence, 8.7 states that hree-base wealth W_t^{hree} grows linearly in t and that the only source for

[7] The hree-price was derived in section 6.5. It is a linear function of agent's risk aversion, the risk-free interest rate, the theoretical dividend mean, speed of mean reversion of the dividend process, and the variance of the dividend noise process.

wealth accumulation in the SFI-ASM are the risk premiums that agents collect.[8]

Table 8.2. Some numerical examples for hree-prices and hree-base wealth levels at different periods. Other parameter values were $\rho = 0.95$, $r_f = 0.1$, $\bar{d} = 10$, and $C_0 = 20,000$.

risk aversion	$\overline{p^{hree}}(\bar{d} = 10)$	$W^{hree}_{t=250,000}$	$W^{hree}_{t=300,000}$
$\lambda = 0.0$	100.00	20,000	20,000
$\lambda = 0.3$	88.00	320,000	380,000
$\lambda = 0.5$	80.00	520,000	620,000

It is important to realize at this point that base wealth usually exceeds hree-base wealth since the long-term average of simulated prices p_t is smaller than the theoretical hree-price average. This reflects an additional risk premium since the constant learning of agents introduces some extra noise. Equation 8.7 also implies that an increase in the risk premium due to higher volatility, reflected in smaller values for p_t, will result in higher wealth levels in the SFI-ASM. It is easy to see that wealth could be maximized if stock prices would be zero. Contrary to the statements made by Joshi et al. [206], an efficient market outcome in the SFI-ASM is thus not characterized by high wealth levels, but quite the contrary. Considering that simulated prices are usually below the hree-prices and seldom overshoot, the hree-benchmark equilibrium constitutes a lower bound for non-hree wealth levels.

8.4.2 Average Stock Holdings and Wealth Levels

The previous section only focused on the problem of aggregate wealth. It found that when the risk premium in the stock market increases, all agents are able to acquire more wealth. The question as to where the wealth differences come from between technical and fundamental traders or fast and slow learners reported by Joshi et al. [206, 207] was not addressed. Given that different trader types in a particular simulation face the same risk premium, the question remains why technical

[8] Another reason why equation 8.7 is only an estimate is that it considers the theoretical averages for prices and dividends and not the empirical averages for a particular simulation run. In the long run, however, these averages converge to their theoretical means. The exact hree-base wealth in t for an actual simulation run would be $W^{hree}_t = C_0 + t(\bar{d} - r_f \overline{p^{hree}}) + p^{hree}_t$, where $\bar{d}$ and $\overline{p^{hree}}$ now represent the empirical averages.

traders accumulate different wealth levels than fundamental traders. Since we have established in chapter 7 that agents endogenously give up the use of their classifier system, it seems highly unlikely that there are possible gains from technical trading. It is implied that the patterns agents might temporarily detect are random and not profitable in the long run.

When investigating the possible impact of the classifier system on wealth accumulation, it seems reasonable to introduce another trader type. Since most bit-neutral agents end up not using their classifier systems in the long run, it seems obvious to shorten the adjustment process and to have non-classifier agents to begin with. Non-classifier agents do not check any environmental information at all and have trading rules that only consist of prediction parts and fitness information. How would these agents compare if they had to compete with old-fashioned SFI-classifier agents?

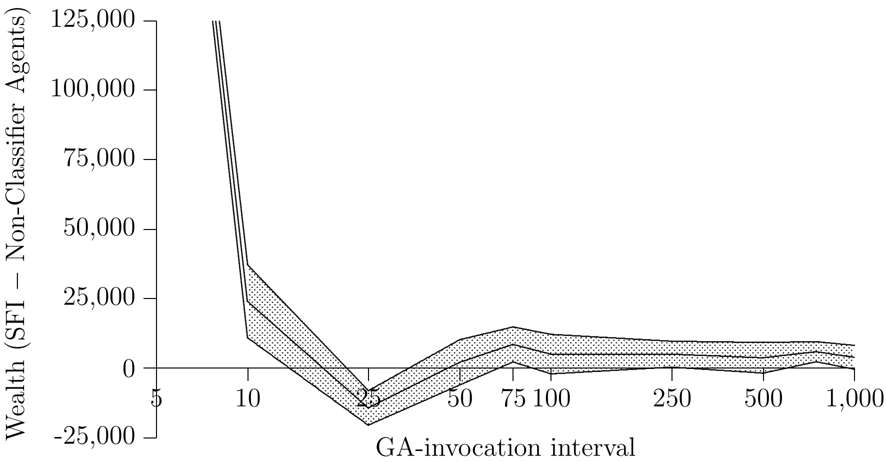


Fig. 8.3. Wealth differences between SFI and non-classifier agents, both trader types using select best for rule selection, recorded for different GA-invocation intervals and averaged over 25 simulation runs. Positive values indicate that SFI agents outperform the non-classifier agents. The shaded area indicates mean wealth difference ± one standard deviation.

It was my initial hypothesis was that non-classifier agents should never outperform classifier agents since only the latter were equipped with an additional tool to analyze and exploit potentially useful information.[9] Figure 8.3, however, tells a different story. While for most GA-invocation intervals the classifier agents acquire more wealth, there is a significant dip in the curve when non-classifier agents do better than their SFI-counterparts. Better performing non-classifier agents, however, are a clear indicator that reasons other than pattern recognition and exploitation are responsible for differences in accumulated wealth levels.[10]

Since the risk premium is the only source for wealth accumulation in the model, one obvious conclusion is that wealthier trader groups hold, on average, more risky stocks. Figure 8.4 verifies this hypothesis by plotting the acquired wealth levels of individual SFI- and non-classifier traders as a function of their average stock holdings over their entire lifetime.[11]

Besides the highly linear relationship between average stock holdings and final wealth, figure 8.4 also shows that the non-classifier agents, as a group, hold more stock than SFI agents. When regressing individual wealth W_i on average stock holdings x_i,

$$W_i = a_0 + a_1 x_i, \tag{8.8}$$

the regression coefficients as shown in table 8.3 are highly significant.

Table 8.3. Regression results for the linear regression $W_i = a_0 + a_1 x_i$

regr. coefficient	std. error	t-value	coeff. of determination
$a_0 = 85,380$	7,400	11.5	0.9506
$a_1 = 223,342$	7,350	33.4	

[9] That hypothesis was posited before I actually established that risk premiums are the only source of possible wealth accumulation in the SFI-ASM.

[10] Initially, I doubted the validity of the simulation results in [206]. In [105], I reported that wealth levels of classifier and non-classifier agents rise equally after some initial divergence during the warm-up phase, which is in contradiction to figure 8.1 with diverging wealth slopes. I then noticed, however, that while the absolute wealth differences were tiny, the classifier agents did slightly better in 23 out of 25 simulation runs. A paired two sided t-test then revealed that these minuscule wealth differences were significant, sometimes even at the 1% percent level. It was this outperformance of the classifier agents that prompted me to look for alternative explanations as to how wealth differences in the SFI-ASM come about.

[11] The specific parameter combination was chosen in order to clearly visualize that the two trader groups hold, on average, different amounts of stock.

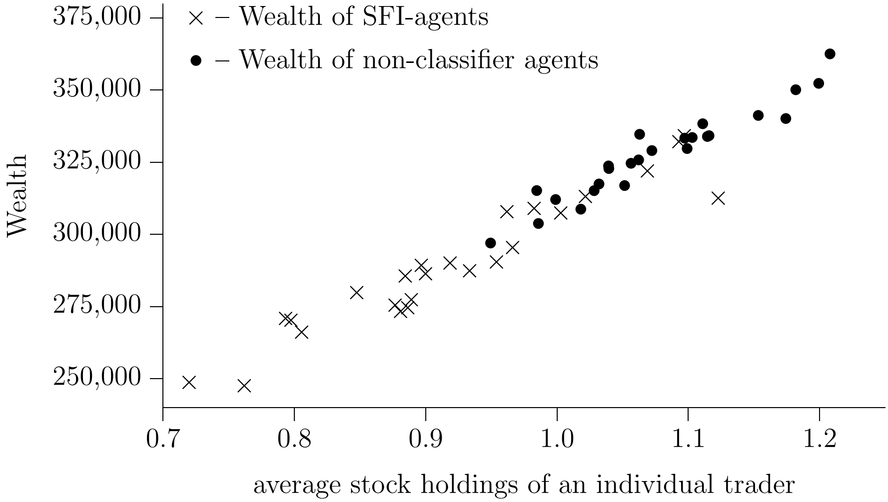


Fig. 8.4. Final wealth levels of SFI- and non-classifier agents as a function of average stock holdings (period 250,000; GA-interval = 250; risk aversion =0.3; 10 trading rules; 25 SFI- and 25 non-classifier agents).

The coefficient of determination indicates that more than 95% of wealth variations are explained through average stock holdings. While this seems like a reasonably high power of explanation, the remaining five percent of unexplained wealth differences could be, in principle, attributed to profitable pattern recognition and exploitation, perhaps of the "buy low and sell high" type.

However, equations 8.4 and 8.7 lead me to conclude that the ***only*** source of wealth accumulation in the SFI-ASM framework is the risk premium that agents collect. As long as stock carries a positive risk premium at all, selling stock inevitably reduces the wealth that an agent could otherwise acquire. This seems counterintuitive since myopic CARA-agents maximize their expected utility, which is a monotone function of expected wealth in the next period. Even if they expect a sharp drop in stock value, they should never sell. The explanation for desired stock sales as a result of utility maximization is that agents maximize their pre-tax wealth, neglecting that shifting wealth from stock to cash will hurt them through subsequent taxation. The remaining five percent of unexplained wealth variation thus, most likely reflect differences in taxation over the lifetime of different agents. Even if the average stock holdings of two agents happen to be identical in a given period, they probably were not identical in all the periods before. This led to different tax payments which, when added up, may explain the difference in final wealth.

8.4.3 Activated Rules and Rule Selection

A first step in identifying possible reasons why distinct trader groups consistently hold more or less risky stock is to look at how they possibly differ from each other in their properties. A plausible explanation is the number of activated trading rules that systematically varies for different trader types. For instance, technical agents who check technical conditions in addition to fundamental ones end up with less activated rules than fundamental agents. Fast learning SFI agents invoke their biased mutation operator more often per period than slow learning SFI agents and thus have more specific and illogical rules. The size of the activated rule set for SFI agents is generally smaller than for non-classifier agents since for the latter, it is largely independent of the GA-invocation interval and usually coincides with the total number of trading rules they possess.[12]

It is thus, my hypothesis that it is the number of activated rules that is mainly responsible for wealth differences between different trader groups. Since the size of the set of activated trading rules depends on the number of trading rules that agents possess, it seems reasonable to compare agents with varying numbers of trading rules. In addition, one also notices an important difference between the SFI-ASM version used by LeBaron et al. and those used by Joshi et al. [206, 207]. In LeBaron et al., always the best, i.e., the rule with the lowest variance estimate from equation 6.7, is selected. Joshi et al., however, use roulette wheel selection in which a rule is randomly chosen from the active rule set with a probability proportional to its fitness value.

To test whether the number of activated trading rules may have an influence on wealth accumulation of different trader types, wealth differences were recorded for different rule set sizes under both selection procedures. Figures 8.5 for select best and 8.6 for the roulette wheel

[12] There is one caveat at this point. For both agent types, trading rules that had been activated less than “mincount” times since their creation are not included in the active rule set either. The mincount-parameter was set at a value of five and thus, tends to decrease the number of activated rules only for very fast learning speeds. Even though non-classifier agents have all their trading rules activated by definition, a rule that is younger than mincount-periods will not be considered for rule selection. For SFI agents, on the other hand, the requirement of minimum activation may take longer to fulfill. The higher the learning rate of SFI agents, the longer it takes, on average, for rules to become eligible for use.

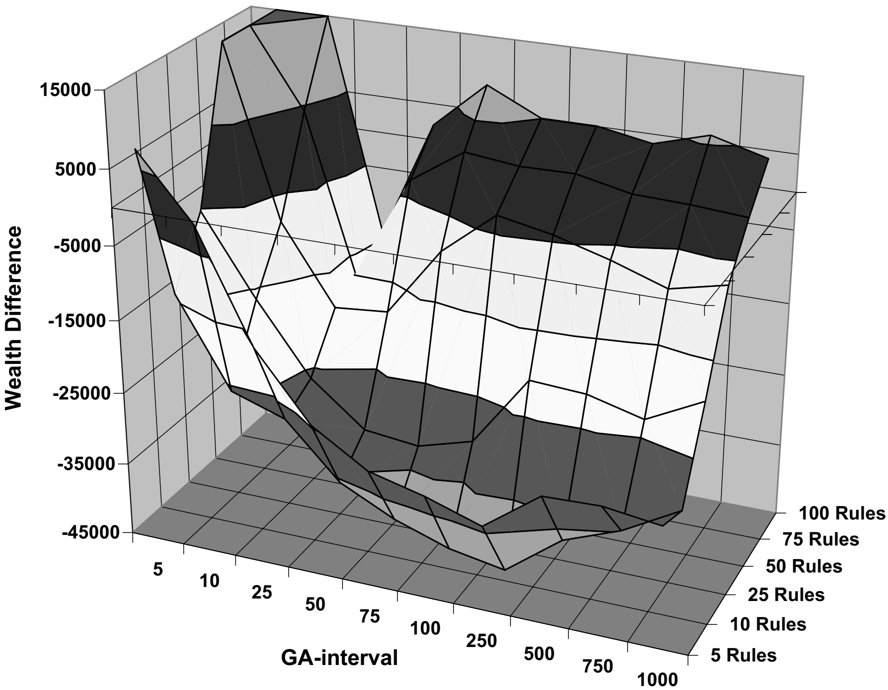


Fig. 8.5. Wealth differences between SFI and non-classifier agents for select best as a function of learning speed and number of trading rules they possess. Data were averaged over 25 simulation runs. Note that non-classifier agents outperform the SFI agents for most parameter combinations (negative wealth differences) and that a peak in the upper left corner reaching as high as 300,000 was truncated for better visibility.

mechanism thus give a more complete impression of wealth behavior than figure 8.3 does. First, one notices that wealth differences between SFI- and non-classifier agents do not obey any simple relationship when changing either the learning speed or the rule set size. Both directions have pronounced "wealth valleys", thus indicating at least two or more overlapping factors that influence an agent's wealth accumulation. Secondly, while the general shape in both wealth graphs seems similar, there are some noticeable differences such as the pronounced peak in the upper left corner for select best.

Instead of focusing on the wealth differences between competing SFI- and non-classifier agents, it might be more insightful to investigate

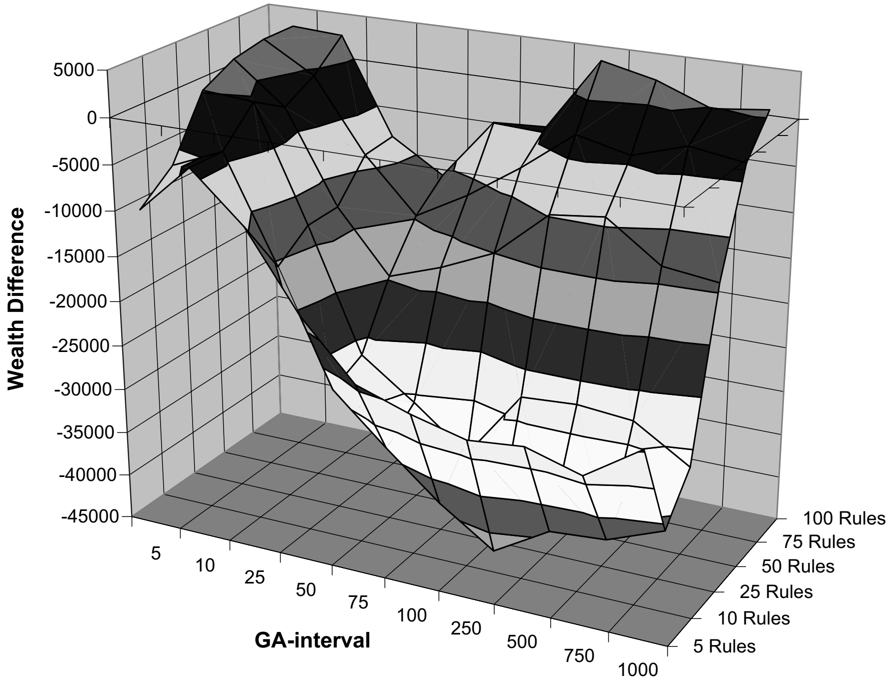


Fig. 8.6. Wealth differences between SFI and non-classifier agents for select roulette as a function of learning speed and number of trading rules they possess. Data were averaged over 25 simulation runs. Note that non-classifier agents outperform the SFI agents for most parameter combinations (negative wealth differences).

the absolute wealth levels of one agent type alone. They are shown in the top sections of figure 8.7 for SFI agents and 8.8 for non-classifier agents. It is interesting to note that subtracting the absolute wealth levels of non-classifier agents from those of SFI agents does not yield the same wealth differences as depicted in figures 8.5 and 8.6. Obviously, the co-evolution of two competing agent types creates a different economic environment in terms of aggregate price level and risk premium than just having one agent type in a simulation.

The middle and bottom sections in figures 8.7 and 8.8 point to another reason why wealth levels may vary for different trader types and learning speeds. Even when told to use select best or roulette wheel selection for forecast production, SFI agents may not be able to do so

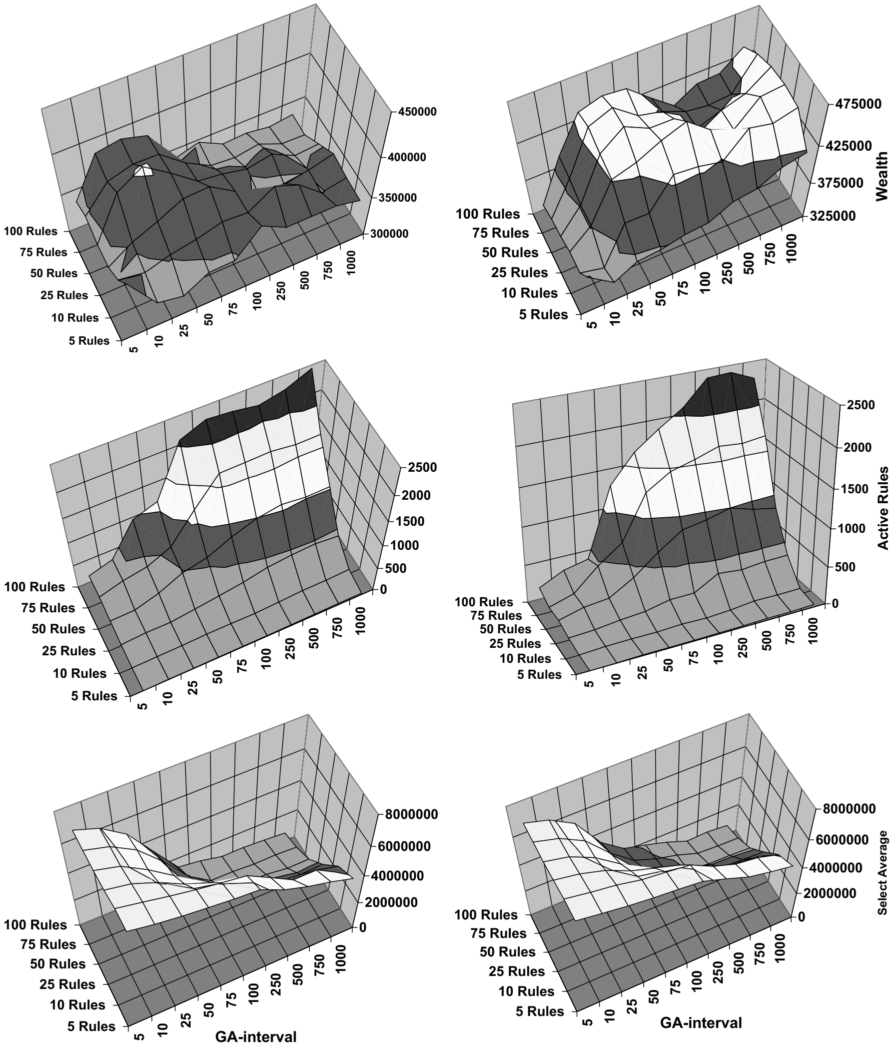


Fig. 8.7. Top: Final wealth levels for SFI agents with select best (left) and roulette wheel selection (right). Middle: Number of activated rules. Bottom: Number of "select averages" during simulation. Data were averaged over 10 simulation runs.

and employ certain fall-back methods. By increasing the learning speed for SFI agents or by reducing the number of rules they possess, the size of their activated rule set shrinks as shown in the middle section in figure 8.7. When there are no active rules at all, SFI agents derive their forecast parameters as a fitness-weighted average of all the rules

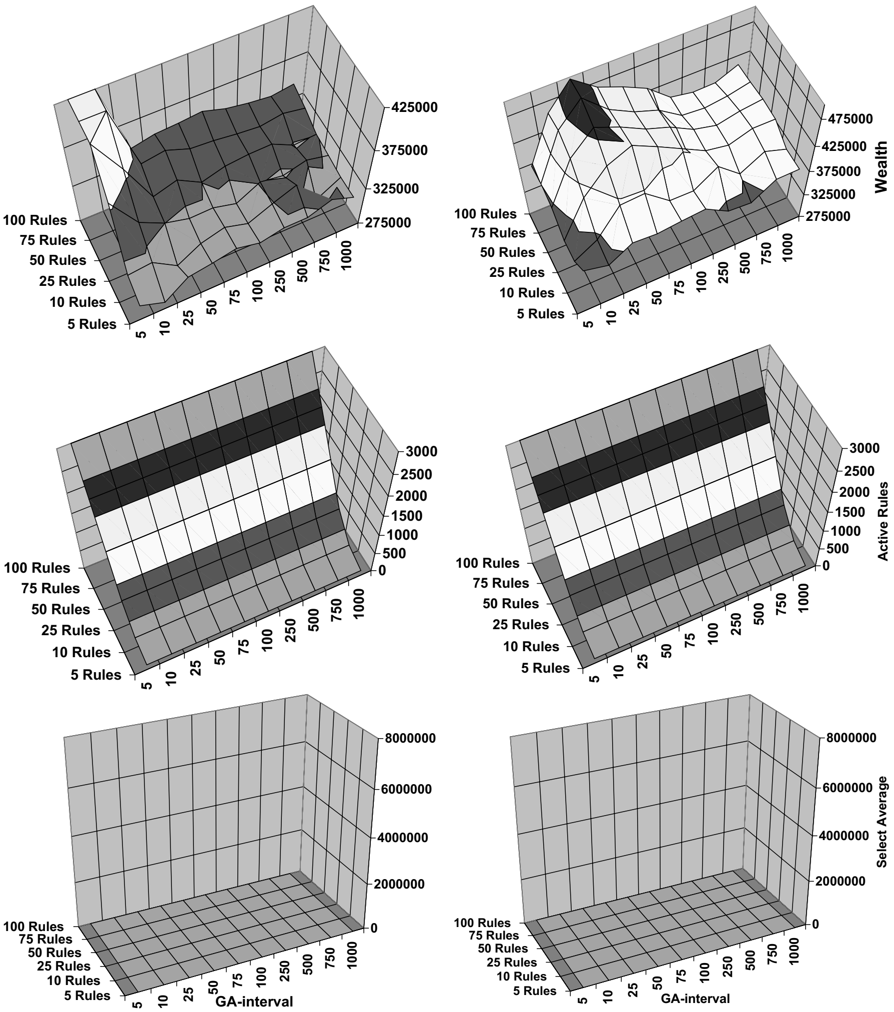


Fig. 8.8. Top: Final wealth levels for non-classifier agents with select best (left) and roulette wheel selection (right). Middle: Number of activated rules. Bottom: Number of "select averages" during simulation. Data were averaged over 10 simulation runs. Note that the wealth peak in the upper left corner for select best, reaching as high as 1.7 million, was truncated for better visibility.

that have been activated at least mincount-times in their lifetime. If none of the rules have been activated at least mincount times, SFI agents resort to using the global average of past prices and dividends as their forecast. Each of the intended or fall-back selection methods, however, has an influence on the quality of the forecast production and thus influences the model behavior. The relative frequencies with which those methods are used depend on the the parameterization, especially on learning speed, and will thus lead to different model behaviors. I subsume both fall-back mechanism in the SFI-ASM as *select average*. The bottom section in figure 8.7 shows how SFI agents increasingly resort to select average when the number of activated rules becomes smaller.

Non-classifier agents, on the other hand, are not plagued by the problem of a diminishing active rule set. Their rules are always activated (middle section of figure 8.8) and hence, there is no need for them to resort to any of the fall-back methods (bottom section). Even though non-classifier agents indeed stick to the intended selection mechanism, we still see a wealth peak for GA-intervals around 50 and a valley at about 500 for roulette wheel selection. These remaining wealth differences most likely reflect the working of the GA which changes the forecast parameters, yet other still unidentified influences are possible, too. Separating and attributing price and wealth effects to each individual factor seems impossible, though. Changing one factor affects the price series, which in turn will affect the GA and the forecast parameters, which in turn will affect the triggering change. Because of these interdependencies, an exhaustive explanation of price and wealth behavior cannot be given.

Identifying all causes and interdependencies of wealth dynamics, however, is not necessary at this point and would not justify the computational expenses. All that is necessary is to show that *some* reasons other than pattern recognition and exploitation influence the wealth distribution to a much greater extent than technical pattern exploitation ever could. Even though the game-theoretic analyses by Joshi et al. themselves are correct, they are based on wealth levels that should not be given any economic meaning.

Up to this point, we have not addressed whether one of the two selection mechanisms, select best or select roulette, is better suited. However, section 4.4.2 has already discussed the problem of scaling invariance for simple roulette wheel selection. Remember that scaling invariance means that adding an offset to all fitness values tends to equalize the selection probabilities in roulette wheel selection. In order to avoid negative fitness for trading rules, a constant C with a value of 100 is added to each raw fitness.[13] This effectively leads to almost uniform selection probabilities. The fitness information is thus widely neglected, especially in large active rule sets. Hence, we can safely as-

[13] See also equation 6.8 on page 98.

sume that unscaled roulette wheel selection as used by Joshi et al. is inferior to the select best mechanism used by LeBaron et al.

8.5 A Verdict on Wealth Analysis in the SFI-ASM

This chapter has clearly demonstrated that the only source for wealth accumulation in the SFI-ASM framework is the aggregate risk premium. Since simulated prices are usually below the hree-benchmark, wealth generated in the hree-regime can be said to constitute a lower bound for wealth in non-hree simulations. In a way, the smartest agents who are able to coordinate on the correct hree-solution make the least amount of money. It was furthermore shown that an agent's wealth is highly affected by the size of his activated rule set and the selection procedure that he uses to choose a trading rule to act upon. Last, but not least, agents in the SFI-ASM are modeled as myopic by having a constant absolute risk aversion utility function. Since the CARA-assumption implies that an agent's wealth level has no impact on his demand function, wealthier traders do not have a greater impact on market prices than poorer traders [242]. Because of all these issues, the SFI-ASM is highly unsuitable to address issues of wealth accumulation. The game-theoretic analyses by Joshi et al. thus rest on simulation results that should not be given any economic meaning.

It was also noted before that the trading rules in the SFI-ASM are not evaluated according to their wealth generating capabilities, but on their forecast accuracy. Wilpert [437] based the fitness evaluation of trading rules according to the wealth they would have generated if used by agents.[14] Equation 6.7 on page 98 as the main determinant of rule fitness (equation 6.8) would have to be adapted to

$$f_{W,t} = \left(1 - \frac{1}{\theta}\right) f_{W,t-1} + \frac{1}{\theta}\left[p_t + d_t - (1 + r_f)p_{t-1}\right] \widehat{x_{t-1}}, \qquad (8.9)$$

with $f_{W,t}$ being a raw fitness values based on wealth, and $\widehat{x_{t-1}}$ the optimal demand of last period's portfolio decision. The term in brackets multiplied by that optimal demand would denote an excess profit. It is now a legitimate question to ask whether this type of rule evaluation is still justified since we just have derived that wealth levels in the SFI-ASM are economically not meaningful. There is, however, nothing wrong with this profit based rule evaluation. Agents derive their optimal demands by neglecting any subsequent wealth taxation. Since

[14] This type of fitness evaluation was first suggested by Brock and Hommes [54].

the profitability of trading rules, too, is based on their excess profit generation before taxation, it would be completely justified to use this fitness evaluation.

9

Selection, Genetic Drift, and Technical Trading

I have yet to see any problem, however complicated, which, when you look at it in the right way, did not become more complicated.

Poul W. Anderson

9.1 Introduction

A mutation operator modification suggested in chapter 7 resulted in drastically different model behavior. Instead of discovering profitable technical trading possibilities, most agents voluntarily gave up the use of their classifier system. On the one hand, this model behavior prompted me to analyze the implications for wealth accumulation by agents. On the other hand, the model behavior was so radically different from the original results that I kept designing tests and procedures that could bolster my argument of an inherent uselessness of the classifier system. In the course of these tests, the robustness of the zero-bit solution proved to be very robust. In fact, it still emerged when I did not expect it to under the given parameter combinations. Before presenting this simulation evidence in section 9.3, I will discuss a misconception about the proper way of detecting technical trading in the SFI-ASM. The main part of this chapter will then identify and discuss the zero-bit solution as a side effect of genetic drift. In population genetics and in evolutionary programming, genetic drift is known to affect the genetic variation in finite population sizes even in the absence of any selection pressure. By transferring the longstanding selectionist-neutralist debate from evolutionary biology to the field of genetic algorithms, I will separate the effects caused by genetic drift and those caused by selective forces in the two SFI-ASM versions. As a result, I will determine that the original mutation operator is effectively combatting the influences of genetic drift while the modified operator is subject to genetic drift itself. By using different tests than a superficial look at the aggregate bit level, I will finally resolve the question whether agents in the SFI-ASM use technical trading information or not.

9.2 Technical Trading and the Aggregate Bit Level

The first signs that a superficial look at the aggregate bit level is inappropriate in judging the emergence of technical trading appeared in figure 7.3 on page 122. It was obvious that certain bits within the economy were often set in multiples of 100, which is the number of trading rules agents possess. This implied that some bit positions in the rule sets of bit-neutral agents were completely filled with 0- or 1-bits while other bit positions seemed to be completely neglected. While figure 7.3 was derived for a noiseless periodic dividend process, the bit behavior should have been similar for stochastic dividend processes if technical trading had emerged. Under the standard AR(1) dividend process, the number of instances of bit fixations at the 100% level almost completely disappeared for bit-neutral agents. The few exceptions were seen as temporary lock-ins at a suboptimal solution since in the long run, they also ended up with the zero-bit solution.

The simulation evidence of certain bit positions in an agent's rule becoming completely filled with non-# bits is consistent with the expected long run asymptotic behavior of genetic algorithms. It was shown by Holland that once an allele or a schema has been found to be useful, it will spread through the whole population [182]. In the long run, all chromosomes in a population will become identical. Once an agent has found a good trading rule, one would expect the complete rule set to be characterized by bit fixations at either the 0 or 100% level.[1]

Therefore, the remark by LeBaron et al. that bit costs were intended to ensure that *each bit* is actually useful in improving the forecast ability of a trading rule takes on a different meaning [244, p. 1497]. For a long time I, had interpreted that remark literally, i.e., I simply assumed that it meant one single bit in one particular trading rule. It now becomes clear that "each bit" only makes sense if it refers to

[1] GAs that employ *niching* methods [155] such as *fitness sharing* and *crowding* are capable of locating and maintaining several solutions within a single population. Mahfoud [270] derives lower bounds for the population size to maintain a desired number of niches. An overview of niching methods can be found in [272].

a particular bit position in the entire rule set of an agent.[2] In order to infer whether technical trading has emerged or not, one should not look at the general bit level. The proper way of assessing the existence of technical trading in the SFI-ASM is to check whether there are bit positions in agents' rule sets that are completely filled with non-# bits.

Since John Holland, the creator of genetic algorithms, was one of the founding fathers of the SFI-ASM, it is hard to believe that the SFI-ASM architects were not aware of this. It is more likely that they did not take the implications of bit convergence fully into account. Nowhere in their series of papers that describe the SFI-ASM [331, 11, 330, 244] do they refer to bit positions and bit fixations at the 100% level. Instead, they always focused on the aggregate bit level, averaged over all trading rules of all agents. Only vaguely can one sometimes infer an interest in the use of individual bits. In Arthur et al., for instance, one finds the statement that "*it can be said that technical trading has emerged if bits 7–10 become set significantly more often, statistically, than the control bits*" [11, p. 27]. First, I suppose that this statement was still intended to apply to the aggregate bit level. To detect a supposedly useful trading condition on the level of an individual trader, one does not need elaborate statistical tests since it is easy to spot a bit fixation at the 100% level, especially when all other trading conditions should have very low bit frequencies. More importantly, though, is that the significance test as suggested above cannot be applied in the original model. Remember that the control bits were, by definition, always set to zero or one. It is, therefore, impossible for any trading condition to be set more often. Another indication of an interest in disaggregated bit information can be found on page 1507 of the article by LeBaron et al. [244]. There, they narrow down their focus to the use of one particular bit position, yet it is still averaged over all agents in the economy, and the question of possible bit fixations was not addressed.

Since the aggregate bit level may get artificially inflated at higher learning speeds through the mutation operator, it is unsuitable to infer the degree of technical trading in the model as long as we do not know the exact impact that mutation has on it. There may be a high correlation between the aggregate bit usage and bit fixations, but we cannot simply assume this correlation, especially when knowing about the upward bias of the mutation operator. In order to measure the de-

[2] Depending on the interpretation, the "each bit" formulation in their papers may not be wrong, but I would call it at least unfortunate. It may have caused considerable confusion and misinterpretation in the area of genetic algorithms among novices, such as I was when I started working with the SFI-ASM.

gree of technical trading in the model, one has to investigate whether there are bit fixations. To assess whether there is emergent technical trading at higher learning speeds, one has to test whether the number of bit fixations increases for faster learning speeds.

9.3 The Zero-Bit Solution: Some Disturbing Evidence

When looking at figure 7.1 on page 116, one notices that there are some latecomer agents who experience difficulties in finding the correct hree-solution. They appear to be temporarily locked into a suboptimal solution, i.e., one or two specific trading conditions are set to either zero or one in all of their trading rules. Since these bit fixations cannot be changed by the the crossover operator, they are effectively fixed. Only mutations and generalizations are able to change several of these bits such that the #-bits can finally take over. It is obvious that this may take a long time to happen. Some long run testing with bit costs greater than or equal to 0.01, however, yielded that the latecomers are eventually able to arrive at the zero-bit solution, too. The question arises, though, as to why some of those bit fixations arose to begin with. Was it a random mistake on the agents' part, or did they temporarily discover some useful trading information?

Another peculiarity that figure 7.1 reveals is the apparent speed with which the majority of agents is able to find the zero-bit solution. No matter what GA-invocation interval is used, the first agents to discard their classifier system need only between 10 to 20 GA-invocations. Sixty to seventy percent of the other agents follow shortly after. This, however, seems quite fast for a GA-learning algorithm to discover the optimum solution. It is only later on that the discovery process slows down.[3]

The simulation results with the modified SFI-ASM have so far unequivocally shown that the zero-bit solution is robust. My growing conviction of its validity, however, was shaken when I ran the model under no bit cost. If small efficiency gains from using condition bits had been overcompensated by the associated bit costs, agents would have been forced into the zero-bit solution. Thus, the model behavior was investigated under zero bit cost. To exclude that the bit decreasing effect

[3] I should add at this point that figure 7.1 was derived with the consistency check switched on. Since the rule set then contains only logical trading rules, the number of rules were reduced to 50. The reason why this might have an impact on the speed with which agents find the zero-bit solution will become clear in the following sections.

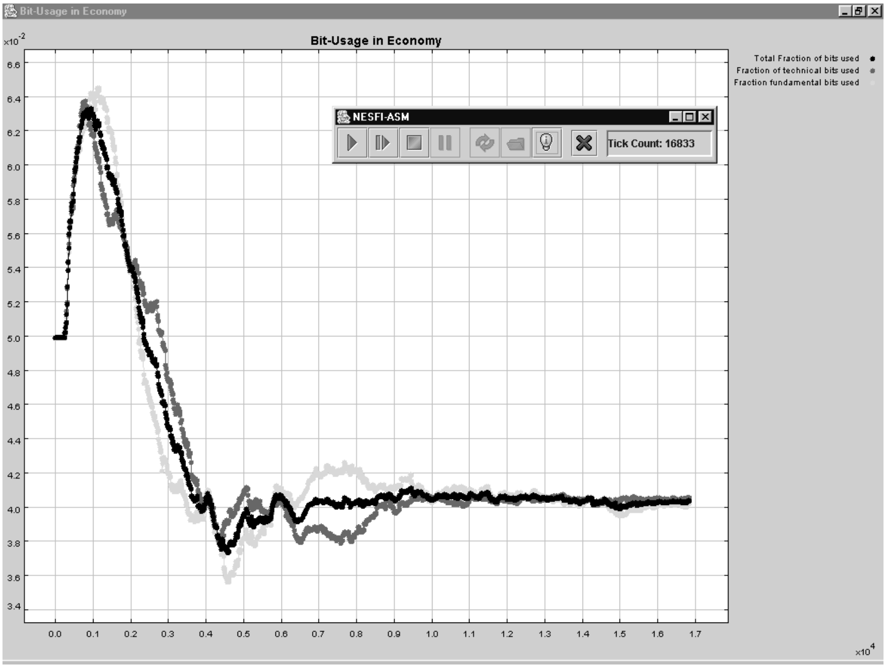


Fig. 9.1. Aggregate bit level in the modified SFI-ASM for zero-bit costs, no generalization, 25 bit-neutral agents with 100 trading rules, an initial bit probability of 0.05, a crossover probability of 0.1, and a GA-invocation interval of 100. In spite of zero-bit costs, 4 agents arrived at the zero bit level for fundamental bits and 5 agents for technical bits within less than 17,000 periods.

of the generalization procedure was responsible for the appearance of zero-bit agents, generalization was deactivated.

For no bit cost at all I intuitively expected a global bit fraction fluctuating around its initial level and no zero-bit agents. It was surprising when many, albeit less agents, still ended up in the zero-bit state. In figure 9.1, a somewhat stable aggregate bit level materialized only after four agents had given up all their fundamental trading bits and five other agents had given up their technical trading bits. The continuing existence of the zero-bit solution under no bit cost may point towards a systematic influence that exerts a downward pressure on bit usage.

The hypothesis of a stationary bit level and no zero-bit agents for zero-bit costs, on the other hand, was supported for an initial bit probability greater than a certain threshold level and for an inactive generalization procedure. In the simulation run shown in figure 9.2, no zero-bit

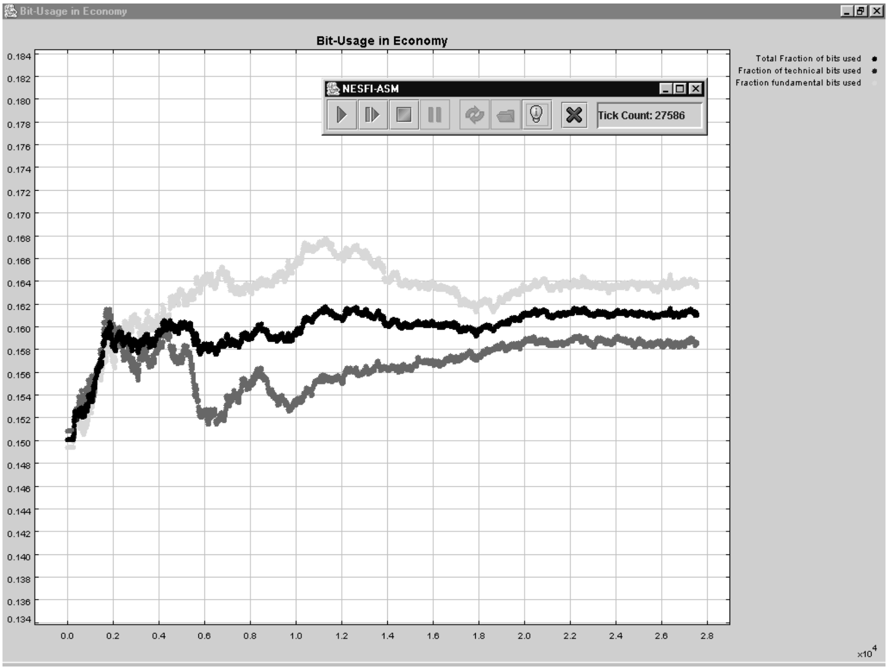


Fig. 9.2. Aggregate bit level in the modified SFI-ASM for zero-bit costs, no generalization, 25 bit-neutral agents with 100 trading rules, an initial bit probability of 0.15, a crossover probability of 0.1, and a GA-invocation interval of 100. Under this parameterization, no zero-bit agents appeared within the first 30,000 periods.

agents appeared within the first 30,000 periods and the aggregate bit level seemed to have reached its long run equilibrium. Note that the only difference in the parameterization between figures 9.1 and 9.2 is the initial bit level with which trading rules were initialized.

9.4 Random Genetic Drift in Genetic Algorithms

Because of the emergence of zero-bit agents under no bit costs, the zero-bit solution turns out to be more robust than one would normally expect. It became increasingly clear to me that zero-bit agents most likely were a result of a phenomenon known as *genetic drift*. In population genetics and in evolutionary programming, this concept refers to the loss of genetic diversity due to accumulated stochastic selection errors in finite populations, even in the absence of any selection pressure.[4]

After I had realized that useful trading bits should spread through the whole population of trading rules, I had assumed, in line with Holland's theoretical results on genetic algorithms, that this can only be the result of natural selection. Holland's results, however, were derived under some simplifying assumptions, such as an infinite population size, an accurate mapping of the utility values of a solution into a fitness function, and the absence of gene interaction in a chromosome (epistasis) [182].

In their short review of the history of evolutionary computation, DeJong et al. mentioned that the initial lack of computational resources led GA-researchers to do very few simulation runs with populations consisting of generally less than 20 individuals [93]. Those first experiments often produced results that deviated from the theoretically expected behavior. It soon became clear that those deviations were the consequence of genetic drift within finite populations.

Goldberg and Segrest [156] distinguished two types of stochastic errors in GAs with a finite population size which may prevent a GA from converging to the optimal solution. *Sampling errors* may occur because the population is not representative in that it does not contain a highly fit schema that can propagate. The stochastic errors in selection, however, are more important and were given a special name—genetic drift.

Because a new population is generated by selecting offspring from a finite sized parent generation, allele frequencies in this population are subject to random selection errors. With infinite population sizes, these errors will cause only negligible changes, yet in small populations, they will be more than noticeable. Even under no selection pressure, i.e., with a constant fitness function, the accumulation of these stochastic errors will cause certain alleles to become more predominant, and the members of a genetic population will eventually converge to a single instance in the solution space. Beasley et al. [27] describe the mechanism of genetic drift as a ratchet effect. That is, once an allele has spread to all members of a finite population, it is effectively fixed since crossover cannot change it anymore. This ratchet effect will eventually cause each individual gene in the population to become fixed. The effect of genetic

[4] Genetic drift is also sometimes characterized as random fluctuations in finite populations, "evolutionary forgetting", or allele loss. It is occasionally referred to as the *Sewall Wright effect*, named after one of the founders of population genetics. Genetic drift in the context of population genetics is discussed, for instance, in [414, 170, 174].

drift thus, can lead to a premature convergence of the GA before the optimal solution has been found.

Goldberg and Segrest analyzed genetic drift in a GA acting with the operators of reproduction and mutation on a one-locus binary structure. By using a Markov chain analysis, they calculated the expected time of first passage to various convergence levels. They found that increasing the mutation rates increased the expected time until a particular convergence level was reached for the first time. This implies that an increase in mutation rates reduces the impact of genetic drift. For mutation rates that are too high, on the other hand, the search becomes effectively random since the gradient information in the fitness function is not exploited, as Beasley et al. point out.

Unlike the calculations in terms of convergence time, approaches to quantify the effect of genetic drift based on population variances allow exact analytical expressions. Beyer [31] calculates genetic drift in unrestricted and binary search spaces. Let us assume that there is a parent population of μ single-locus individuals whose gene values are $b_i \in 0, 1$ for $i = 1, \ldots, \mu$ for the binary case. Beyer defines the population variance $\sigma_P^2[b]$ as

$$\sigma_P^2[b] := \frac{1}{\mu}\sum_{i=1}^{\mu}(b_i - \bar{b})^2, \qquad \bar{b} = \frac{1}{\mu}\sum_{i=1}^{\mu} b_i, \tag{9.1}$$

with $\bar{b}$ being the average gene value. The offspring generation is created through random sampling, i.e., one offspring is obtained by randomly choosing one of the μ parents, a procedure that is repeated μ times. Beyer then derives the expected population variance of the offspring generation as

$$\sigma_O^2 = \frac{\mu - 1}{\mu}\sigma_P^2. \tag{9.2}$$

Equation 9.2 demonstrates that random selection in finite populations reduces the expected population variances. A similar calculation of genetic drift based on population variance was done by Rogers and Prügel-Bennet who then applied their results to various selection schemes such as steady generational replacement, state selection, and generation gap methods [358].

9.5 The Neutralist–Selectionist Controversy

Even though genetic drift is often defined as the phenomenon of allelic convergence under no selection pressure, it is important to realize that it is still at work when there is a fitness function. This immediately leads to the question of whether one is able to separate the effects of "natural" selection from those of genetic drift, and which of the two evolutionary forces are more important in the evolutionary process.

In fact, Harrison et al. point out that the question about the relative importance of genetic drift versus natural selection in determining evolutionary change was and still remains one of the most controversial issues in population genetics [170]. The so-called neutralist-selectionist debate arose between Sewall Wright and Ronald A. Fisher in the 1930's and 1940's when they started to investigate the fixation probabilities of newly mutant alleles in a given population. It has been rekindled with the ascent of the neutral and near-neutral theories of molecular evolution by Kimura [222, 223] and Ohta [326, 327] about the neutrality or near-neutrality of genes.

The following description of the Wright-Fisher model is intended to highlight the differences between these two different schools of thought in population genetics. The Wright-Fisher model will also form the basis for our following analysis of whether selection or random genetic drift is more important in the two versions of the SFI-ASM. The Wright-Fisher model assumes a population of N haploid organisms[5] who create an offspring generation of N individuals through random mating. Let us consider the case for two alleles A and a with corresponding allele frequencies p and q. Assume that there are i copies of A at a certain gene position, i.e., $p = i/A$ and $q = 1 - i/A$. A's relative fitness w is defined as its absolute fitness, often the number of expected offspring, divided by the average fitness in the population. The selection coefficient favoring A is denoted by $s = w - 1$ and can take any value in the interval $[-1, \infty)$.[6] We are interested in the fixation probability $U(p)$ of this allele for varying degrees of selection pressure and differing population sizes. There is, however, no closed-form solution to this problem in discrete time, and enumerative algorithms work only for very small population sizes. A diffusion approximation in a continuous time formulation was developed by Kimura [220, 221],[7] and A's probability of fixation is determined as

[5] A haploid organism contains only one copy of each chromosome. Diploid organisms, i.e., most sexually reproducing species, contain two copies of each chromosome, one from each parent. The genetic algorithm in the SFI-ASM, as most GAs, is a haploid GA-representation.

[6] The selection coefficient could also be defined as being against A. In this case, $s = 1 - w$ would be in the interval $(-\infty, 1]$.

[7] A relatively easy to follow derivation of this result can be found in [224] and [126]. Since the diploid case, i.e., each individual owning two sets of chromosomes,

$$U(p) \approx \frac{1 - e^{-2Nsp}}{1 - e^{-2Ns}}. \tag{9.3}$$

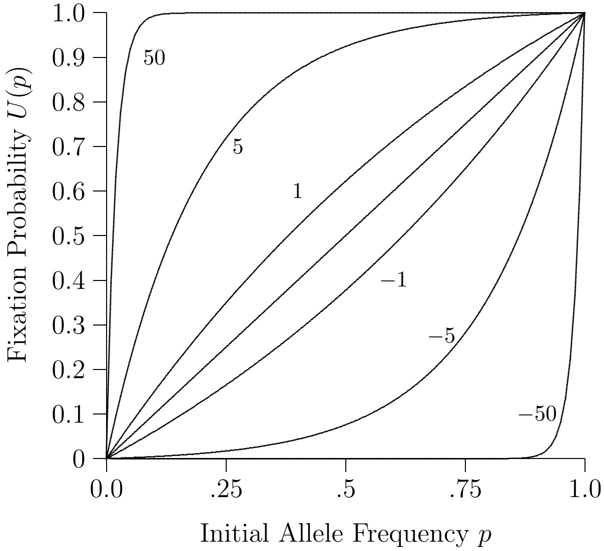


Fig. 9.3. Probabilities of fixation of an allele for various values of $2Ns$ and p. The values of $2Ns$ are shown adjacent to the curves, except for the diagonal with no selection pressure ($2Ns = 0$).

Figure 9.3 plots the fixation probability $U(p)$ as a function of p for various values of $2Ns$. Let us first consider the case under no selection pressure. For $s = 0$, both the numerator and denominator of equation 9.3 are zero, but by using L'Hôpital's theorem, we obtain that the probability of fixation for an allele equals its initial allele frequency p. For positive selection coefficients, the fixation probabilities of A exceed its initial allele frequencies. It is, however, obvious from equation 9.3 that the selection coefficient s, alone, is not sufficient to decide about the relative importance of natural selection versus genetic drift. It has become commonplace in evolutionary biology to regard

$$s > \frac{1}{2N} \tag{9.4}$$

as the threshold at which natural selection begins to have a significant impact on allele fixation. Otherwise, the alleles are either called *nearly*

is more common in nature, one often finds the diploid formula $U(p) \approx (1 - e^{-4Nsp})/(1 - e^{-4Ns})$.

neutral or *effectively neutral*, i.e., their frequencies are determined more by genetic drift [80].

While Fisher explicitly assumed that the population size (up to 10^{12} individuals) is extremely large, Wright thought more of local populations and explicitly considered effective population sizes.[8] For extremely large N as Fisher had assumed, very small selective forces are sufficient, given enough time, to cause large adaptive changes. In contrast to this Darwinian or Fisherian gradualism as it is sometimes called [432], Wright espoused a much larger role of genetic drift in evolutionary adaptation due to smaller population sizes. He rejected Fisher's idea that either selection or genetic drift had to alone be responsible for the observed fluctuations in gene frequencies. Instead, he accepted that several evolutionary forces—mutation, random genetic drift, selection, and migration—act simultaneously. As a result of this controversy, random genetic drift was dubbed the *Sewall Wright Effect* by Fisher. The modern version of this debate, the neutral and nearly neutral theories of molecular evolution, are mainly concerned with the percentage of alleles that are neutral or near-neutral, i.e., with zero or only negligible effects on an individual's fitness. Similar to Wright, the difference in the traditional selection theory lies in the magnitude of the Ns coefficient [327].

9.6 Fitness Driven Selection or Genetic Drift?

9.6.1 Selection or Genetic Drift in the Modified SFI-ASM?

The idea that the fixation of trading bits at the 100% level in the SFI-ASM necessarily reflects their advantageousness is obviously equivalent to the position held by the selectionist faction. The view that vanishing trading bits are likewise a result of selection would equally be favored by the selectionists.

The emergence of zero-bit agents under no bit costs, however, clearly points to genetic drift as a major evolutionary force in the modified SFI-ASM. Since genetic drift has not been taken into account yet, the sim-

[8] The N in equation 9.3 actually refers to effective population size. In order to determine effective population sizes, evolutionary biologists make adjustments for unequal sex ratios, strong sexual selection, or fluctuating population sizes [80, 174]. In genetic algorithms, such adjustments are usually unnecessary. If, however, the SFI-ASM founders would have opted to prevent illogical trading rules from reproducing, the effective population size would have been lower than the number of trading rules.

ulation results presented so far are in need of reinterpretation. Judgement about whether there is technical trading in the SFI-ASM or not has to be postponed.

The Zero-Bit Solution and Initial Bit Probability

In hindsight, the choice of initial parameter values made bit-neutral SFI agents particulary vulnerable to getting locked into the zero-bit solution through genetic drift. The trading rules were initialized with a bit probability of only 0.05. Since 95 % of all trading bits were already don't care signs #, random genetic drift was much more likely to fixate the #-bits than to propagate the 0 and 1-alleles through the population. When switching off mutation and setting bit costs to zero, equation 9.3 tells us that, on average, 95% of all bit positions would fixate at the zero-bit level. The remaining five percent will eventually fixate at the 100% level.[9]

This also explains why a simple increase in the initial bit level of trading rules drastically reduced the emergence of zero-bit agents in figure 9.2 on page 152. Even though no zero-bit agent appeared within the first 30,000 periods in that particular simulation, additional tests have shown that further increasing the initial bit probability cannot prevent zero-bit agents from eventually showing up in the long run. Zero-bit agents are, in a way, inevitable. In cases with high initial bit probabilities, the generalization procedure first lowers the bit level, which paves the way for genetic drift to work its way toward the zero-bit solution.

Differentiation of Trading Rule Sets

It is also known from population genetics that random genetic drift tends to create subpopulation differentiation [80]. The analogy between an agent's rule set and the subpopulation concept in population genetics is hardly contentious. All rule sets were initialized randomly and thus, had the same stochastic properties such as allelic frequencies. Yet, after they started to independently evolve, random genetic drift led to an increase in the among-population variance (differentiation), and to a decrease in the within-population variation.[10] When discussing the bit distribution under periodic dividends in figure 7.3 on page 122, it was pointed out that each agent converged to a single trading rule that he was using, yet all agents seemed to have discovered a different techni-

[9] The Wright-Fisher model in the previous section did not take random mutations into account and assumed random mating. Mutation and the fact that the two SFI-ASM versions choose parent rules through tournament selection will have an impact on the exact fixation probabilities.

[10] The decrease in the within-population variance was already captured in equation 9.2.

cal trading rule. In simulation runs with many traders, subpopulation differentiation may prevent one from detecting bit fixations when only looking at the bit usage across all agents.

LATECOMER BIT-NEUTRAL AGENTS

Let us assume for a while that the majority of bit-neutral agents who arrive at the zero-bit level rather quickly do so through genetic drift and not by fitness driven selection. The latecomer agents then suddenly become of particular interest to us.

Table 9.1. Typical rule set of a bit-neutral "latecomer" agent. The rule set contains 100 trading rules, each with 64 trading bits. The last row shows the sum of all non-# bits.

	1	2	3	4	5	6	7		58	59	60	61	62	63	64
1	0	#	0	#	#	#	#	······	#	1	#	#	#	#	#
2	#	#	0	#	#	#	#	······	#	1	#	#	#	#	#
3	#	#	0	#	#	#	#	······	#	1	#	#	#	#	#
⋮	⋮														⋮
98	#	#	0	#	#	#	#	······	#	1	#	#	#	#	#
99	#	#	0	#	#	#	#	······	#	#	#	#	1	#	#
100	#	#	1	#	#	#	#	······	#	1	#	#	#	#	#
$\sum$	1	0	100	0	0	1	0		0	99	0	0	1	0	0

The fixation of certain bit positions in the trading rules of latecomers could be either a mistake on their part, i.e., a result of genetic drift, or indeed reflect fitness driven selection. A typical rule set of a latecomer agent is shown in figure 9.1. Long run simulations have shown though that those agents will sooner or later become zero-bit agents themselves. The reversal of bit fixations, however, is not necessarily a sign of fitness-driven forces finally gaining ground. This interpretation is likely, but again, equation 9.3 tells us that there is a small, but positive probability of this happening, even under no selection pressure.

9.6.2 Selection or Genetic Drift in the Original SFI-ASM?

Having inferred from the zero-bit solution under no bit costs that random genetic drift plays a major role in the modified SFI-ASM, does its absence in the SFI-ASM indicate the dominance of selective forces? Not necessarily. Given the upward biased mutation operator, a zero-bit solution in the SFI-ASM is impossible to begin with.

Table 9.2. Typical rule set of a SFI agent. The rule set contains 100 trading rules, each with 64 trading bits. The last row shows the sum of all non-# bits.

	1	2	3	4	5	6	7		58	59	60	61	62	63	64
1	#	0	#	0	#	1	#	······	#	#	1	#	0	#	#
2	#	#	#	1	#	1	#	······	#	#	1	#	0	#	#
3	#	#	#	0	#	1	#	······	#	#	1	1	0	#	#
⋮	⋮														⋮
98	#	#	#	#	#	1	#	······	#	#	#	#	#	1	#
99	#	0	#	0	#	1	#	······	#	#	1	1	0	#	#
100	#	#	#	#	#	1	#	······	#	#	#	#	0	1	#
$\sum$	1	4	0	96	0	100	0		0	1	93	5	99	5	0

A typical rule set of SFI agents shown in table 9.2 reveals that there are indeed bit positions that have fixated. Instead of providing data on the aggregate bit level in the economy, this could have been a much stronger argument for emergent technical trading. After all, those bits close to the 100% level have emerged from low initial bit frequencies. But again, it is possible that those bit positions became fixed through random genetic drift and not through selective forces. If we compare the two rule sets shown in tables 9.1 and 9.2, one notices several things. First, compared to bit-neutral latecomers, SFI agents have more bit positions that are close to zero, but not exactly zero. Second, some bit positions in SFI-rule sets tend to be only "nearly" fixated, i.e., they are also close to, but not exactly at the 100% level. In the rule sets of bit-neutral agents, the fixation is usually complete. Both effects are probably an effect of the SFI-mutation operator which is more likely, i.e., with probabilities of two thirds, to introduce a 0- or a 1-bit when it finds a #-bit at a particular position, or to change non-# bits to the don't care symbol.

It is also possible to offer an explanation for the apparent path dependency of equilibrium bit levels. In section 6.7.2, it was pointed out that the equilibrium bit frequencies differ when the trading rules are initialized with different bit probabilities. Equation 9.3 showed that the fixation probability of trading bits increases for higher initial allele frequencies. Starting at higher initial bit probabilities, thus, will lead to more bit fixations and a higher aggregate bit level. It does not matter at this point whether bit fixation is a result of genetic drift or selection, but once it has occurred, it is hard to reverse.

9.6.3 Genetic Drift, Fitness Gradient, and Population Size

How much selective pressure is necessary to ensure that the two model versions of the SFI-ASM are working in the regime where selection dominates genetic drift? Remember that section 9.5 established that the condition $s > 1/(2N)$ has to be satisfied for selection to be primarily responsible for allele frequencies. Certainly, this is more a rule of thumb than a strict border separating the two regimes, but it will help us get an idea of how much a trading bit has to improve a rule's forecast accuracy in order to be selected.

An individual's i relative fitness w_i was defined as its absolute fitness f_i over mean population fitness $\bar{f}$, and the selection coefficient in favor of individual i as $s = w - 1$ with $s \in [-1, \infty)$. Inspection of typical trading rule sets shows the mean fitness values of trading rules to be mostly in the interval 92–94. Since we are only interested in an estimate of the magnitude of the fitness differential, the exact value does not matter that much. If we then assume that mean fitness is 93.0, by how much does a rule's fitness value have to be improved by an additional trading bit such that this bit is likely to become fixated through selection? Substituting s with $w - 1$ and rearranging equation 9.4 yields

$$\begin{aligned} f_i &> \bar{f}\left(1 + \frac{1}{2N}\right) \qquad (9.5) \\ &> 93.000\left(1 + \frac{1}{200}\right) \\ &> 93.465. \end{aligned}$$

Each trading bit in a rule should, thus, improve a rule's forecast accuracy not only by the associated bit costs of 0.005, but by an additional offset of about 0.5 which is necessary to counter the effect of genetic drift.[11]

However, having no idea whether, and if so, by how much, a single trading bit is able to improve a rule's forecast accuracy, this calculation does not help us in deciding whether the gradient information contained in fitness functions is sufficient for selection to dominate genetic drift. It does illustrate, however, that a trading bit offsetting only slightly more than its associated bit costs is not guaranteed for being selected and for spreading through the whole rule set. If, however, we would

[11] There is, however, the possibility of epistasis, that is, a certain trading bit is only useful in conjunction with another trading bit. In the derivation of 9.3, epistasis has not been accounted for.

know by how much a certain allele would contribute to the fitness of a chromosome, equation 9.5 could be solved for the minimum effective population size N for a given GA-problem.

9.7 The Effect of Mutation on Genetic Drift

Since we do not know the possible magnitude of the fitness contribution of useful trading bits, it is also possible that the fixated bit positions in both model versions reflect genetic drift and not selection. Dynamic stability or instability of fixated bit position can also be both a sign of genetic drift or selection. If a gene is found to be useful, one could argue, its alleles should remain fixed over time. On the other hand, the SFI-ASM was developed to show that the stock market is in a constantly co-evolving dynamic environment. Fixation reversal thus could indicate changes in the economic environment.

We can do nothing other than concede at this moment that we are in the midst of the neutralist-selectionist debate. How can we judge whether the fixation of some trading bits reflects technical trading or not? How can we separate whether the effects on the bit level have mainly been caused by genetic drift or by fitness based selection?

In evolutionary biology, this issue is often resolved by taking the neutral theory as a "null model" and testing whether observed differences in DNA sequences can be explained by genetic drift or not.[12] In computer simulations with genetic algorithms, we do not need to resort to the theoretical predictions by the neutral model since we are able to simply test our genetic algorithm under no selection pressure. The observed behavior can then serve as the null model with which the full-blown model behavior is compared.

9.7.1 Genetic Drift, Mutation, and Crossover Only in SFI Agents

To obtain the "null behavior" of the classifier system under no selection pressure, the model was run in GA-only mode. In this mode, the only activity of agents was to alter their trading rules by invoking the GA. In addition to disabling all other model routines, certain adjustments

[12] Evolutionary biologists resort to this type of test when they already have an estimate of how much time has passed since two lineages have diverged. They then compare whether the observed genetic differences may be explained by the neutral model of genetic drift or not, i.e., whether selection has also had an impact on the genetic sequences.

to the GA had to be made. Since there are no "bad" trading rules under a constant fitness function, newly created rules had to randomly replace older trading rules in the rule set. The GA-parameters were set at the same values that were used in realistic simulations runs.

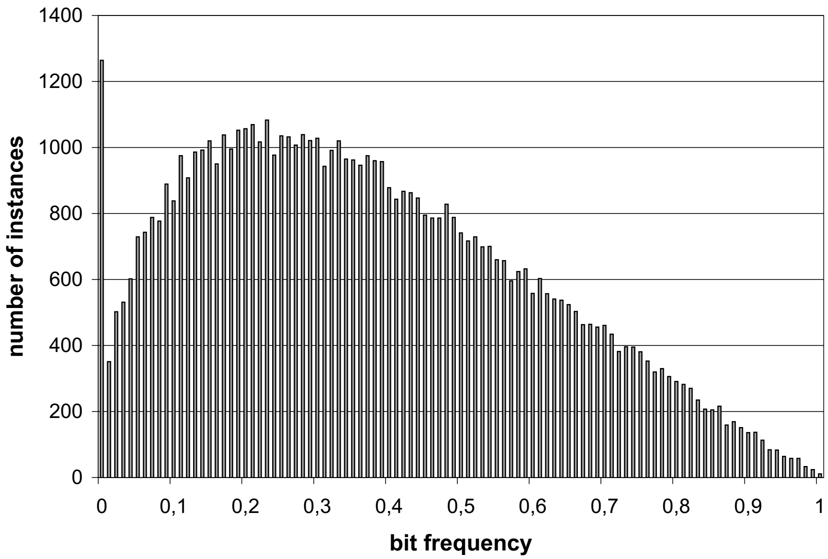


Fig. 9.4. Number of bit positions that have a certain bit frequency within an agent's rule set. Data were obtained at period 250,000 from a single GA-only simulation run with 1,000 SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs = 0, GA-interval = 1,000).

Figure 9.4 shows the bit frequency of bit positions in the rule sets of SFI agents. A bit frequency of zero means that a certain bit position in an agent's rule set has no set trading bits at all. A bit frequency of one refers to the situation when all 100 trading bits are set to either zero or one. In a simulation run with 1,000 SFI agents who all possessed 100 trading rules consisting of 64 trading bits, a total of 64,000 bit positions could be investigated at once. Figure 9.4 shows that a bit frequency of zero is the most common one, even though it does not fit very well into the overall shape of the remaining histogram. Apart from this outlier, the bit frequencies between 0.15 to 0.35 are the most common ones. It is also worth noticing that a few bit positions with a bit frequency of one have emerged even though the initial bit probability was set at 0.05.

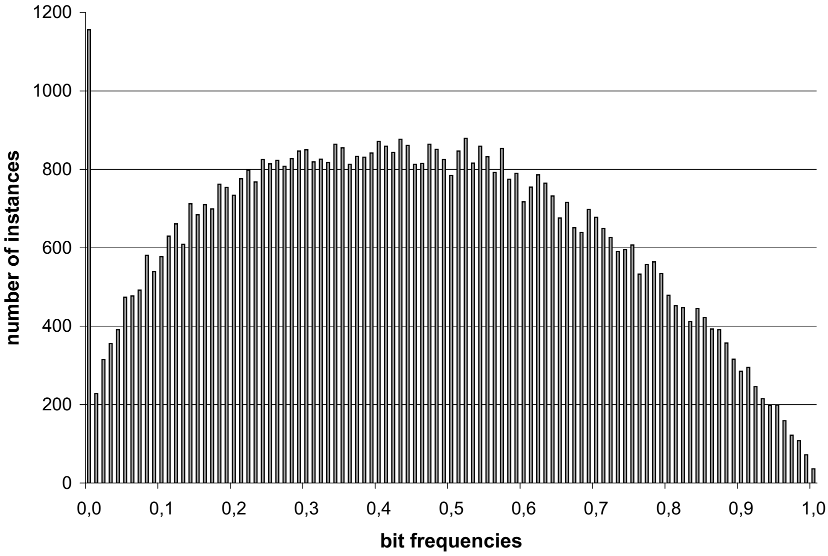


Fig. 9.5. Number of bit positions that have a certain bit frequency within an agent's rule set. Data were obtained at period 250,000 from a single GA-only simulation run with 1,000 SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs = 0, GA-interval = 50).

Figure 9.5 was obtained by increasing the GA-invocation interval to 50. One notices that the mode of the distribution has shifted to the right, that the bit frequency of zero has become less common, and that the number of bit fixations at the 100 % level has increased. Since SFI agents do nothing other than randomly altering the condition bits of their trading rules, increasing the "learning speeds" simply speeds up convergence to the equilibrium bit frequency distribution. Convergence would be fastest if agents would call the GA in every period. The shifting shape of the bit frequency distribution towards the center suggests that the equilibrium bit distribution is centered around one half. The remaining peak at a bit frequency of zero could be caused through the crossover operator. Its decrease for faster GA-rates makes it more likely though that the bit distribution has not yet reached its equilibrium. I suppose that in the limit, the number of instances with a bit frequency of zero equals the number of bit fixations at the 100% level.

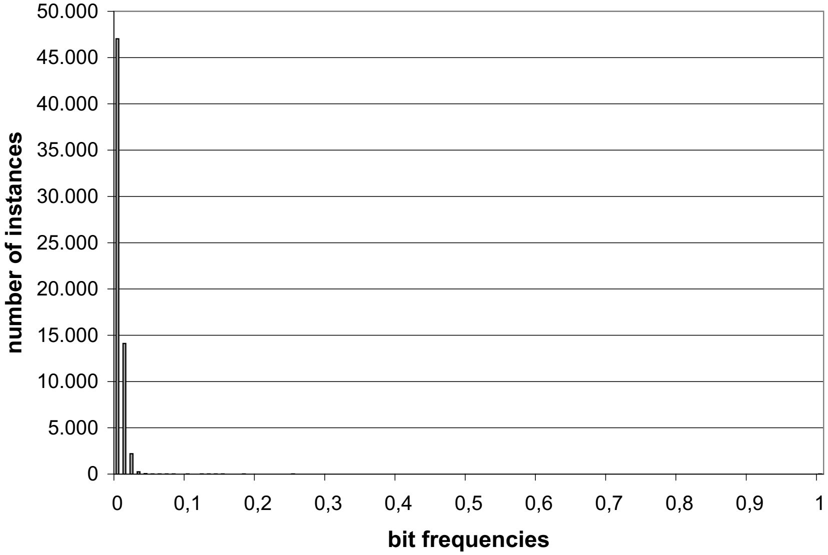


Fig. 9.6. Number of bit positions that have a certain bit frequency within an agent's rule set when bit costs are positive. Data were obtained at period 250,000 from a single GA-only simulation run with 1,000 SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs = 0.005, GA-interval = 5).

Figure 9.6 shows a dramatic change in equilibrium bit frequencies when bit costs are included.[13] Most of the bit positions are completely filled with #-signs. It is remarkable though that genetic drift has resulted in 11 bit fixations at the 100% level and 3 instances at the 99%-level.

In the presence of bit costs, the proper null-model with which the bit distribution of real SFI agents should be compared is figure 9.6. The significance of the equilibrium bit distribution under no selection pressure at all will become clear in section 9.7.3.

[13] Under positive bit costs, the 20 worst trading rules were replaced. With no other fitness influence than bit costs, the exact value does not matter, i.e., the bit distributions for different bit costs would look identical.

9.7.2 Genetic Drift, Mutation, and Crossover Only in Bit-Neutral Agents

For bit-neutral agents, the bit distribution under no selection pressure is shown in figure 9.7. Qualitatively, it resembles the bimodal bit distribution of ordinary SFI agents in the presence of bit costs.

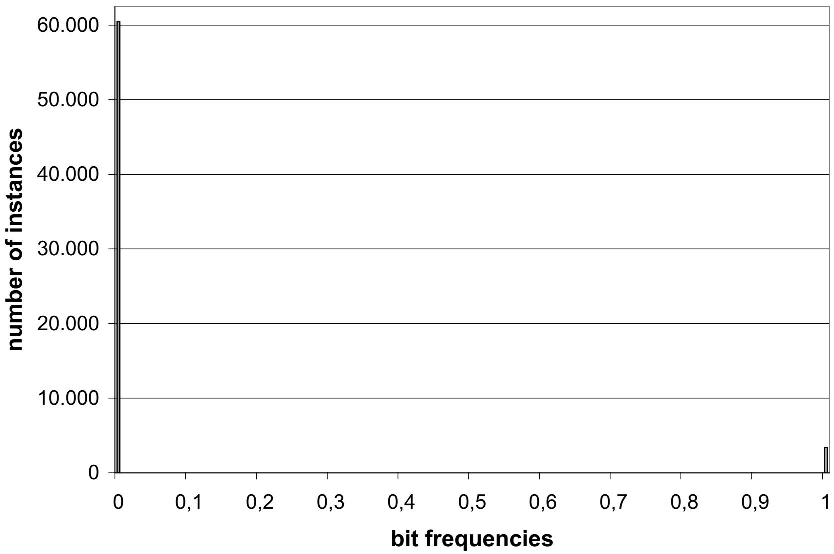


Fig. 9.7. Number of bit positions that have a certain bit frequency in the rule sets of all bit-neutral agents when bit costs are zero. Data were obtained at period 250,000 from a single GA-only simulation run with 1,000 bit-neutral SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs $= 0$, GA-interval $= 5$).

There are 3,403 out of 64,000 possible bit fixations at the 100% level, that is, 5.3% of trading bits became fixated. Given an initial bit probability of 0.05, this is remarkably close to the theoretical prediction of 5% by equation 9.3. The bit equilibrium distribution in the presence of bit costs is trivial—there are no remaining bits at all.

9.7.3 An Equilibrium Analysis of Genetic Drift and Mutation

The previous sections have shown that replacing the original mutation operator with an unbiased operator will result in drastically different bit equilibrium distributions under no selection pressure. These equilibrium distributions are the result of the interaction of genetic drift, mutation, and crossover. Since the equilibrium properties of mutation and genetic drift have already been studied in population genetics , we are able to investigate the two mutation operators in more detail. Consider a neutral model from population genetics with reversible mutation, i.e., an allele A mutates to a at rate μ and a alleles mutate to A with probability u. If we denote the allele frequency of A at time t with p_t, A's frequency in $t+1$ is then given as

$$p_{t+1} = (1 - p_t)u - p_t\mu. \tag{9.6}$$

By using a diffusion approximation approach, the final equilibrium density for the haploid case can be shown to be

$$\phi(p) = Cp^{2Nu-1}(1-p)^{2N\mu-1}. \tag{9.7}$$

A derivation for the diploid model can be found in Felsenstein [126, chapter VII.9] and Bustamante [60]. Because equation 9.7 is the density of a Beta distribution with parameters $2Nu$ and $2N\mu$, the constant C is

$$C = \frac{\Gamma(2Nu + 2N\mu)}{\Gamma(2Nu)\Gamma(2N\mu)}. \tag{9.8}$$

Figure 9.8 shows some probability density plots of allele frequencies when mutation rates from A to a and vice versa are symmetrical. For values of $2Nu = 2N\mu > 1$, the Beta distribution is unimodal and the probability mass of allele frequencies is centered around its theoretical mean of half.[14] For large values of $2Nu = 2N\mu \gg 1$, allele frequencies are tightly clustered around the deterministic equilibrium, because once genetic drift moves the allele frequencies away from this equilibrium, mutation pushes it right back. Mutation thus dominates genetic drift.

For small values of $2Nu = 2N\mu < 1$, the Beta distribution is U-shaped with most of its probability mass concentrated near the absorbing states at $p = 0$ and $p = 1$. In this case, genetic drift dominates mutation and has pushed the equilibrium allele frequencies to the absorbing states with little resistance from mutation. There are some occasional bit reversals from one tail of the curve to the other when

[14] The mean of a Beta distribution with parameters α and β is $\alpha/(\alpha+\beta)$. The mode of a Beta distribution $(\alpha - 1)/(\alpha + \beta - 2)$ is only defined for $\alpha > 1$ and $\beta > 1$.

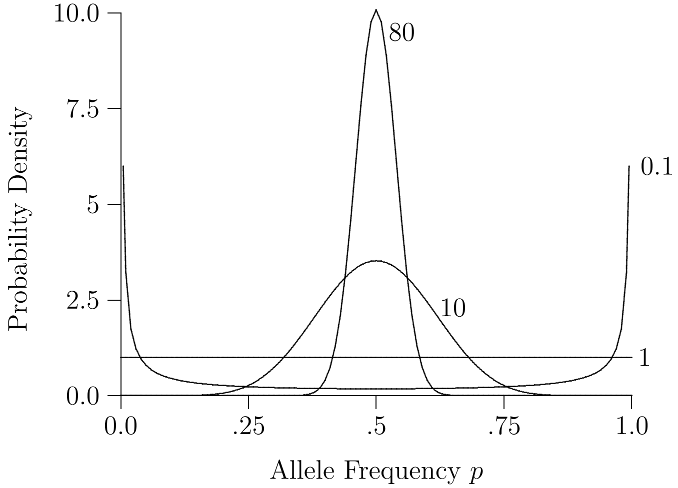


Fig. 9.8. Equilibrium distribution of allele frequencies under genetic drift and mutation for symmetric mutation ratios $2Nu = 2N\mu$. The values for $2Nu$ are shown next to the curves.

a new mutation succeeds in spreading through the population.[15] The uniform distribution at $2Nu = 2N\mu = 1$ is the dividing line where mutation and genetic drift balance each other out exactly.

It is now straightforward to apply equation 9.7 to the SFI-mutation operator and to interpret the resulting equilibrium density function of allele frequencies. Technically, our "genes" in the SFI-ASM come in three alleles, 0-, 1-, or #-bits. We can, however, subsume the 0- and 1-bits into one "super allele" A with the meaning "bit is set". From the bit transition matrix 6.9 on page 101, we see that the probability of A being mutated to # is two thirds. Similarly, the mutation rate of #-bits to zero or one is also two thirds. Since the mean of a Beta distribution with parameters α and β is $\alpha/(\alpha + \beta)$, mutation presses the gene frequencies towards an equilibrium of one half. Note that this is the same fixed point that we have determined with our Markov chain analysis of the SFI-mutation operator. In order to determine the effective mutation rates of alleles in the SFI-ASM, one has to consider the probability with which individual bits are chosen for selection, i.e.,

[15] Technically, the Beta distribution for $2Nu = 2N\mu < 1$ is not defined at its tails for $p = 0$ and $p = 1$, but we are only approximating a discrete histogram with a continuous density function. The diffusion approximation never predicts an allele frequency of exactly zero or one. Practically, there will be numerous genes with either none or all of their alleles set at a common value.

we have to take predictor mutation probability Π and bit mutation probability π into account. The effective mutation rates u and μ are then determined as

$$u = \mu = \frac{2}{3}\Pi\pi = \frac{2}{3} \times 0.9 \times 0.03 = 0.018. \tag{9.9}$$

With our population of 100 haploid trading rules, the numerical values for $2Nu = 2N\mu$ equal 3.6, and the resulting probability density function of allele frequencies is shown in figure 9.9. Since $2Nu$ and $2n\mu$ are greater than one in the SFI-ASM, the equilibrium density of allele frequencies is clustered around the deterministic equilibrium of one half. The unimodal probability density of allele frequencies shows that the original SFI-mutation operator dominates genetic drift. Now it becomes clear that the two empirical bit distributions in figures 9.4 and 9.5 are on their way to their final equilibrium distribution shown in figure 9.9.

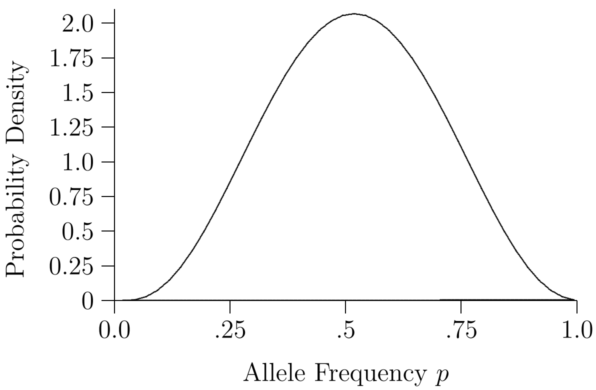


Fig. 9.9. Equilibrium distribution of allele frequencies under genetic drift and mutation with the original SFI-mutation operator.

The effective mutation rates u and μ for the bit-neutral agents are dynamically self-adjusting. We can, however, generate a probability density plot for equilibrium allele frequencies if we keep the bit transition probabilities from the bit transition matrix fixed at their values for the initial bit frequency of 0.05. The effective mutation rates from A to #-bits and vice versa are not symmetrical anymore and equal

$$u = 0.95\Pi\pi = 0.95 \times 0.9 \times 0.03 = 0.02565 \tag{9.10}$$

$$\mu = 0.05\Pi\pi = 0.05 \times 0.9 \times 0.03 = 0.00135. \tag{9.11}$$

The parameters for the Beta distribution are then calculated as

$$2Nu = 200 \times 0.02565 = 5.13 \tag{9.12}$$

$$2N\mu = 200 \times 0.00135 = 0.27. \tag{9.13}$$

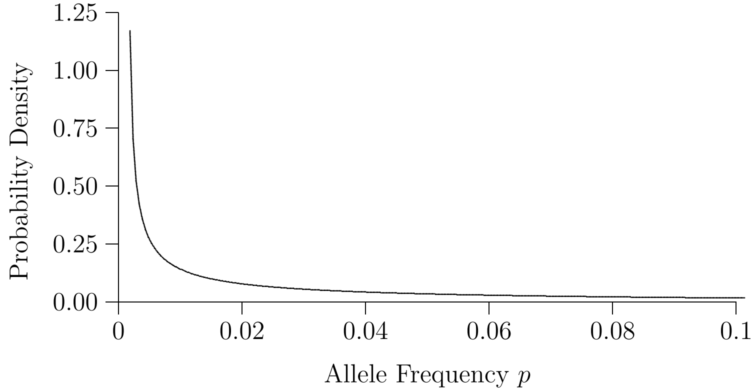


Fig. 9.10. Equilibrium distribution of allele frequencies under genetic drift and mutation with the modified mutation operator when population size is 100 trading rules and effective mutation rates are kept fixed at the initial bit probability. Note that the probability density values for allele frequencies > 0.1 have been cut off for better visibility.

The corresponding density plot of equilibrium allele frequencies is shown in figure 9.10. It clearly shows that most of the probability mass is torn towards the tail with $p = 0$. With only 100 trading rules and the newly suggested mutation operator, genetic drift dominates mutation. The simulation evidence of bit fixation at the zero-bit level is, thus, evidently a result of genetic drift and not of selective forces as previously thought.

Equation 9.7 also allows us to determine a critical effective population size for any given mutation operator. For symmetrical mutation rates, the dividing line at which mutation and genetic drift balance each other out is where $2Nu = 2N\mu = 1$. With these parameters, the Beta distribution turns into a uniform distribution. For unequal mutation rates, the dividing line between genetic drift and mutation is crossed when the smaller of the two parameters exceeds 1.0.
With the new mutation operator, the condition

$$\begin{aligned} 2N\mu &\geq 1.0 \\ N &\geq 371 \end{aligned}$$

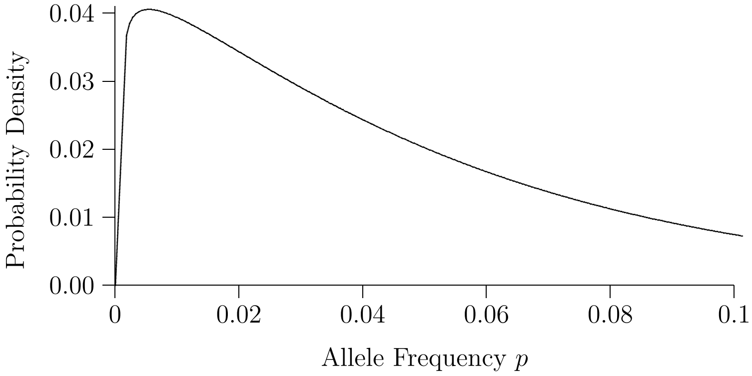


Fig. 9.11. Equilibrium distribution of allele frequencies under genetic drift and mutation with the modified mutation operator when population size is 400 trading rules and effective mutation rates are kept fixed at the initial bit probability. Note that the probability density values for allele frequencies > 0.1 have been cut off for better visibility.

has to be satisfied for mutation to dominate genetic drift.[16] Figure 9.11 shows a unimodal probability density function of equilibrium allele frequencies when agents possess 400 trading rules. Note that even though the continuous approximation of the density function starts at zero, the first class in the discrete histogram is very likely filled with many instances of allele frequencies of zero. It should be mentioned at this point that the prior intuition about reducing the rule set size for an activated rule consistency check was not justified. It just aggravated the problem of genetic drift and resulted in an even earlier convergence for most bit-neutral agents at the zero-bit solution.

9.7.4 A Final Assessment of the Two Mutation Operators

The previous analysis has shown that the original mutation operator is more efficient in combating genetic drift than the new bit-neutral operator. Its effective mutation rates for individual bits are higher and it is more disruptive. The difference in disruptiveness between the two operators is highest when the bit level has reached zero. For the bit-neutral mutation operator, the zero-bit level is truly an absorbing state. It becomes inactive and the genetic algorithm converges prematurely.

[16] The critical population size for the original mutation operator turns out to be only 28 trading rules.

Even though the built-in upward bias of the original mutation operator may be questionable at first sight, a deeper analysis reveals that it is superior to the bit-neutral design. With the benefit of hindsight, the new mutation operator did not effectively combat genetic drift; it even became subject to genetic drift itself. Since the new operator was designed to keep the bit fraction before and after mutation, on average, unaltered, the bit level in an agent's rule set can successively be changed in one direction several periods in a row by pure chance.[17] Eventually, one of the two absorbing states will be reached. With low initial bit probabilities, the absorbing state at the zero-bit level is the more likely one.

9.8 Detection of Emergence of Technical Trading

As it turns out, the problem with the original SFI-ASM is not in model design, but in the interpretation of certain simulation results. Given the original mutation operator's fixed point of one half, an isolated look at the aggregate bit level to detect the existence of technical trading is not sufficient. Without prior knowledge of the degree to which the increase in learning speed affects the aggregate bit level through the biased mutation operator, additional tests need to be done to establish the existence of technical trading in the model.

9.8.1 Predictability in the Price Series

Instead of providing hard and unquestionable evidence, LeBaron et al. often present arguments that do not establish the existence of technical trading in the model beyond any doubt. For instance, they showed that the simulated price series contain useful technical information that could be successfully exploited. LeBaron et al. regressed a simulated price series on lagged prices and dividends and added an extra explanatory variable, e.g.,

$$p_{t+1} + d_{t+1} = a + b(p_t + d_t) + cI_{t,MA500} + \epsilon_{t+1}. \tag{9.14}$$

This extra variable was either an indicator showing whether the price is above or below a 5-period or a 500-period moving average, or an indicator using the price dividend ratio. The results reported by LeBaron et al. are reproduced in table 9.3.

[17] Under no selection pressure and mutation only, the bit level in an agent's rule set would follow a random walk.

Table 9.3. Predictability in the simulated price time series as documented in [244, p. 1503].

Forecasting regressions

Description	Fast learning	Slow learning
$MA(5)$	0.009	− 0.008
	(0.013)	(0.007)
$MA(500)$	0.074	− 0.025
	(0.014)	(0.015)
$rP/D > 3/4$	− 0.443	0.050
	(0.104)	(0.093)

Note: Means over 25 runs. Numbers in parenthesis are standard errors estimated using the 25 runs.

LeBaron et al. find no extra predictability coming from the 5-period MA for both the slow and fast learning case. While not being significant for the slow learning case, the coefficients for the 500-period MA and the price dividend ratio are, on the other hand, statistically significant for the fast learning case. This points to some technical predictability in the price series that agents could exploit by using their classifier system.

While this regression demonstrates that there is indeed predictability remaining in the simulated price series, it does not show whether SFI agents are able to discover this predictability with their classifier system. Neither the increase in the aggregate bit level nor the sign of remaining predictability are convincing evidence whether agents really use specific trading rules more often than general ones.

9.8.2 Trading Bits and Fitness Values

To really establish whether set trading bits reflect technical trading or not, one could utilize the fitness information attached to each trading rule. If they indeed contain useful forecasting information, specific trading rules should then be more accurate and possess higher fitness values.

In order to test this hypothesis, 10 simulations were run over 500,000 periods in which 25 ordinary SFI agents were competing against 25 non-classifier agents. As learning speeds, GA-invocation intervals of 100 and 1,000 were chosen. There are a variety of significance tests possible with this setup. First, I tested the mean fitness values of all trading rules over all simulation runs. The mean fitness values of SFI agents in one simulation run were adjusted for bit costs, i.e., the values reported in table 9.4 show the pure forecast accuracy of trading rules, independent of their specificity.

Table 9.4. Comparison of bit cost adjusted mean fitness values for SFI- and non-classifier agents. for. Fitness values were averaged over all rules of one agent type within one simulation run.

Run	GA-1,000		GA-100	
	Mean Fitness Adjusted for Bit Costs for SFI Agents	Mean Fitness over All Non-Classifier Agents	Mean Fitness Adjusted for Bit Costs for SFI Agents	Mean Fitness over All Non-Classifier Agents
1	92.672	93.449	93.767	93.771
2	92.912	92.442	94.612	93.975
3	92.737	92.732	94.094	94.506
4	91.123	91.159	92.824	92.432
5	92.774	93.121	93.531	93.613
6	91.940	92.617	94.365	93.654
7	90.666	91.263	93.330	93.340
8	93.280	92.404	94.226	94.001
9	92.672	92.370	92.888	92.738
10	91.977	91.923	93.740	93.411
mean	**92.275**	**92.348**	**93.738**	**93.544**
std.-dev.	0.837	0.731	0.605	0.609
t-stat.	-0.430		1.792	

In the slow learning case, non-classifier agents developed rule sets that have a slightly higher mean fitness than those of their SFI counterparts. For the fast learning case, that relationship was reversed though. Even though not yet significant at the 5%-level, SFI agents seem to generate trading rules that indeed produce more accurate forecasts. The average specificity of their trading rules is about 4 trading bits per rule, with a small increase in specificity for the fast learning case.

While an analysis of the mean fitness values in the economy showed no evidence of advantageousness for technical trading in the slow learning regime and weak evidence for it in the fast learning case, the picture looks quite different if we only focus on the best trading rules. As a maximum fitness value for each simulation run in table 9.5, the bit cost-adjusted fitness value of the best trading rule in the whole economy per agent type was used. The use of only the best trading rules makes sense since agents use the best of all activated trading rules for

Table 9.5. Comparison of bit cost adjusted maximum fitness values for SFI- and non-classifier agents. For each simulation run, the maximum fitness value of all rules over all agents within the economy was used.

Run	**GA-1,000**		**GA-100**	
	Max Fitness Adjusted for Bit Costs for SFI Agents	Max Fitness of Non-Classifier Agents	Max Fitness Adjusted for Bit Costs for SFI Agents	Max Fitness of Non-Classifier Agents
1	97.390	97.113	96.474	96.258
2	96.881	96.462	97.181	96.977
3	96.827	96.467	96.759	96.594
4	96.812	95.012	95.544	95.112
5	96.875	96.470	96.325	96.031
6	96.646	96.346	96.933	96.401
7	96.085	94.939	96.248	95.418
8	96.729	96.403	97.099	96.528
9	96.998	96.811	95.998	95.272
10	96.393	95.896	96.309	95.899
mean	**96.664**	**96.192**	**96.487**	**96.049**
std.-dev.	0.459	0.713	0.513	0.620
t-stat.	5.16		6.11	

their forecast production. When using only the best trading rule per simulation run, one suddenly realizes that SFI agents are able to produce significantly better trading rules than non-classifier agents. Both test statistics from a paired t-test are highly significant for both the one- and the two-sided versions of the test. One also notices that in the slow learning regime, the best trading rules produce slightly better forecasts than in the fast learning case.[18]

The analysis of fitness information finally proves beyond a doubt that there is indeed technical trading in the model. A superficial look at the aggregate bit level or a reference to remaining predictability in the price series does not have the same power of persuasion than the fitness information stored for each trading rule. If the authors of the original SFI-ASM would have performed such a fitness analysis to begin with, no questions would have arisen about whether the continued existence of trading bits reflects technical trading or not.

[18] For the best trading rules of SFI agents, the t-statistics for an unpaired t-test with unknown and unequal variances is 0.81. The t-value is only 0.48 for the best trading rules of non-classifier agents, hence, in both cases, the forecasting quality in the slow learning regime is better, but not significantly better.

9.8.3 Equilibrium Bit Frequencies and Bit Fixations

While the previous section proves that there is technical trading in the model, it does not yet prove whether there is emergent technical trading for faster learning speeds.[19] The difference lies in the degree of technical trading at varying learning speeds. Is there really an increase in the amount of technical trading when agents learn faster, or does the augmented bit level reflect a rise in the background level of bit frequencies because of the biased mutation operator?

A definite answer to this question can be given when analyzing the bit frequencies at different GA-invocation intervals. Figure 9.6 on page 165 provides an approximation to the null model of equilibrium bit frequencies with which the bit frequencies of SFI agents in normal simulation runs should be compared. Remember that 9.6 was derived under bit costs of 0.005, but with no selection pressure stemming from forecast evaluation. Any differences in bit frequencies should then reflect the influence that the fitness evaluation of forecasts exerts on the bit dynamics.

Figure 9.12 shows the distribution of bit frequencies after 250,000 periods for 30 slow learning SFI agents. First, one notices that there about 10,000 fewer occurrences with a bit frequency of zero. The absolute occurrences, however, cannot be compared since we analyzed 64,000 bit positions at once in figure 9.6 (1,000 agents x 64 conditions), and only 48,000 bit positions in figure 9.12 (25 runs x 30 agents x 64 conditions). If we compare the relative commonness of zero-bit fixations, we have a drop from 0.735 down to 0.660. This reduction certainly implies an increase in bit frequencies greater than zero. Of main interest, however, is the change in complete or near bit fixations at the 100% level. While the null-model has only 0.02% possible bit fixations, 3.35% of all possible bit positions are completely filled with non-# bits under real simulation conditions. The increase compared to the null model is another irrefutable test that proves the existence of technical trading in the model.

The distribution of bit frequencies of SFI agents in the fast learning regime is shown in figure 9.13. To test whether there is indeed an increase in the degree of technical trading, we need to check whether there are more bit fixations at the 100% level compared to the slow

[19] The increased fitness differential between the best rules of SFI and non-classifier agents in the fast learning case suggests that there is a potential for higher degrees of technical trading. But again, the fitness values alone do not allow a statement about levels of technical trading at both learning speeds.

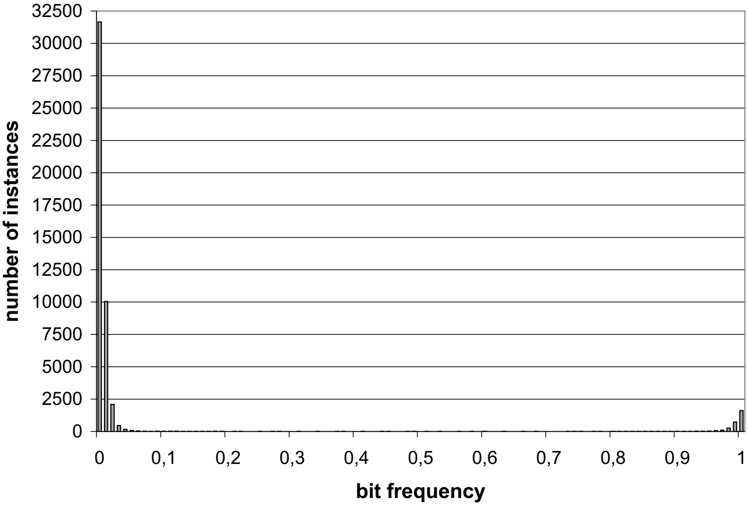


Fig. 9.12. Number of bit positions that have a certain bit frequency in the rule sets of slow learning SFI agents. Data were obtained at period 250,000 from 25 simulation runs with 30 SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs = 0.005, GA-interval = 1,000).

learning case.[20] This is definitely the case, with 4.46% of all possible bit positions fixed. When checking the significance of the increase from 3.35% to 4.46% with an unpaired t-test with unknown and unequal variances, the t-value turns out to be 8.9. Since this is significant at even the highest levels of confidence, we can finally establish the emergence of technical trading for faster learning speeds beyond a doubt.

9.9 An Evolutionary Perspective on Technical Trading

In section 9.8.2, we saw that the average specificity of successful trading rules is about 4 bits per rule, with a small increase in specificity for the fast learning case. Remember that we determined the necessary

[20] Even though we know from figure 6.1 on page 107 that the final equilibrium distributions of bit frequencies have not yet been reached, we also know that we can already test for differences in bit fixations. At period 250,000, the aggregate bit levels for the two GA-invocation intervals have already reached their relative position. The enormous requirements in terms of computational time led me to shorten the time horizon. Figure 6.1 suggests that in the final bit equilibrium, the difference between bit fixations will be even more pronounced than at period 250,000.

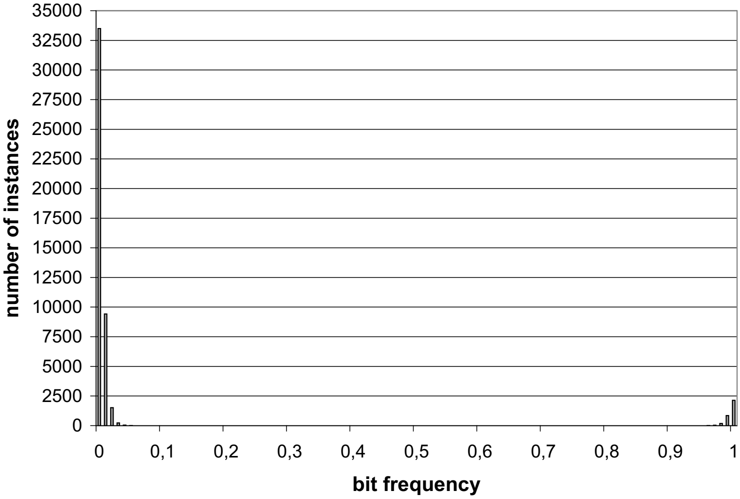


Fig. 9.13. Number of bit positions that have a certain bit frequency in the rule sets of fast learning SFI agents. Data were obtained at period 250,000 from 25 simulation runs with 30 SFI agents, each endowed with 100 trading rules ($\Pi = 0.9$, $\pi = 0.03$, bit costs = 0.005, GA-interval = 100).

fitness increase per single trading bit to be around 0.5 in order for selection to offset the effect of genetic drift (section 9.6.3 on page 161). With an average specificity of four, the fitness increases of only 0.2 for mean fitness and of about 0.5 for maximum fitness values in the fast learning case suggest that selection cannot guarantee that favorable trading bits will spread through the population. The observed effect of subpopulation differentiation, i.e., different agents discovering different trading rules, is most likely a result of insufficient selection pressure in favor of successful trading bits. With fitness gradients too small to counter genetic drift, it is hard to say which trading bits will be selected since bit fixation remains a noise-prone process.

While we have developed this argument within a model that is characterized by successful technical trading rules, the existence of technical trading strategies in real financial markets remains a puzzle for many adherents of the EMH. However, they can now use the same argument to explain that existence even in the presence of efficient markets.

Those who voiced the earliest evolutionary arguments in favor of emergent rationality or efficient markets seemed to have come from the selectionist camp. They implicitly assumed that selection is the only evolutionary force at work and, “*if things have had time to hammer* *logic into men*" [382, p. 80], a rational and/or efficient outcome will be inevitable. An evolutionary perspective that extends its scope beyond selection is not that certain anymore.

In the light of the above argument, the sheer amount of different technical trading strategies in real financial markets suggests two things. First, their possible profits are at best relatively small. Only highly profitable and easily detectable trading strategies could, in principle, spread through the whole population, which would then lead to their own demise. Second, to the extent that these trading strategies are unprofitable (or less profitable than the market portfolio), the losses they produce are equally too small to quickly eradicate them.[21] With only small selective forces in favor of or against most trading strategies, "financial evolution" cannot guarantee that only the fittest strategy will survive.

Under small selection pressures in either direction, random evolutionary forces other than selection might prevail and maintain a sufficiently diverse population of trading strategies over extended periods of time. Keeping in mind that a moderate increase in checked trading conditions in the SFI-ASM drastically prolonged the adjustment period towards equilibrium, the "evolutionary time scale" in real financial markets may be too long to observe evolutionary change in real time.[22] The existence of technical trading strategies does not necessarily prove their profitability. It does not disprove the validity of the EMH either. Technical trading rules may simply be a sign of insufficient selection pressures against them.

[21] Positive profits below the returns of the market portfolio do not strike one as a particulary strong selective force against a particularly trading strategy.

[22] A similar concern was voiced in [242].

10

Summary and Future Research

"... the outcome of any serious research can only be to make two questions grow where only one grew before."
Thorstein Veblen

It was argued in the beginning of this book that the current surge in agent-based simulation models in economics cannot be explained by the sudden increase in computational resources alone. Thus, the first part discussed agent-based simulations in the light of the limitations of current research methodologies and practices in the economics field. It was highlighted that agent-based simulation approaches are highly suited to address heterogeneous agents problems. Giving up the assumption of agent homogeneity necessarily leads to an abandoning of the rational expectations paradigm. In addressing what degree of rationality should be chosen instead, chapter 3 argued in favor of the principle of minimal rationality [70]. When considering cognitive resources as scarce, agents should be modeled as economizing on reasoning processes and employing only as much reasoning power as necessary to perform the task they are facing. It was shown at the end of chapter 3 that the trigger for the rational expectations revolution, the famous Phillips-curve debate, closely followed the principle of minimal rationality.

After discussing some stylized facts of financial markets in the light of the EMH, chapter 5 introduced several agent-based simulations. Through their attempt to replicate certain stylized facts, additional insight can be gained as to which assumptions are essential for the emergence of these stylized facts. The selection and sequencing of presentation of agent-based simulations was guided by the principle of minimal rationality, discussing first the zero-intelligence approach by

Gode and Sunder [150], followed by models with random communication structures which are, too, populated by agents with random demands. Models of chartist-fundamentalist interactions and many-strategy models with learning were the last to be discussed in chapter 5.

Within the hierarchy of agent rationality, the agents of the SFI-ASM, the main focus of this book, ranked at the top. Only because they were modeled as being heterogeneous in analyzing a common information set, the assumption of rational expectations had to be replaced by that of boundedly rational agents that use adaptive learning procedures. Since these agents derive their demands through optimization, the deviations from a typical neoclassical framework were only minimal.

The main part of this book analyzed a Java-replication of the original SFI-ASM. The replication of existing simulation models is, as Axelrod points out, an important, but often neglected step in the field of agent-based simulations [13]. Recreating a simulation model from scratch and confirming the basic results is much more powerful than simply utilizing an existing source code. Unlike many other models, the SFI-ASM, as one of the first and most famous artificial stock market simulations, has been extensively tested and it exists on several different programming platforms.

During the normal process of discovering and eliminating my own implementation errors, I stumbled upon an interesting model behavior. The unexpected and previously undocumented increase in the aggregate bit-level for higher mutation rates prompted me to analyze the GA in more detail in chapter 6. It was shown that the mutation operator had a fixed point of one half, i.e., it was upwardly-biased under normal model parameterization. This upward bias caused me to question whether the increase in the aggregate bit level for faster learning speeds is really a reflection of emergent technical trading. Since, at faster learning speeds, the mutation operator is called more often per time interval and injects new trading bits, the increase in the aggregate bit level could also be a design artifact.

An alternative mutation operator was suggested in chapter 7. This mutation operator has the property of being bit-neutral, i.e., the number of trading bits before and after mutation was left, on average, unaltered. The updated model behavior then supported the Marimon-Sargent Hypothesis with respect to bit usage, namely that adaptive boundedly rational agents in an artificial stock market would discover the homogeneous rational expectations equilibrium. Independent of learning speed, all agents discovered the zero-bit solution which temporarily suggested two things. First, the classifier system does not provide any informational advantage that agents could exploit. Second, the emergence of technical trading in the original SFI-ASM is an artifact of design assumption.

The tentative result of a useless classifier system stood in contradiction to two studies by Joshi, Parker, and Bedau [206, 207]. Based on the terminal wealth levels acquired by agents in the SFI-ASM, they performed two game-theoretic analyses. In the first study, they asked whether it would be optimal for agents to include technical trading rules or not. They concluded that technical trading constitutes a typical prisoner's dilemma. They argued that it is rational for each agent to use technical trading, yet at the same time, they found the aggregate market outcome to be less efficient than when agents had no access to technical trading rules. In the second study, they endogenized the learning speed of agents and found the optimal learning rate to be at a GA-invocation interval of 100. This clearly falls into the less efficient complex regime, and Joshi, Parker, and Bedau inferred that financial markets can operate at sub-optimal equilibria [207]. While the game-theoretic analyses themselves are correct, chapter 8 deduced that the SFI-ASM is highly unsuitable for addressing wealth issues. Because of taxation in the model, a programming trick to avoid a long-run explosion of wealth, a wedge between long-term wealth and maximized short-term wealth was introduced. It was analytically shown that the only source of wealth accumulation in the SFI-ASM is the aggregate risk premium that agents collect. Wealth contributions of the kind "buy low and sell high" are impossible. A series of simulations then demonstrated that wealth differences between different trader types generally arise for various reasons, none of which has any economic meaning.

In the final chapter, some simulation evidence was presented that showed the zero-bit solution to be too robust, i.e., it even emerged under parameterizations under which it should not have emerged. As a reason for this behavior, the effect of genetic drift was identified. Subsequently, the two mutation operators were analyzed in light of genetic drift. By transferring the Wright-Fisher model from population genetics to the field of genetic algorithms, conditions were derived under which selection dominates genetic drift. In a second step, a so-called neutral model from population genetics was used to analyze under which conditions mutation dominates genetic drift. While the original mutation operator proved to dominate genetic drift, the same could not be said for the suggested bit-neutral operator. In addition to being unable to effectively combat genetic drift, it even became subject to genetic drift itself.

Even though the original mutation operator looked suspicious at first, it turned out that there was no problem in model design as previously thought. With the benefit of hindsight, however, it became clear to me that there was a problem in the interpretation of the original simulation results. The aggregate bit level, which was used to bolster the argument of emergent technical trading, is highly unsuitable for advancing such claims, especially in the light of an upwardly-biased mutation operator.[1] The proper way to test for technical trading is to utilize the attached fitness information for each trading rule. To test for emergent technical trading, i.e., whether the level of technical trading is higher at faster learning speeds, one has to analyze whether the number of bit fixations at the 100%-level has significantly increased. Both tests unequivocally showed in the end that there is indeed emergent technical trading in the SFI-ASM.

The final disapproval of the bit-neutral mutation operator made it necessary to reconsider some of the claims and results that were presented in the course of this book. Even though the initial reason for reconsidering two game-theoretic analyses by Joshi, Parker, and Bedau turned out to be ungrounded, the finding that accumulated wealth levels should not be given any economic meaning in the SFI-ASM remains unaffected.

Therefore, future research again needs to tackle the problem of an optimal learning speed of agents. Since wealth levels are not appropriate in addressing this issue, one either has to find another venue for doing this, or one has to alter the economic structure such that accumulated wealth becomes economically meaningful. A related issue is the question whether there is a maximum learning speed beyond which the model structure would break down. A GA-invocation interval of one has been shown to be such a case.

Figure 7.2 on page 117 showed that bit-neutral agents took longer to arrive at the zero-bit solution the closer the stochastic dividend process came to a random walk. This increase in time suggests that there are

[1] It was the aggregate bit level argument that caused me to suggest the bit-neutral mutation operator. My own statements in favor of the bit-neutral mutation operator were, too, based on the aggregate bit level. The approval of my arguments from the scientific economics community at various conferences, the acceptance of an article to be published in a leading economics journal, and the absence of a well-founded rebuttal by the original authors show me that there was, indeed, substantial room for misinterpretation, and that the dynamics of GA-behavior were not fully understood by the groups involved.

more possibilities for emerging patterns to occur at near random walk behavior. It is, therefore, my hypothesis that a similar effect should be prevalent with the original mutation operator, i.e., the closer the dividend is to a random walk, the higher the degree of technical trading in the model. Having now pointed to a proper way of testing the degree of technical trading in the model, this hypothesis could be easily tested in further research.

An important task that remains to be done is the calibration of the model to real empirical data. This includes refining the concept of a trading period and possibly allowing for intra-period trading. Replacing the CARA-assumption with constant relative risk aversion would allow for wealth to impact the optimal demands.[2] Instead of using myopic agents, agents could be modeled as having fully intertemporal preferences. While classifier systems lend themselves to an easy interpretation and analysis of the learning dynamics, LeBaron points out that they remain a controversial modeling tool in economics [242]. Other approaches such as using neural networks for forecast production seem to be becoming more popular [240, 241]. It is however, likely, that in addressing some or all of these issues, the economic structure of the model has to be altered in a way that one should no longer speak of "the SFI-ASM".

Once the various approaches of designing agent-based artificial markets result in an adoption of commonly accepted principles or building blocks, it is likely that agent-based financial markets will be able to address practical problems. Farmer predicts that in several years, they will be used for investment applications [120] while LeBaron conceives of a role for agent-based models in evaluating the stability and efficiency of different trading mechanisms [239]. Until then, however, much work remains to be done.

[2] CRRA preferences are used, for instance, in [247, 241].

11

Appendix

11.1 Timing in the Stock Market

The general sequence of activities in the SFI-ASM is summarized and represented in figure 11.1. At the beginning of period t, a new dividend d_t is announced. For each agent $i = 1, \ldots, N$, the GA is invoked with probability $1/K$ to change an agent's rule set $\Im_i$. Afterwards, agents determine their active rule set $\omega_i \subseteq \Im_i$ by comparing the condition parts of each trading rule $j = 1, \ldots, 100$ with the binary market descriptor. From that active rule set, agents select the rule with the highest forecast accuracy for their forecast production. The price formation process is then initiated by the specialist who announces a trial price p_{trial} equal to last period's price. Based on this trial price, agents form their expectation about next period's price and dividend, determine their optimal demand for the risky stock, and then submit their offers and bids to the specialist. If the bids and offers cannot be matched, the specialist determines a trial price and the whole process starts all over. This iterative process ends when the offers are balanced by the submitted bids or after 10 unsuccessful trial rounds. In the latter case, one side of the market will be proportionally rationed. The last trial price is announced to be the stock price for period t and all trades between agents are executed at that price. Since at this time, the dividend and the stock price in period t are known to the agents, they are now able to update the forecast performance of all rules which were activated in the last period and actually produced a forecast about this period's price and dividend. After completion, a new period starts when the next stock dividend is revealed.

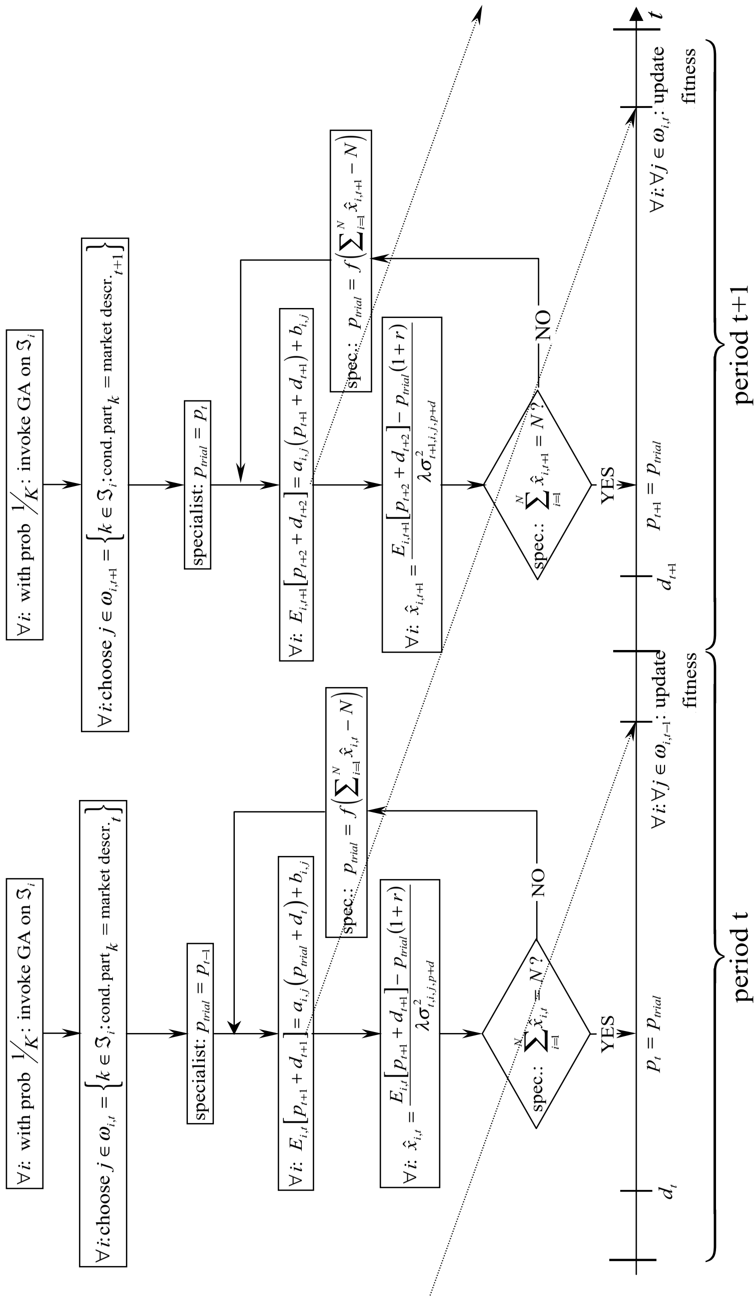


Fig. 11.1. Timing sequence of major activities in the Santa Fe Institute Artificial Stock Market

11.2 Fundamental and Technical Trading Bits

This section documents the differences in condition bits used by several SFI-ASM versions. Furthermore, it explains why certain changes are made to the Java version of the SFI-ASM.

Table 11.1. Condition bits documented in [11, p. 27] and [244, p. 1494].

Bit	Condition
1	Price $*$ interest/dividend $> 1/4$
2	Price $*$ interest/dividend $> 1/2$
3	Price $*$ interest/dividend $> 3/4$
4	Price $*$ interest/dividend $> 7/8$
5	Price $*$ interest/dividend > 1
6	Price $*$ interest/dividend $> 9/8$
7	Price $>$ 5-period MA
8	Price $>$ 10-period MA
9	Price $>$ 100-period MA
10	Price $>$ 500-period MA
11	On: 1
12	Off: 0

Arthur et al. [11] and LeBaron et al. [244] report only 12 conditions for the SFI-ASM. Bits 1 to 6 are the fundamental bits. Note that even though these are 6 individual bits, they can only code one piece of fundamental information since they logically belong together. The rule consistency check introduced in section 7.4.3 compresses this bit sequence to a maximum of 2 bits, and a neural network trading rule would only use one real valued number. Bits 7 to 10 are the technical trading bits. LeBaron et al. remark that removing 3 out of the 4 technical bits can have a big impact, but removing only one will not change things. They do not, however, mention what would happen if the list of trading bits were substantially extended.

The last two bits convey no useful market information, but Arthur et al. explain that they were intended as experimental controls to check to which degree agents act upon useless information. They also assert that these two bits would allow them to detect emergence of technical trading "*if bits 7–10 become set significantly more often, statistically, than the control bits*" [11, p. 27]. Emergence of technical trading in the SFI-ASM has surely been declared without ever having done such a test. Since the two control bits are always set, i.e., they are never #, the "benchmark" of control bits is always 100 %. It is, therefore, impossible for any trading bit to be set significantly more often. Since these bits are not mentioned anywhere in the discussions of simulation results, they are not implemented in the Java version of the SFI-ASM.

Table 11.2. First part of the condition bits in the SFI-ASM Objective-C version 7.1.2 used by Joshi et al. (1998 and 2002).

Bit	Condition
1	dummy bit – always on
2	dummy bit – always off
3	dummy bit – random on or off
4	dividend went up this period
5	dividend went up one period ago
6	dividend went up two periods ago
7	dividend went up three periods ago
8	dividend went up four periods ago
9	5-period moving average of dividend went up
10	20-period moving average of dividend went up
11	100-period moving average of dividend went up
12	500-period moving average of dividend went up
13	dividend $>$ 5 period moving average
14	dividend $>$ 20 period moving average
15	dividend $>$ 100 period moving average
16	dividend $>$ 500 period moving average
17	dividend: 5-period moving average $>$ 20-period moving average
18	dividend: 5-period moving average $>$ 100-period moving average
19	dividend: 5-period moving average $>$ 500-period moving average
20	dividend: 20-period moving average $>$ 100-period moving average
21	dividend: 20-period moving average $>$ 500-period moving average
22	dividend: 100-period moving average $>$ 500-period moving average
23	dividend/mean-dividend $> 1/4$
24	dividend/mean-dividend $> 1/2$
25	dividend/mean-dividend $> 3/4$
26	dividend/mean-dividend $> 7/8$
27	dividend/mean-dividend > 1
28	dividend/mean-dividend $> 9/8$
29	dividend/mean-dividend $> 5/4$
30	dividend/mean-dividend $> 3/2$
31	dividend/mean-dividend > 2
32	dividend/mean-dividend > 4

The Objective-C version 7.1.2 used by Joshi et al. [206], which served as the blueprint for the current Java implementation of the

FI-ASM, had a total of 61 conditions. These conditions are documented in tables 11.2 and 11.3. Besides checking many more condition bits than in the model version by LeBaron et al., the ratios for the price*interest/dividend (pr_f/d) bit sequence were also changed by Joshi et al.

Table 11.3. Second part of the condition bits in the SFI-ASM Objective-C version 7.1.2 used by Joshi et al. (1998 and 2002).

Bit	Condition
33	price*interest/dividend > 1/4
34	price*interest/dividend > 1/2
35	price*interest/dividend > 3/4
36	price*interest/dividend > 7/8
37	price*interest/dividend > 1
38	price*interest/dividend > 9/8
39	price*interest/dividend > 5/4
40	price*interest/dividend > 3/2
41	price*interest/dividend > 2
42	price*interest/dividend > 4
43	price went up this period
44	price went up one period ago
45	price went up two periods ago
46	price went up three periods ago
47	price went up four periods ago
48	5-period moving average of price went up
49	20-period moving average of price went up
50	100-period moving average of price went up
51	500-period moving average of price went up
52	price > 5-period moving average
53	price > 20-period moving average
54	price > 100-period moving average
55	price > 500-period moving average
56	price: 5-period moving average > 20-period moving average
57	price: 5-period moving average > 100-period moving average
58	price: 5-period moving average > 500-period moving average
59	price: 20-period moving average > 100-period moving average
60	price: 20-period moving average > 500-period moving average
61	price: 100-period moving average > 500-period moving average

The cutoff points for the pr_f/d-ratios had been chosen by LeBaron et al. to have good coverage of the range of relevant ratios. Note that the cutoff points around one are non-symmetrical, i.e., it is very unlikely for the stock to be highly overvalued. LeBaron et al. reported that the extreme events at 1/4 and 9/8 were visited with probabilities of less than 0.01. Hence, the bits 39 – 42 used in Joshi et al. can be safely neglected. Since the dividend process is symmetrical around its mean, the $d/\bar{d}$-ratios are also symmetrical. Applying the same principle of good coverage to the dividend/mean-dividend ratios ($d/\bar{d}$), however, led me to propose cutoff points for the $d/\bar{d}$-ratios that are much closer than those used by Joshi et al. The condition bits for the Java version of the SFI-ASM are shown in table 11.4.

The reprogrammed SFI-ASM differs in the number of checked conditions. All conditions labeled as fundamental are coded in a long integer which consists of 8 bytes (64 bits). Since a trading bit can take on three allele values (0, 1, and #), one encoded trading condition requires 2 bytes. Therefore, a long integer can hold 32 conditions. One long integer is used to hold all 32 fundamental conditions; a second holds the 32 technical conditions.

Table 11.4. Fundamental and technical condition bits in the current Java version of the SFI-ASM.

Bit	Fundamental conditions	Technical conditions
0	dividend / dividend-mean > 0.6	price / price-mean > 0.25
1	dividend / dividend-mean > 0.8	price / price-mean > 0.5
2	dividend / dividend-mean > 0.9	price / price-mean > 0.75
3	dividend / dividend-mean > 1.0	price / price-mean > 0.875
4	dividend / dividend-mean > 1.1	price / price-mean > 1.0
5	dividend / dividend-mean > 1.12	price / price-mean > 1.125
6	dividend / dividend-mean > 1.4	price / price-mean > 1.25
7	price * interest / dividend > 0.25	price went up this period
8	price * interest / dividend > 0.5	price went up one period ago
9	price * interest / dividend > 0.75	price went up two periods ago
10	price * interest / dividend > 0.875	price went up three periods ago
11	price * interest / dividend > 0.95	price went up four periods ago
12	price * interest / dividend > 1.0	5-period price-MA went up
13	price * interest / dividend > 1.125	10-period price-MA went up
14	dividend went up this period	20-period price-MA went up
15	dividend went up one period ago	100-period price-MA went up
16	dividend went up two periods ago	500-period price-MA went up
17	dividend went up three period ago	price > 5-period price-MA
18	5-period dividend MA went up	price > 10-period price-MA
19	10-period dividend MA went up	price > 20-period price-MA
20	100-period dividend MA went up	price > 100-period price-MA
21	500-period dividend MA went up	price > 500-period price-MA
22	dividend > 5-period div.-MA	price: 5-period MA > 10-period
23	dividend > 10-period div.-MA	price: 5-period MA > 20-period
24	dividend > 100-period div.-MA	price: 5-period MA > 100-period
25	dividend > 500-period div.-MA	price: 5-period MA > 500-period
26	div.: 5-period MA > 10-period MA	price: 10-period MA > 20-period
27	div.: 5-period MA > 100-period MA	price: 10-period MA > 100-period
28	div.: 5-period MA > 500-period MA	price: 10-period MA > 500-period
29	div.: 10-period MA > 100-period MA	price: 20-period MA > 100-period
30	div.: 10-period MA > 100-period MA	price: 20-period MA > 500-period
31	div.: 100-period MA > 500-period MA	price: 100-period MA > 500-period

References

[1] C. Ahn, E.-B. Lee, and E.-H. Suh. Noise trading in the Korean foreign exchange market: Some questionnaire evidence. *Bank of Korea Economic Papers*, 5(2):133–155, 2002.

[2] V. Akgiray, G. G. Booth, and O. Loistl. Stable laws are inappropriate for describing German stock returns. *Allgemeines Statistisches Archiv*, 73:115–121, 1989.

[3] A. A. Alchian. Uncertainty, evolution, and economic theory. *Journal of Political Economy*, 58(3):211–222, 1950.

[4] M. Allais. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, 21(4):503–546, 1953.

[5] F. Allen and R. Karjalainen. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51(2):245–271, 1999.

[6] F. Allen, S. Morris, and H. S. Shin. Beauty contests, bubbles and iterated expectations in asset markets. *Cowles Foundation Discussion Paper 1406*, 1406, March 2003.

[7] H. L. Allen and M. P. Taylor. Charts, noise, and fundamentals in the London foreign exchange market. *Economic Journal*, 100(400):49–59, 1990.

[8] H. M. Amman, D. A. Kendrick, and J. Rust, editors. *Handbook of Computational Economics*, volume 1. Elsevier, Amsterdam, New York, 1996.

[9] J. Arifovic. Genetic algorithm learning and the cobweb model. *Journal of Economic Dynamics and Control*, 18(1):3–28, 1994.

[10] B. W. Arthur, S. Durlauf, and D. Lane, editors. *The Economy as an Evolving Complex System II.* SFI Studies in the Sciences of Complexity, vol. 27. Addison-Wesley, Reading, MA, 1997.

[11] W. B. Arthur, J. H. Holland, B. LeBaron, R. Palmer, and P. Tayler. Asset pricing under endogenous expectations in an artificial stock market. In W. B. Arthur, S. N. Durlauf, and D. A. Lane, editors, *The Economy as an Evolving Complex System II*, volume 27 of *SFI Studies in the Sciences of Complexity*, pages 15–44. Santa Fe Institute, Addison-Wesley, 1997.

[12] R. J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, 1976.

[13] R. Axelrod. Advancing the art of simulation in the social sciences: Obtaining, analyzing, and sharing results of computer models. *Complexity*, 3(2):16–22, 1997.

[14] C. Azariadis and R. Guesnerie. Sunspots and cycles. *Review of Economic Studies*, 53(5):725–737, 1986.

[15] L. Bachelier. Théorie de la spéculation. *Annales Scientifiques de lÉcole Normale Supérieure*, 17:21–86, 1900.

[16] T. Bäck. *Evolutionary Algorithms in Theory and Practice.* Oxford University Press, New York, 1996.

[17] T. Bäck. Mutation: Binary strings. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C3.2.1. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[18] R. T. Baillie, T. Bollerslev, and H. O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.

[19] P. Bak, M. Paczuski, and M. Shubik. Price variations in a stock market with many agents. *Physica A*, 246:430–453, 1997.

[20] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Physical Review A*, 38:364–373, 1988.

[21] R. Ball. Anomalies in relationships between securities' yields and yield-surrogates. *Journal of Financial Economics*, 6(2/3):103–126, 1978.

[22] S. C. Bankes. Agent-based modeling: A revolution? *Proceedings of the National Academy of Science of the USA*, 99:7199–7200, 2002.

[23] Y. Bar-Yam. *Dynamics of Complex Systems.* Addison-Wesley, Reading, MA, 1997.

[24] N. Barberis, A. Shleifer, and R. Vishny. A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343, 1998.

[25] R. J. Barro. Rational expectations and macroeconomics in 1984. *American Economic Review, Papers and Proceedings*, 74(2):179–182, May 1984.

[26] T. A. Bass. *The Predictors.* Henry Holt, New York, 1999.

[27] D. Beasley, D. R. Bull, and R. R. Martin. An overview of Genetic Algorithms: Part 1, Fundamentals. *University Computing*, 15(2):58–69, 1993.

[28] G. S. Becker. Irrational behavior and economic theory. *Journal of Political Economy*, 70(1):1–13, February 1962.

[29] J. Benhabib and R. E. A. Farmer. Indeterminacy and sunspots in macroeconomics. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1a, chapter 6, pages 387–448. North-Holland, Amsterdam, New York, 1999.

[30] A. K. Bera and C. M. Jarque. Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*, 7(4):313–318, 1981.

[31] H.-G. Beyer. On the dynamics of EAs without selection. In W. Banzhaf and C. Reeves, editors, *Foundations of Genetic Algorithms*, volume 5, pages 5–26. Morgan Kaufmann, 1999.

[32] H. S. Bhamra. Imitation in financial markets. *International Journal of Theoretical and Applied Finance*, 3(3):473–478, 2000.

[33] B. R. Biais and P. Bossaerts. Asset prices and trading volume in a beauty contest. *Review of Economic Studies*, 65(2):307–340, 1998.

[34] B. R. Biais and R. Shadur. Darwinian selection does not eliminate irrational traders. *European Economic Review*, 44(3):469–490, 2000.

[35] S. Bikhchandani and S. Sharma. Herd behavior in financial markets: A review. *IMF Working Paper*, 00/48, 2000.

[36] F. Black. Noise. *Journal of Finance*, 41(3):529–543, 1986.

[37] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[38] T. Blickle. Tournament selection. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 2.3. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[39] T. Blickle and L. Thiele. A comparison of selection schemes used in genetic algorithms. TIK Report 11, Version 2, Computer Engineering and Communication Networks Lab TIK, Zurich, December 1995.

[40] L. E. Blume and D. Easley. Learning to be rational. *Journal of Economic Theory*, 26(2):340–351, 1982.

[41] L. E. Blume and D. Easley. Evolution and market behavior. *Journal of Economic Theory*, 58(1):9–40, 1992.

[42] L. E. Blume and D. Easley. Evolution and rationality in competetive markets. In A. Kirman and M. Salmon, editors, *Learning and Rationality in Economics*, chapter 11, pages 324–342. Blackwell, Oxford, Cambridge, 1995.

[43] L. E. Blume and D. Easley. Optimality and natural selection in markets. *Journal of Economic Theory*, 107(1):95–135, 2002.

[44] N. Boitout and L. Ureche-Rangau. Towards a multifractal paradigm of stochastic volatility? *International Journal of Theoretical and Applied Finance*, 7(7):823–851, 2004.

[45] T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[46] T. Bollerslev, R. Y. Chou, and K. F. Kroner. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52(1):5–59, 1992.

[47] L. B. Booker, D. B. Fogel, D. Whitley, and P. J. Angeline. Recombination. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 3.3. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[48] T. Börgers. On the relevance of learning and evolution to economic theory. *Economic Journal*, 106(438):1374–1385, 1996.

[49] K. E. Boulding. *Evolutionary Economics.* Sage, Beverly Hills and London, 1981.

[50] G. H. Bower and E. R. Hilgard. *Theories of Learning.* Prentice Hall, Englewood Cliffs, N.J., 1981.

[51] T. Brenner. Can evolutionary algorithms describe learning processes? *Journal of Evolutionary Economics*, 8(3):271–283, 1998.

[52] T. Brenner. *Modelling Learning in Economics.* Edward Elgar, Cheltenham, Northampton, 1999.

[53] T. Brenner. Agent learning representation: Advice in modelling economic learning. Papers on Economics & Evolution 0416, Max Planck Institute for Research into Economic Systems, Evolutionary Economics Group, 2004.

[54] W. A. Brock and C. H. Hommes. A rational route to randomness. *Econometrica*, 65(5):1059–1095, September 1997.

[55] W. A. Brock and C. H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8-9):1235–1274, 1998.

[56] W. A. Brock, J. Lakonishok, and B. LeBaron. Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance*, 47(5):1731–1764, 1992.

[57] W. A. Brock and B. LeBaron. A dynamic structural model for stock return volatility and trading volume. *Review of Economics and Statistics*, 78(1):94–110, 1996.

[58] W. A. Brock, J. A. Scheinkmann, W. D. Dechert, and B. LeBaron. A test for independence based on the correlation dimension. *Econometric Reviews*, 15(3):197–235, 1996.

[59] H. S. Buscher. Angewandte Zeitreihenanalyse. In M. Schröder, editor, *Finanzmarkt-Ökonometrie: Basistechniken, Fortgeschrittene Verfahren, Prognosemodelle*, chapter III, pages 131–212. Schäffer-Poeschel, Stuttgart, 2002.

[60] C. D. Bustamante. Population genetics of molecular evolution. In R. Nielsen, editor, *Statistical Methods in Molecular Evolution*, Statistics for Biology and Health, chapter 4. Springer, New York, 2005.

[61] R. J. Caballero. A fallacy of composition. *American Economic Review*, 82(5):1279–1292, 1992.

[62] P. Cagan. The dynamics of hyperinflation. In M. Friedman, editor, *Studies in the Quantity Theory of Money*, pages 25–117. University of Chicago Press, Chicago, London, 1956.

[63] L. Calvet and A. Fisher. Multifractality in asset returns: Theory and evidence. *Review of Economics and Statistics*, 84(3):381–406, 2002.

[64] C. Camerer. Individual decision making. In J. H. Kagel and A. E. Roth, editors, *The Handbook of Experimental Economics*, chapter 8, pages 587–703. Princeton University Press, Princeton, N.J., 1995.

[65] J. Y. Campbell, A. W. Lo, and A. C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ, 1997.

[66] I. Chang, D. Stauffer, and R. B. Pandey. Asymmetries, correlations and fat tails in percolation market model. *International Journal of Theoretical and Applied Finance*, 5(6):585–597, 2002.

[67] S.-H. Chen, T. Lux, and M. Marchesi. Testing for non-linear structure in an artificial financial market. *Journal of Economic Behavior and Organization*, 46(3):327–342, 2001.

[68] S.-H. Chen and C.-H. Yeh. 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(3-4):363–393, 2001.

[69] S.-H. Chen and C.-H. Yeh. On the emergent properties of artificial stock markets: the efficient market hypothesis and the ratio-

nal expectations hypothesis. *Journal of Economic Behavior and Organization*, 49(2):217–239, 2002.

[70] C. Cherniak. *Minimal Rationality.* MIT Press, Cambridge, MA, 1986.

[71] Y.-W. Cheung and M. D. Chinn. Macroeconomic implications of the beliefs and behavior of foreign exchange traders. *NBER working paper 7417*, 7417, 1999.

[72] Y.-W. Cheung, M. D. Chinn, and I. W. Marsh. How do UK-based foreign exchange dealers think their market operates? *CEPR Discussion Paper 2230*, 2230, 1999.

[73] C. Chiarella, R. Dieci, and L. Gardini. Speculative behaviour and complex asset price dynamics: A global analysis. *Journal of Economic Behavior and Organization*, 49(2):173–197, 2002.

[74] P. K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–156, 1973.

[75] D. Cliff and J. Bruten. More than zero intelligence needed for continuous double-auction trading. Technical Report HPL-97-157, Hewlett-Packard Laboratories, Bristol, UK, December 1997.

[76] D. Cliff and J. Bruten. Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets. Technical Report HPL-97-141, Hewlett-Packard Laboratories, Bristol, UK, 1997.

[77] K. C. Cole. *The Universe and the Teacup: The Mathematics of Truth and Beauty.* Harcourt Brace, New York, 1998.

[78] N. Collier, T. Howe, and M. North. Onward and upward: The transition to Repast 2.0. In *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference, Electronic Proceedings*, Pittsburgh, PA, 2003.

[79] J. Conlisk. Why bounded rationality? *Journal of Economic Literature*, 34(2):669–700, June 1996.

[80] J. K. Conner and D. L. Hartl. *A Primer of Ecological Genetics.* Sinauer Associates, Sunderland, MA, 2004.

[81] R. Cont and J.-P. Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4(2):170–196, 2000.

[82] P. H. Cootner, editor. *The Random Character of Stock Market Prices.* MIT Press, Cambridge, MA, 1964.

[83] T. E. Copeland. A model of asset trading under the assumption of sequential information arrival. *Journal of Finance*, 31(4):1149–1168, 1976.

[84] K. Cuthbertson. *Quantitative Financial Economics: Stocks, Bonds, and Foreign Exchange.* Wiley, Chichester, UK, 1996.

[85] D. M. Cutler, J. M. Poterba, and L. H. Summers. Speculative dynamics and the role of feedback traders. *American Economic Review, Papers and Proceedings*, 80(2):63–68, 1990.

[86] R. B. D'Agostino and M. A. Stephens. *Goodness-of-fit Techniques.* Marcel Dekker, New York, 1986.

[87] F. J. Damerau and B. B. Mandelbrot. Tests of the degree of word clustering in samples of written english. *Linguistics*, 102:58–75, 1973.

[88] K. Daniel, D. Hirshleifer, and A. Subrahmanyam. Investor psychology and security market under- and overreactions. *Journal of Finance*, 53(6):1839–1885, 1998.

[89] H. Dawid. *Adaptive Learning by Genetic Algorithms.* Springer, Berlin et al., 2nd, revised and enlarged edition, 1999.

[90] R. Dawkins. *The Selfish Gene.* Oxford University Press, Oxford, UK, 1976.

[91] K. De Jong. Adaptive system design: A genetic approach. *IEEE Transactions on Systems, Man and Cybernetics*, 10(9):566–574, Sept. 1980.

[92] K. De Jong. Genetic algorithms are not function optimizers. In L. D. Whitley, editor, *Foundations of Genetic Algorithms*, pages 5–17. Morgan Kaufmann, San Mateo, CA, 1993.

[93] K. De Jong, D. B. Fogel, and H.-P. Schwefel. A history of evolutionary computation. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter A2.3. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[94] K. Deb. Encoding and decoding functions. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C4.2. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[95] W. F. M. DeBondt and R. Thaler. Does the stock market overreact? *Journal of Finance*, 40(3):793–805, 1985.

[96] J. B. DeLong, A. Shleifer, L. H. Summers, and R. J. Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, August 1990.

[97] J. B. DeLong, A. Shleifer, L. H. Summers, and R. J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance*, 45(2):379–395, June 1990.
[98] J. B. DeLong, A. Shleifer, L. H. Summers, and R. J. Waldmann. The survival of noise traders in financial markets. *Journal of Business*, 64(1):1–19, 1991.
[99] A. Devenow and I. Welch. Rational herding in financial economics. *European Economic Review*, 40(3-5):603–615, 1996.
[100] E. Dimson and M. Mussavian. A brief history of market efficiency. *European Financial Management*, 4(1):91–103, 1998.
[101] K. Dopfer, editor. *Evolutionary Economics: Program and Scope.* Kluwer, Boston, 2001.
[102] G. Dosi and R. R. Nelson. An introduction to evolutionary theories in economics. *Journal of Evolutionary Economics*, 4(3):153–172, 1994.
[103] J. Duffy. Agent-based models and human subject experiments. In K. L. Judd and L. Tesfatsion, editors, *Handbook of Computational Economics*, volume 2, chapter 4. North-Holland, Amsterdam, 2006.
[104] F. Y. Edgeworth. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences.* Augustus M. Kelley, New York, 1881.
[105] N. Ehrentreich. The Santa Fe Artificial Stock Market reexamined: Suggested corrections. Betriebswirtschaftliche Diskussionsbeiträge 45/02, Martin Luther University Halle-Wittenberg, 2002.
[106] N. Ehrentreich. Technical trading in the Santa Fe Institute Artificial Stock Market revisited. *Journal of Economic Behavior and Organization*, 61(4):599–616, 2006.
[107] R. N. Elliott. *The Wave Principle.* Elliott, New York, 1938.
[108] R. N. Elliott. *Nature's Law—The Secret of the Universe.* Elliott, New York, 1946.
[109] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75(4):643–669, 1961.
[110] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1008, 1982.
[111] T. W. Epps and M. L. Epps. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica*, 44(2):305–321, 1976.

[112] L. J. Eshelman. Genetic algorithms. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter B1.2. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[113] E. F. Fama. Mandelbrot and the stable Paretian hypothesis. *Journal of Business*, 36(4):420–429, 1963.

[114] E. F. Fama. The behavior of stock-market prices. *Journal of Business*, 38(1):34–105, 1965.

[115] E. F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417, 1970.

[116] E. F. Fama. Efficient capital markets II. *Journal of Finance*, 46(5):1575–1617, 1991.

[117] E. F. Fama, L. Fisher, M. C. Jensen, and R. Roll. The adjustment of stock prices to new information. *International Economic Review*, 10(1):1–21, 1969.

[118] E. F. Fama and K. R. French. Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1):3–25, 1988.

[119] J. D. Farmer. Physicists attempt to scale the ivory towers of finance. *Computing in Science & Engineering (IEEE)*, 1(6):26–39, 1999.

[120] J. D. Farmer. Toward agent-based models for investment. In *Developments in Quantitative Investment Models*, pages 61–71, AIMR
http://www.santafe.edu/˜jdf/papers/aimr.pdf, 2001.

[121] J. D. Farmer and S. Joshi. The price dynamics of common trading strategies. *Journal of Economic Behavior and Organization*, 49(2):149–171, 2002.

[122] J. D. Farmer and A. W. Lo. Frontiers of finance: Evolution and efficient markets. *Proceedings of the National Academy of Science of the USA*, 96:9991–9992, 1999.

[123] J. D. Farmer, P. Patelli, and lija I. Zovko. The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Science of the USA*, 102(6):2254–2259, 2005.

[124] M. J. Farrell. Profitable speculation. *Economica*, 33(130):183–193, 1966.

[125] E. Fehr and J.-R. Tyran. Individual irrationality and aggregate outcomes. *Journal of Economic Perspectives*, 19(4):43–66, 2005.

[126] J. Felsenstein. Theoretical Evolutionary Genetics – Genome 562. http://evolution.gs.washington.edu/pgbook/pgbook.html, March 2005.

[127] F. Fernandez-Rodriguez, C. Gonzalez-Martel, and S. Sosvilla-Rivero. On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. *Economics Letters*, 69(1):89–94, 2000.

[128] R. A. Fisher. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford, 1930.

[129] S. Focardi, S. Cincotti, and M. Marchesi. Self-organization and market crashes. *Journal of Economic Behavior and Organization*, 49(2):241–267, 2002.

[130] T. C. Fogarty. Varying the probability of mutation in the genetic algorithm. In J. D. Schaffer, editor, *Proc. 3rd Int. Conf. on Genetic Algorithms*, pages 104–109. Morgan Kaufmann, San Mateo, CA, 1989.

[131] D. B. Fogel. Other selection methods. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 2.6. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[132] C. Forner and J. Marhuenda. Contrarian and momentum strategies in the Spanish stock market. *European Financial Management*, 9(1):67–88, 2003.

[133] J. Forrester. *Industrial Dynamics*. MIT Press, Cambridge, MA, 1961.

[134] J. Forrester. *Urban Dynamics*. MIT Press, Cambridge, MA, 1969.

[135] J. Forrester. *World Dynamics*. Wright-Allen Press, Cambridge, MA, 1971.

[136] F. D. Foster and S. Viswanathan. Variations in trading volume, return volatility and trading costs: Evidence on recent price formation models. *Journal of Finance*, 48(1):187–211, 1993.

[137] C. Fourgeaud, C. Gourieroux, and J. Pradel. Learning procedures and convergence to rationality. *Econometrica*, 54(4):845–868, July 1986.

[138] J. A. Frankel and K. A. Froot. Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review, Papers and Proceedings*, 80(2):181–185, 1990.

[139] D. Friedman. Monty hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review*, 88(4):933–946, 1998.

[140] M. Friedman. The methodology of positive economics. In *Essays in Positive Economics*, pages 3–43. University of Chicago Press, Chicago, 1953.

[141] M. Friedman. The role of monetary policy. *American Economic Review*, 58(1):1–17, 1968.

[142] M. Friedman. Nobel lecture: Inflation and unemployment. *Journal of Political Economy*, 85(3):451–472, 1977.

[143] A. R. Gallant, P. E. Rossi, and G. Tauchen. Stock prices and volume. *Review of Financial Studies*, 5(2):199–242, 1992.

[144] M. Gardner. Mathematical games: The curious properties of the Gray code and how it can be used to solve puzzles. *Scientific American*, 227(2):106–109, August 1972.

[145] A. Gaunersdorfer. Endogenous fluctuations in a simple asset pricing model with heterogeneous agents. *Journal of Economic Dynamics and Control*, 24(5-7):799–831, 2000.

[146] M. Gell-Mann. What is complexity? *Complexity*, 1(1), 1995.

[147] R. Gençay, F. Selçuk, and A. Ulugülyağci. High volatility, thick tails, and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, 33(2):337–356, 2003.

[148] N. Gilbert and S. Bankes. Platforms and methods for agent-based modeling. *Proceedings of the National Academy of Science of the USA*, 99(3):7197–7198, 2002.

[149] J. Gleick. *Chaos: Making a New Science.* Penguin, New York, 1987.

[150] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual irrationality. *Journal of Political Economy*, 101(1):119–137, Feb. 1993.

[151] D. Goldbaum. Cycles of market stability and instability due to endogenous use of technical trading rules. In Y. S. Abu-Mostafa, B. LeBaron, A. W. Lo, and A. S. Weigand, editors, *Computational Finance 1999.* MIT Press, New York and Cambridge, MA, 2000.

[152] D. Goldbaum. Profitable technical trading rules ans a source of price instability. *Quantitative Finance*, 3(3):220–229, 2003.

[153] D. Goldbaum. Market efficiency and learning in an endogenously unstable environment. *Journal of Economic Dynamics and Control*, 29(5):953–978, 2005.

[154] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Reading, MA, 1989.

[155] D. E. Goldberg and J. Richardson. Genetic algorithms with sharing for multimodal function optimization. In J. J. Grefenstette, editor, *Genetic Algorithms and their Applications: Proceedings of the Second International Conference on Genetic Algorithms Juy*

28-31, 1987 MIT Cambridge, pages 41–49. Lawrence Erlbaum Associates, Hillsdale, NJ, 1987.

[156] D. E. Goldberg and P. Segrest. Finite Markov chain analysis of genetic algorithms. In J. J. Grefenstette, editor, *Genetic Algorithms and their Applications: Proceedings of the Second International Conference on Genetic Algorithms Juy 28-31, 1987 MIT Cambridge*, pages 1–8. Lawrence Erlbaum Associates, Hillsdale, NJ, 1987.

[157] J. Gomez. Self adaptation of operator rates in evolutionary algorithms. In K. Deb, editor, *Genetic and Evolutionary Computation - GECCO 2004 Genetic and Evolutionary Conference Seattle, WA, USA, June 26-30, 2004, Proceedings Part I*, Lecture Notes in Computer Science 3102, pages 1162–1173. Springer, 2004.

[158] P. Gopikrishnan, V. Plerou, L. A. N. Amaral, M. Meyer, , and H. E. Stanley. Scaling of the distribution of fluctuations of financial market indices. *Physical Review E*, 60(5):5305–5316, 1999.

[159] P. Gopikrishnan, V. Plerou, Y. Liu, L. A. N. Amaral, X. Gabaix, and H. E. Stanley. Scaling and correlation in financial time series. *Physica A*, 287:362–373, 2000.

[160] W. M. Gorman. Community preference fields. *Econometrica*, 21(1):63–80, 1953.

[161] H. A. J. Green. *Aggregation in Economic Analysis: An Introductory Survey.* Princeton University Press, Princeton, N.J., 1964.

[162] J. Grefenstette. Proportional selection and sampling algorithms. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C2.2. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[163] J. Grefenstette. Rank-based selection. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 2.4. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[164] S. Grossman and J. E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.

[165] D. M. Guillaume, M. M. Dacorogna, R. R. Davé, U. A. Müller, R. B. Olsen, and O. V. Pictet. From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics*, 1(2):95–129, 1997.

[166] L. Gulyás, B. Adamcsek, and A. Kiss. An early agent-based stock market: Replication and participation. In *Proceedings 4th Soft-*

Computing for Economics and Finance Meeting, May 29, 2003, University Ca' Foscari, Venice, Italy, 2003.

[167] P. J. B. Hancock. A comparison of selection mechanisms. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 2.8. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[168] P. R. Hansen and A. Lunde. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics (forthcoming)*, 2005.

[169] P. R. Hansen, A. Lunde, and J. M. Nason. Choosing the best volatility models: The model confidence set approach. *Oxford Bulletin of Economics and Statistics, Supplement*, 65(0):839–861, 2003.

[170] G. A. Harrison, J. M. Tanner, D. R. Pilbeam, and P. T. Baker. *Human Biology: An Introduction to Human Evolution, Variation, Growth, and Adaptability.* Oxford University Press, Oxford, UK, 3 edition, 1988.

[171] R. F. Harrod. Price and cost in entrepreneurs' policy. *Oxford Economic Papers*, 2(1):1–11, 1939.

[172] J. E. Hartley. *The Representative Agent in Macroeconomics.* Routledge, London, New York, 1997.

[173] F. G. Heath. Origins of the binary code. *Scientific American*, 227(2):76–83, August 1972.

[174] P. W. Hedrick. *Genetics of Populations.* Jones and Bartlett, Sudbury, MA, 2 edition, 1999.

[175] M. Heinemann. Adaptive learning of rational expectations using neural networks. *Journal of Economic Dynamics and Control*, 24(5-7):1007–1026, 2000.

[176] M. F. Hellwig. Zur Informationseffizienz des Kapitalmarktes. *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 102(1):1–27, 1982.

[177] B. Hergenhahn. *An Introduction to Theories of Learning.* Prentice Hall, Englewood Cliffs, N.J., 3 edition, 1988.

[178] I. N. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21(2):291–297, 1953.

[179] J. Hirshleifer. The expanding domain of economics. *American Economic Review*, 75(6):53–68, 1985.

[180] G. M. Hodgson. *Economics and Evolution - Bringing Life Back into Economics.* Economics, Cognition, and Society. University of Michigan Press, Ann Arbor, 1993.

[181] G. M. Hodgson. Darwinism in economics: From analogy to ontology. *Journal of Evolutionary Economics*, 12(3):259–281, 2002.

[182] J. H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* University of Michigan Press, Ann Arbor, MI, 1975.

[183] J. H. Holland. The global economy as an adaptive process. In P. W. Anderson, K. J. Arrow, and D. Pines, editors, *The Economy as an Evolving Complex System*, volume 5 of *Santa Fe Institute studies in the sciences of complexity*, pages 117–124. Westview, Santa Fe, N.M., 1988.

[184] J. H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* MIT Press / Bradford Books edition, Cambridge, MA, 1992.

[185] J. H. Holland, K. J. Holyoak, R. E. Nisbett, and P. R. Thagard. *Induction: Processes of Inference, Learning, and Discovery.* MIT Press, Cambridge, MA, 1986.

[186] J. H. Holland and J. H. Miller. Artificial adaptive agents in economic theory. *American Economic Review, Papers and Proceedings*, 81(2):365–370, 1991.

[187] J. H. Holland and J. S. Reitman. Cognitive systems based on adaptive algorithms. In D. A. Waterman and F. Hayes-Roth, editors, *Pattern-Directed Inference Systems*, pages 313–329. Academic Press, New York, 1978.

[188] R. B. Hollstien. *Artificial Genetic Adaptation in Computer Control Systems.* Phd thesis, University of Michigan, 1971.

[189] C. H. Hommes. Financial markets as nonlinear adaptive evolutionary systems. *Quantitative Finance*, 1(1):149–167, 2001.

[190] C. H. Hommes. Heterogeneous agent models in economics and finance. In K. L. Judd and L. Tesfatsion, editors, *Handbook of Computational Economics*, volume 2. Elsevier, Amsterdam, 2006.

[191] C. H. Hommes and G. Sorger. Consistent expectations equilibria. *Macroeconomic Dynamics*, 2(3):287–321, 1998.

[192] J. Hong and J. Stein. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54(6):2143–2184, 1999.

[193] J. Horgan. From complexity to perplexity. *Scientific American*, 272(6):104–110, 1995.

[194] G. Iori. A threshold model for stock return volatility and trading volume. *International Journal of Theoretical and Applied Finance*, 3(3):467–472, 2000.

[195] G. Iori. A microsimulation of traders activity in the stock market: The role of heterogeneity, agents' interactions and trade frictions. *Journal of Economic Behavior and Organization*, 49(2):269–285, 2002.

[196] J. F. Jaffe. Special information and insider trading. *Journal of Business*, 47(410):410–428, 1974.

[197] C. M. Jarque and A. K. Bera. Efficient tests for normality, heteroscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259, 1980.

[198] T. Jasic and D. Wood. The profitability of daily stock market indices trades based on neural network predictions: Case study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965-1999. *Applied Financial Economics*, 14(4):285–297, 2004.

[199] N. Jegadeesh. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation: Discussion. *Journal of Finance*, 55(4):1765–1770, 2000.

[200] N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91, 1993.

[201] M. C. Jensen. Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2-3):95–101, 1978.

[202] M. C. Jensen and G. A. Benington. Random walks and technical theories: Some additional evidence. *Journal of Finance*, 25(2):469–482, 1970.

[203] M. C. Jerison. Aggregation and pairwise aggregation of demand when the distribution of income is fixed. *Journal of Economic Theory*, 33(1):1–33, June 1984.

[204] P. E. Johnson. Agent-based modeling: What I learned from the artificial stock market. *Social Science Computer Review*, 20(2):174–186, 2002.

[205] S. Joshi and M. A. Bedau. An explanation of generic behavior in an evolving financial market. *Santa Fe Institute Working Paper 98-12-114*, 1998.

[206] S. Joshi, J. Parker, and M. A. Bedau. Technical trading creates a prisoner's dilemma: results from an agent-based model. *Santa Fe Institute Working Paper 98-12-115*, 1998.

[207] S. Joshi, J. Parker, and M. A. Bedau. Financial markets can be at sub-optimal equilibria. *Computational Economics*, 19(1):5–23, 2002.

[208] K. L. Judd. Computational economics and economic theory: Substitutes or complements? *Journal of Economic Dynamics and Control*, 21(6):907–942, 1997.

[209] K. L. Judd and L. Tesfatsion, editors. *Handbook of Computational Economics*, volume 2. Elsevier, Amsterdam, 2006.

[210] J. Jung and R. J. Shiller. Samuelson's dictum and the stock market. *Economic Inquiry*, 43(2):221–228, 2005.

[211] M. A. Kaboudan. Genetic programming prediction of stock prices. *Computational Economics*, 16(3):207–236, 2000.

[212] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[213] J. M. Karpoff. A theory of trading volume. *Journal of Finance*, 41(5):1069–1087, 1986.

[214] J. M. Karpoff. The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22(1):109–126, 1987.

[215] S. Kauffman. *At Home in the Universe: The Search for Laws of Self-Organization and Complexity.* Oxford University Press, New York, Oxford, 1995.

[216] M. Kendall. The analysis of economic time series, part i: Prices. *Journal of the Royal Statistical Society, Series A*, 96:11–25, 1953.

[217] D. A. Kendrick. Sectoral economics. In H. M. Amman, D. A. Kendrick, and J. Rust, editors, *Handbook of Computational Economics*, chapter 6, pages 295–332. Elsevier, Amsterdam, 1996.

[218] J. M. Keynes. *The General Theory of Employment, Interest, and Money.* MacMillan, New York, reprinted 1964 San Diego, Harcourt Brace, 1936.

[219] G. A. Kimble. The definitions of learning and some useful distinctions. In G. A. Kimble, editor, *Foundations of Conditioning and Learning*, The Century Psychology Series, pages 82–99. Appleton-Century-Crofts, New York, 1967.

[220] M. Kimura. Solution of a process of random genetic drift with a continuous model. *Proceedings of the National Academy of Science of the USA*, 41(3):144–150, March 1955.

[221] M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713–719, 1962.

[222] M. Kimura. Evolutionary rate at the molecular level. *Nature*, 217:624–626, 1968.

[223] M. Kimura. *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge, UK, 1983.

[224] M. Kimura and T. Ohta. *Theoretical Aspects of Population Genetics.* Princeton University Press, Princeton, NJ, 1971.

[225] A. P. Kirman. Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136, Spring 1992.

[226] N. Kiyotaki and R. Wright. On money as a medium of exchange. *Journal of Political Economy*, 97(4):927–954, 1989.

[227] T. S. Kuhn. *The Structure of Scientific Revolutions.* University of Chicago Press, Chicago, London, 3 edition, 1996.

[228] K. Y. Kwon and R. J. Kish. A comparative study of technical trading strategies and return predictability: An extension of Brock, Lakonishok, and LeBaron (1992) using NYSE and NASDAQ indices. *Quarterly Review of Economics and Finance*, 42(3):611–631, 2002.

[229] K. Y. Kwon and R. J. Kish. Technical trading strategies and return predictability: NYSE. *Applied Financial Economics*, 12(9):639–653, 2002.

[230] A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, 1985.

[231] R. N. Langlois. Bounded rationality and behavioralism: A clarification and critique. *Journal of Institutional and Theoretical Economics*, 146(4):691–695, 1990.

[232] R. N. Langlois. Rule-following, expertise, and rationality: a new behavioral economics? In K. Dennis, editor, *Rationality in Economics: Alternative Perspectives*, pages 57–80. Kluwer, Dordrecht, 1998.

[233] R. N. Langlois and L. Csontos. Optimization, rule following, and the methodology of situational analysis. In B. Gustafsson, C. Knudsen, and U. Mäki, editors, *Rationality, Institutions, and Economic Methodology.* Routledge, Nondon, New York, 1993.

[234] J. S. Lansing, J. N. Kremer, and B. B. Smuts. System-dependent selection, ecological feedback and the emergence of functional structure in ecosystems. *Journal of Theoretical Biology*, 192(3):377–391, 1998.

[235] T. Lawson. *Economics and Reality.* Routledge, London, New York, 1997.

[236] B. LeBaron. Building financial markets with artificial agents: Desired goals, and present techniques. In G. Karakoulas, editor, *Computational Markets.* MIT Press, 1999.

[237] B. LeBaron. Technical trading rule profitability and foreign exchange intervention. *Journal of International Economics*, 49(1):125–143, 1999.

[238] B. LeBaron. Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, 24(5-7):679–702, 2000.

[239] B. LeBaron. A builder's guide to agent based financial markets. *Quantitative Finance*, 1(2):254–261, March 2001.

[240] B. LeBaron. Empirical regularities from interacting long and short memory investors in an agent based stock market. *IEEE Transactions on Evolutionary Computation*, 5(5):442–455, 2001.

[241] B. LeBaron. Evolution and time horizons in an agent-based stock market. *Macroeconomic Dynamics*, 5(2):225–254, 2001.

[242] B. LeBaron. Building the Santa Fe Artificial Stock Market. working paper, Brandeis University, 2002.

[243] B. LeBaron. Agent-based computational finance. In K. L. Judd and L. Tesfatsion, editors, *The Handbook of Computational Economics*, volume 2, chapter 9. Elsevier, Amsterdam, 2006.

[244] B. LeBaron, W. B. Arthur, and R. Palmer. Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, 23(9-10):1487–1516, 1999.

[245] S. F. LeRoy. Efficient capital markets and martingales. *Journal of Economic Literature*, 27(4):1583–1621, 1989.

[246] M. Lettau. Explaining the facts with adaptive agents: The case of mutual fund flows. *Journal of Economic Dynamics and Control*, 21(7):1117–1147, 1997.

[247] M. Levy, H. Levy, and S. Solomon. A microscopic model of the stock market: Cycles, booms, and crashes. *Economics Letters*, 45(1):103–111, 1994.

[248] M. Levy, H. Levy, and S. Solomon. *Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena*. Academic Press, San Diego et al., 2000.

[249] A. Lewbel. Exact aggregation and a representative consumer. *Quarterly Journal of Economics*, 104(3):621–633, 1989.

[250] R. C. Lewontin. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York, 1974.

[251] W. K. Li, S. Ling, and M. McAleer. Recent theoretical results for time series models with GARCH errors. *Journal of Economic Surveys*, 16(3):245–269, 2002.

[252] J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, 1965.

[253] B. L. Lipman. How to decide how to decide how to ... : Modeling bounded rationality. *Econometrica*, 59(4):1105–1125, 1991.

[254] R. G. Lipsey. The relation between unemployment and the rate of change of money wages in the United Kingdom 1862-1957: A further analysis. *Economica*, 27(105):1–31, 1960.

[255] A. W. Lo. The Adaptive Markets Hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*, 30:15–29, 2004.

[256] A. W. Lo, H. Mamaysky, and J. Wang. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *Journal of Finance*, 55(4):1705–1765, 2000.

[257] I. N. Lobato and C. Velasco. Long memory in stock-market trading volume. *Journal of Business and Economic Statistics*, 18(4):410–426, 2000.

[258] G. Loomes and C. Taylor. Non-transitive preferences over gains and losses. *Economic Journal*, 102(411):357–365, 1992.

[259] M. Loretan and P. C. B. Phillips. Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance*, 1(2):211–248, 1994.

[260] J. Lucas, Robert E. Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–124, April 1972.

[261] J. Lucas, Robert E. Econometric policy evaluation: A critique. In K. Brunner and A. H. Meltzer, editors, *The Phillips Curve and Labor Markets*, volume 1 of *Carnegie-Rochester Conference Series on Public Policy*, pages 19–46. North-Holland, Amsterdam, 1976.

[262] J. Lucas, Robert E. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445, 1978.

[263] Y.-H. Lui and D. Mole. The use of fundamental and technical analyses by foreign exchange dealers: Honk Kong evidence. *International Journal of Finance and Economics*, 17(3):535–545, 1998.

[264] T. Lux. Herd behaviour, bubbles, and crashes. *Economic Journal*, 105(431):881–896, 1995.

[265] T. Lux. The stable Paretian hypothesis and the frequency of large returns: An examination of major German stocks. *Applied Financial Economics*, 6(6):463–475, 1996.

[266] T. Lux. The socio-economic dynamics of speculative markets: Interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior and Organization*, 33(2):143–165, 1998.

[267] T. Lux and M. Ausloos. Market fluctuations I: Scaling, multiscaling and their possible origins. In A. Bunde, J. Kropp, and H.-J. Schellnhuber, editors, *The Science of Desaster: Climate Disruptions, Heart Attacks, and Market Crashes*, pages 373–409. Springer, 2002.

[268] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397:498–500, 1999.

[269] T. Lux and M. Marchesi. Volatility clustering in financial markets: a microsimulationof interacting agents. *International Journal of Theoretical and Applied Finance*, 3(4):675–702, 2000.

[270] S. W. Mahfoud. Population size and genetic drift in fitness sharing. In L. D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, volume 3, pages 185–223. Morgan Kaufmann, San Francisco, CA, 1995.

[271] S. W. Mahfoud. Boltzmann selection. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C 2.5. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[272] S. W. Mahfoud. Niching methods. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C6.1. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[273] U. Mäki. 'The Methodology of Positive Economics' (1953) does not give us the methodology of positive economics. *Journal of Economic Methodology*, 10(4):495–505, 2003.

[274] B. G. Malkiel. *A Random Walk Down Wall Street : Including a Life-Cycle Guide to Personal Investing*. Norton, New York, 1999.

[275] B. G. Malkiel. The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1):59–82, 2003.

[276] B. B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36(4):394–419, 1963.

[277] B. B. Mandelbrot. Forecasts of future prices, unbiased markets, and "martingale" models. *Journal of Business*, 39(1, Part 2: Supplement on Security Pricing):242–255, 1966.

[278] B. B. Mandelbrot. Scaling in financial prices: 3: Cartoon Brownian motions in multifractal time. *Quantitative Finance*, 1(4):427–440, 2001.

[279] B. B. Mandelbrot, A. Fisher, and L. Calvet. A multifractal model of asset returns. *Cowles Foundation Discussion Paper #1164*, 1997.

[280] B. B. Mandelbrot and R. L. Hudson. *The (Mis)behavior of Markets: A Fractal View of Risk, Ruin, and Reward.* Basic Books, New York, 2004.

[281] B. B. Mandelbrot and J. R. Wallis. Noah, Joseph, and operational hydrology. *Water Resources Research*, 4(3):909–918, 1968.

[282] R. N. Mantegna and H. E. Stanley. Scaling behavior in the behavior of an economic index. *Nature*, 376:46–49, 1995.

[283] R. N. Mantegna and H. E. Stanley. *An Introduction to Econophysics: Correlations and Complexity in Finance.* Cambridge University Press, Cambridge, UK, 2000.

[284] R. Marimon. Adaptive learning, evolutionary dynamics, and equilibrium selection in games. *European Economic Review*, 37(2-3):603–611, 1993.

[285] R. Marimon, E. McGrattan, and T. Sargent. Money as a medium of exchange in an economy with artificially intelligent agents. *Journal of Economic Dynamics and Control*, 14(2):329–373, 1990.

[286] S. Markose. Computability and evolutionary complexity: Markets as Complex Adaptive Systems (CAS). *Economic Journal, forthcoming*, 2005.

[287] S. Markose, E. Tsang, and S. Martinez-Jaramillo. The Red Queen principle and the emergence of efficient financial markets: An agent based approach. In T. Lux, editor, *Proceedings of the 8th Annual Workshop on Economics with Heterogeneous Interacting Agents (Kiel, Germany, May 29-31, 2003)*, Heidelberg, New York, 2004.

[288] H. M. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, March 1952.

[289] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments.* Wiley, New York, 1959.

[290] A. Marshall. *Principles of Economics.* Macmillan & Co., Ltd., London, 1890.

[291] T. Mayer. *Truth versus Precision in Economics.* Edward Elgar, Aldershot, 1993.

[292] D. McFadden. Rationality for economists? *Journal of Risk and Uncertainty*, 19(1-3):73–105, December 1999.

[293] G. McQueen and K. P. Vorkink. Whence GARCH? A preference-based explanation for conditional volatility. *Review of Financial Studies*, 17(4):915–949, 2004.

[294] D. H. Meadows, D. L. Meadows, and J. Randers. *Beyond the Limits.* Chelsea Green, Post Mills, VT, 1992.

[295] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens III. *The Limits to Growth.* Universe Books, New York, 1972.

[296] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens III. *The Dynamics of Growth in a Finite World.* Wright-Allen Press, Cambridge, MA, 1974.

[297] D. H. Meadows, J. Randers, and D. L. Meadows. *Limits to Growth: The 30-Year Update.* Chelsea Green, White River Jct., VT, 2004.

[298] L. Menkhoff. Examining the use of technical currency analysis. *International Journal of Finance and Economics*, 2(4):307–318, 1997.

[299] L. Menkhoff. The noise trading approach–questionnaire evidence from foreign exchange. *Journal of International Money and Finance*, 17(3):547–564, 1998.

[300] M. Metghalchi and Y. H. Chang. Profitable technical trading rules for the Italian stock market. *RISEC – International Review of Economics and Business*, 50(4):433–450, 2003.

[301] D. Michie, editor. *Expert Systems in the Micro-electronic Age.* Edinburgh University Press, Edinburgh, 1979.

[302] T. Mikosch and C. Stărică. Is it really long memory we see in financial returns? In P. Embrechts, editor, *Extremes and Integrated Risk Management*, chapter 12, pages 149–168. Risk Books, London, 2000.

[303] T. Mikosch and C. Stărică. Long range dependence effects and ARCH modelling. In P. Doukhan, G. Oppenheim, and M. Taqqu, editors, *Theory and Applications of Long Range Dependence*, pages 439–460. Birkhauser, Boston, 2003.

[304] T. Mikosch and C. Stărică. Changes of structure in financial time series and the GARCH model. *REVSTAT Statistical Journal*, 2(1):41–73, 2004.

[305] T. Mikosch and C. Stărică. Non-stationarities in financial time series, the long range dependence and the IGARCH effects. *Review of Economics and Statistics*, 86(1):378–390, 2004.

[306] P. Milgrom and N. Stokey. Information, trade, and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

[307] P. E. Mirowski. Review of: Fractals and scaling in finance: Discontinuity, concentration, risk. *Journal of Economic Literature*, 39(2):585–587, 2001.

[308] M. Mitchell. *An Introduction to Genetic Algorithms.* MIT Press, Cambridge, MA, 1996.

[309] S. K. Mitra. Profiting from technical analysis in Indian stock market. *Finance India*, 16(1):109–120, 2002.

[310] P. Mongin. The marginalist controversy. In J. Davis, W. Hands, and U. Maki, editors, *Handbook of Economic Methodology*, pages 558–562. Edward Elgar, London, 1997.

[311] R. D. Montassér. *Die Rolle des Volumens bei der Aktienkursprognose unter besonderer Berücksichtigung der AVAS-Transformation.* Gabler Edition Wissenschaft: Hallesche Schriften zur Betriebswirtschaft 12. Deutscher Universitätsverlag, Wiesbaden, 2003.

[312] S. Moss and B. Edmonds. Modeling economic learning as modeling. *Cybernetics and Systems*, 29(3):215–247, 1998.

[313] J. Mossin. Equilibrium in a capital asset market. *Econometrica*, 34(4):768–783, 1966.

[314] R. F. Mulligan. Fractal analysis of highly volatile markets: An application to technology equities. *Quarterly Review of Economics and Finance*, 44(1):155–179, 2004.

[315] J. F. Muth. Rational expectations and the theory of price movements. *Econometrica*, 29(3):315–335, 1961.

[316] A. Nagurney. Parallel computation. In H. M. Amman, D. A. Kendrick, and J. Rust, editors, *Handbook in Computational Economics*, chapter 7, pages 335–404. Elsevier, Amsterdam, 1996.

[317] C. J. Neely. The temporal pattern of trading rule returns and exchange rate intervention: Intervention does not generate technical trading profits. *Journal of International Economics*, 58(1):211–232, 2002.

[318] D. B. Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991.

[319] R. R. Nelson and S. G. Winter. *An Evolutionary Theory of Economic Change.* MIT Press, Cambridge, Mass., 1982.

[320] S. A. Nelson. *The ABC of Stock Speculation.* Fraser Publ. Co., New York, 1903.

[321] M. Nerlove. Adaptive expectations and cobweb phenomena. *Quarterly Journal of Economics*, 72(2):227–240, 1958.

[322] G. Nicolis and I. Prigogine. *Exploring Complexity : An Introduction.* Freeman, New York, 1989.

[323] W. D. Nordhaus. World dynamics: Measurement without data. *Economic Journal*, 83(332), 1973.

[324] W. D. Nordhaus. Lethal model 2: The limits to growth revisited. *Brookings Paper on Economic Activity*, 2, 1992.

[325] T. Oberlechner. Importance of technical and fundamental analysis in the European foreign exchange market. *International Journal of Finance and Economics*, 6(1):81–93, 2001.

[326] T. Ohta. Slightly deleterious mutant substitutions in evolution. *Nature*, 246:96–98, 1973.

[327] T. Ohta. Near-neutrality in evolution of genes and gene regulation. *Proceedings of the National Academy of Science of the USA*, 99(25):16134–16137, December 2002.

[328] A. Pagan. The econometrics of financial markets. *Journal of Empirical Finance*, 3(1):15–102, 1996.

[329] S. E. Page. Computational models from A to Z. *Complexity*, 5(1):35–41, 1999.

[330] R. G. Palmer, W. B. Arthur, J. H. Holland, and B. LeBaron. An artificial stock market. *Artificial Life and Robotics*, 3(1):27–31, 1999.

[331] R. G. Palmer, W. B. Arthur, B. LeBaron, and P. Tayler. Artificial economic life: A simple model of a stockmarket. *Physica D*, 75:264–274, 1994.

[332] R. Pancs and N. J. Vriend. Schelling's spatial proximity model of segregation revisited. Queen Mary University of London, Dept. of Economics Working Paper 487, 2003.

[333] K. Pearson. The problem of the random walk. *Nature*, 72:294, July 1905.

[334] M. H. Pesaran. *The Limits to Rational Expectations*. Basil Blackwell, Oxford, New York, 1987.

[335] E. E. Peters. *Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility*. Wiley, New York, 1991.

[336] E. E. Peters. *Fractal market Analysis: Applying Chaos Theory to Investment and Economics*. Wiley, New York, 1994.

[337] J. Pheby. *Methodology and Economics: A Critical Introduction*. MacMillan Press, Houndsmill et al., 1988.

[338] E. S. Phelps. Phillips curves, expectations of inflation, and optimal unemployment over time. *Economica*, 34(135):254–281, 1967.

[339] A. W. Phillips. The relation between unemployment and the rate of change of money wages in the United Kingdom 1861-1957. *Economica*, 25(100):283–299, 1958.

[340] K. R. Popper. *The Open Society and Its Enemies*, volume 2. Princeton University Press, Princeton, 2 edition, 1966.

[341] K. R. Popper. *Objektive Erkenntnis: Ein evolutionärer Entwurf.* Hoffman and Campe, Hamburg, 3 edition, 1995.

[342] R. R. Prechter. *The Major Works of R.N. Elliott.* New Classics Library, New York, 1980.

[343] W. H. Press, S. A. Teukolski, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C. The Art of Scientific Computing.* Cambridge University Press, Cambridge, UK, 2 edition, 1992.

[344] M. Raberto, S. Cincotti, S. M. Focardi, and M. Marchesi. Agent-based simulation of a financial market. *Physica A*, 299:319–327, 2001.

[345] M. Raberto, S. Cincotti, S. M. Focardi, and M. Marchesi. Traders' long-run wealth in an artificial financial market. *Computational Economics*, 22(2-3):255–272, 2003.

[346] S. Rachev and S. Mittnik. *Stable Paretian Models in Finance.* Wiley, Chichester, UK, 2000.

[347] N. J. Radcliffe. Schema processing. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter B2.5. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[348] M. Ratner and R.-P. C. Leal. Tests of technical trading strategies in the emerging equity markets of Latin America and Asia. *Journal of Banking and Finance*, 23(12):1887–1905, 1999.

[349] J. W. S. Rayleigh. The problem of the random walk. *Nature*, 72:318, August 1905.

[350] M. J. Ready. Profits from technical trading rules. *Financial Management*, 31(3):43–61, 2002.

[351] I. Rechenberg. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der technischen Evolution.* Frommann-Holzboog, Stuttgart, 1973.

[352] I. Rechenberg. *Evolutionsstrategie '94.* Frommann-Holzboog, Stuttgart, 1994.

[353] M. W. Reder. Remarks on the 'The methodology of positive economics'. *Journal of Economic Methodology*, 10(4):527–530, 2003.

[354] T. Riechmann. Learning and behavioral stability: An economic interpretation of genetic algithms. *Journal of Evolutionary Economics*, 9(2):225–242, 1999.

[355] T. Riechmann. Genetic algorithm learning and evolutionary games. *Journal of Economic Dynamics and Control*, 25(6-7):1019–1037, 2001.

[356] T. Riechmann. *Learning in Economics: Analysis and Application of Genetic Algorithms*. Physica, Heidelberg, 2001.

[357] A. J. Robson. The evolution of rationality and the Red Queen. *Journal of Economic Theory*, 111(1):1–22, 2003.

[358] A. Rogers and A. Prügel-Bennett. Genetic drift in genetic algorithm selection schemes. *IEEE Transactions on Evolutionary Computation*, 3(4):298–303, November 1999.

[359] R. Roll. What every ceo should know about scientific progress in economics: What is known and what remains to be resolved. *Financial Management*, 23(3):69–75, 1994.

[360] S. A. Ross. Finance. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave: A Dictionary of Economics*. W. W. Norton, New York, London, 1987.

[361] J. B. Rosser Jr. On the complexities of complex economic dynamics. *Journal of Economic Perspectives*, 13(4):169–192, Fall 1999.

[362] M. Rubinstein. Rational markets: Yes or no? The affirmative case. *Financial Analysts Journal*, 57(3):15–29, May-Jun 2001.

[363] P. Saacke. Technical analysis and the effectiveness of central bank intervention. *International Journal of Money and Finance*, 21(4):459–479, 2002.

[364] L. Samuelson. *Evolutionary Games and Equilibirum Selection*. MIT Press Series on Ecvonomic Learning and Social Evolution. MIT Press, Cambridge, MA, 1997.

[365] P. A. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(1):41–50, 1965.

[366] P. A. Samuelson. Summing up on business cycles: Opening address. In J. C. Fuhrer and S. Schuh, editors, *Beyond Shocks: What Causes Business Cycles?* Federal Reserve Bank of Boston, Boston, 1998.

[367] P. A. Samuelson and R. M. Solow. The problem of achieving and maintaining a stable price level: Analytical aspects of anti-inflation policy. *American Economic Review*, 50(2):177–194, 1960.

[368] A. M. Santomero and J. J. Seater. The inflation-unemployment trade-off: A critique of the literature. *Journal of Economic Literature*, 16(2):499–544, June 1978.

[369] A. Sapusek. *Informationseffizienz auf Kapitalmärkten: Konzepte und empirische Ergebnisse.* Gabler, Wiesbaden, 1998.

[370] T. J. Sargent. *Macroeconomic Theory.* Academic Press, New York, 1979.

[371] T. J. Sargent. Beyond demand and supply curves in macroeconomics. *American Economic Review, Papers and Proceedings*, 72(2):382–389, May 1982.

[372] T. J. Sargent. *Bounded Rationality in Macroeconomics.* Clarendon Press, Oxford, 1993.

[373] J. Sarma and K. D. Jong. Generation gap methods. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter C2.7. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[374] L. J. Savage. *The Foundations of Statistics.* Wiley, New York, 1954.

[375] T. C. Schelling. Models of segregation. *American Economic Review, Papers and Proceedings*, 59(2):488–493, 1969.

[376] T. C. Schelling. *Micromotives and Macrobehavior.* W.W. Norton & Co., New York, London, 1978.

[377] K. H. Schlag. Why imitate, and if so, how? a boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130–156, 1998.

[378] K. H. Schlag. Which one should I imitate? *Journal of Mathematical Economics*, 31(4):493–522, 1999.

[379] J. Schmidt. *Die Grenzen der Rational Choice Theorie: Eine kritische theoretische und empirische Studie.* Leske + Budrich, Opladen, 2000.

[380] R. Schmidt and S. Wulff. Zur Entdeckung von Insider-Aktivitäten am deutschen Aktienmarkt. *Zeitschrift für Bankrecht und Bankwirtschaft*, 5(1):57–68, 1993.

[381] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley, New York, 2003.

[382] J. A. Schumpeter. *The Theory of Economic Development.* Harvard University Press, 1934.

[383] P. Schuster and K. Sigmund. Replicator dynamics. *Journal of Theoretical Biology*, 100(3):533–538, 1983.

[384] F. Schweitzer. Modelling migration and economic agglomeration with active Brownian particles. *Advances in Complex Systems*, 1(1):11–37, 1998.

[385] F. Schweitzer. *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*. Springer Series in Synergetics. Springer, Berlin, 2003.

[386] F. Schweitzer and J. A. Holyst. Modelling collective opinion formation by means of active Brownian particles. *European Physical Journal B*, 15(4):723–732, 2000.

[387] S. Sehgal and A. Garhyan. Abnormal returns using technical analysis: The Indian experience. *Finance India*, 16(1):181–203, 2002.

[388] R. Selten. What is bounded rationality? In G. Gigerenzer and R. Selten, editors, *Bounded Rationality: The Adaptive Toolbox*, chapter 2, pages 13–36. MIT Press, Cambridge, MA, 2001.

[389] W. F. Sharpe. Capital asset prices: A theory of equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

[390] R. J. Shiller. Investor behavior in the October 1987 stock market crash: Survey evidence. *NBER working paper 2446*, 1987.

[391] R. J. Shiller. From efficient market theory to behavioral finance. *Cowles Foundation Discussion Paper 1385*, 2002.

[392] A. Shleifer. *Inefficient Markets : An Introduction to Behavioral Finance*. Oxford University Press, Oxford, New York, 2000.

[393] A. Shleifer and L. H. Summers. The noise trader approach to finance. *Journal of Economic Perspectives*, 4(2):19–33, 1990.

[394] H. A. Simon. *Models of Man*. Wiley, New York, 1957.

[395] H. A. Simon. From substantive to procedural rationality. In S. J. Latsis, editor, *Method and Appraisal in Economics*, pages 129–148. Cambridge University Press, Cambridge, New York, 1976.

[396] H. A. Simon. Rationality as a process and as product of thought. *American Economic Review*, 68(2):1–16, 1978.

[397] H. A. Simon. Rational decision making in business organizations. *American Economic Review*, 69(4):493–513, 1979.

[398] H. A. Simon. Complex systems: The interplay of organizations and markets in contemporary society. *Computational and Mathematical Organization Theory*, 7(2):79–85, 2001.

[399] B. F. Skinner. Are theories of learning necessary? *Psychological Review*, 57(4):193–216, 1950.

[400] T. Slembeck. Learning in economics: Where do we stand? a behavioral view on learning in theory, practice and experiments. *University of St. Gallen, VWA Discussion Paper No. 9907*, 1999.

[401] T. Slembeck. Evolution und Lernen im Vergleich: Individuelles Lernen als Voraussetzung für Selbsttransformation und die Bedeutung von Lernbedingungen. Paper delivered at the Annual

Meeting of the Committee for Evolutionary Economics of the Verein für Socialpolitik, Schloss Reisensburg, July 2000, 2000.

[402] R. E. Smith. Learning classifier systems. In T. Bäck, D. B. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, chapter B1.5.2. IOP Publishing Ltd and Oxford University Press, Oxford, UK, 1997.

[403] S. F. Smith. *A learning system based on genetic adaptive algorithms*. Phd thesis, University of Pittsburgh, Pittsburgh, PA, 1980.

[404] V. L. Smith. An experimental study of competitive market behavior. *Journal of Political Economy*, 70(3):111–137, 1962.

[405] J. Sobel. Economists' model of learning. *Journal of Economic Theory*, 94(2):241–261, 2000.

[406] L. Sögner and H. Mitlöhner. Consistent expectations equilibria and learning in a stock market. *Journal of Economic Dynamics and Control*, 26(2):171–185, 2002.

[407] D. Stauffer, P. M. C. de Oliveira, and A. T. Bernardes. Monte Carlo simulation of volatility clustering in market model with herding. *International Journal of Theoretical and Applied Finance*, 2(1):83–94, 1999.

[408] D. Stauffer and T. J. P. P. Penna. Crossover in the Cont-Bouchaud percolation model for market fluctuations. *Physica A*, 256:284–290, 1998.

[409] D. Stauffer and D. Sornette. Self-organized percolation model for stock market fluctuations,. *Physica A*, 271:496–506, 1999.

[410] M. Steiner and H.-G. Wittkemper. Portfolio optimization with a neural network implementation of the coherent market hypothesis. *European Journal of Operational Research*, 100(1):27–40, 1997.

[411] J. E. Stiglitz. The allocation role of the stock market. *Journal of Finance*, 36(2):235–251, 1981.

[412] T. M. Stoker. Exact aggregation and generalized slutsky conditions. *Journal of Economic Theory*, 33:368–377, 1984.

[413] S. Sunder. Markets as artifacts: Aggregate efficiency from zero-intelligence traders. *Yale ICF Working Paper No. 02-16*, May 2003.

[414] D. T. Suzuki, A. J. F. Griffiths, J. H. Miller, and R. C. Lewontin. *An Introduction to Genetic Analysis*. Freeman, New York, 4 edition, 1989.

[415] G. Syswerda. Uniform crossover in genetic algorithms. In J. D. Schaffer, editor, *Proc. 3rd Int. Conf. on Genetic Algorithms*, pages 2–9. Morgan Kaufmann, San Mateo, CA, 1989.

[416] M. S. Taqqu. Bachelier and his times: A conversation with Bernard Bru. *Finance and Stochastics*, 5(1):3–32, 2001.

[417] G. E. Tauchen and M. Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505, 1983.

[418] N. S. P. Tay and S. C. Linn. Fuzzy inductive reasoning, expectation formation and the behavior of security prices. *Journal of Economic Dynamics and Control*, 25(3-4):321–361, 2001.

[419] M. P. Taylor and H. L. Allen. The use of technical analysis in the foreign exchange market. *International Journal of Money and Finance*, 11(3):304–314, 1992.

[420] L. Tesfatsion. Economic agents and markets as emergent phenomena. *Proceedings of the National Academy of Science of the USA*, 99:7199–7200, 2002.

[421] L. S. Tesfatsion. Agent-based computational economics: Modelling economies as complex adaptive systems. *Information Sciences*, 149(4):263–269, 2003.

[422] J. Tirole. On the possibility of speculation under rational expectations. *Econometrica*, 50(5):1163–1182, 1982.

[423] R. Tobias and C. Hofman. Evaluation of free java-libraries for social-scientific agent based simulation. *Journal of Artificial Societies and Social Simulation*, 7(1), 2004.

[424] H. Tong and K. S. Lim. Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society, Series B (Methodological)*, 42(3):245–292, 1980.

[425] R. M. Townsend. Forecasting the forecasts of others. *Journal of Political Economy*, 91(4):546–588, 1983.

[426] T. Vaga. The coherent market hypothesis. *Financial Analysts Journal*, 46(6):36–49, 1990.

[427] L. van Valen. A new evolutionary law. *Evolutionary Theory*, 1(1):1–30, 1973.

[428] H. R. Varian. Differences of opinion in financial markets. In C. C. Stone, editor, *Financial Risk: Theory, Evidence and Implications*, Proceedings of the 11th Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis, pages 3–37, Boston, 1988. Kluwer.

[429] T. Veblen. Why is economics not an evolutionary science? *Quarterly Journal of Economics*, 12(4):373–397, 1898.

[430] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

[431] N. J. Vriend. An illustration of the essential difference between individual and social learning, and its consequences for computational analyses. *Journal of Economic Dynamics and Control*, 24(1):1–19, 2000.

[432] M. Wade. Evolutionary genetics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/evolutionary-genetics/, 2005.

[433] M. M. Waldrop. *Complexity: The Emerging Science at the Edge of Order and Chaos*. Touchstone, New York, 1992.

[434] D. A. Waterman. *A Guide to Expert Systems*. Addison-Wesley, Reading, MA, 1986.

[435] J. W. Weibull. *Evolutionary Game Theory*. MIT Press, Cambridge, London, 1996.

[436] U. Wilensky. Netlogo segregation model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL., 1998.

[437] M. Wilpert. *Künstliche Aktienmarktmodelle auf Basis von Classifier-Systems*. Knapp, Frankfurt/M., 2004.

[438] S. Winter. Economic 'natural selection' and the theory of the firm. *Yale Economic Essays*, 4:225–272, 1964.

[439] U. Witt, editor. *Evolutionary Economics*, volume 25 of *International Library of Critical Writings in Economics*. Edward Elgar, Aldershot, 1993.

[440] U. Witt. What evolutionary economics is all about? In U. Witt, editor, *Evolutionary Economics*, volume 25 of *International Library of Critical Writings in Economics*, pages xiii–xxvii. Edward Elgar, Aldershot, 1993.

[441] W. K. Wong, M. Manzur, and B. K. Chew. How rewarding is technical analysis? Evidence from Singapore stock market. *Applied Financial Economics*, 13(7):543–551, 2003.

[442] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proc. 6th Int. Congress on Genetics*, volume 1, pages 356–366, 1932.

[443] J. Yao, C. L. Tan, and H. L. Poh. Neural networks for technical analysis: A study on KLCI. *International Journal of Theoretical and Applied Finance*, 2(2):221–241, 1999.

[444] C. E. Zeeman. On the unstable behavior of the stock exchanges. *Journal of Mathematical Economics*, 1(1):39–44, 1974.

Index

Lecture Notes in Economics and Mathematical Systems

For information about Vols. 1–512
please contact your bookseller or Springer-Verlag

Vol. 513: K. Marti, Stochastic Optimization Techniques. VIII, 364 pages. 2002.

Vol. 514: S. Wang, Y. Xia, Portfolio and Asset Pricing. XII, 200 pages. 2002.

Vol. 515: G. Heisig, Planning Stability in Material Requirements Planning System. XII, 264 pages. 2002.

Vol. 516: B. Schmid, Pricing Credit Linked Financial Instruments. X, 246 pages. 2002.

Vol. 517: H. I. Meinhardt, Cooperative Decision Making in Common Pool Situations. VIII, 205 pages. 2002.

Vol. 518: S. Napel, Bilateral Bargaining. VIII, 188 pages. 2002.

Vol. 519: A. Klose, G. Speranza, L. N. Van Wassenhove (Eds.), Quantitative Approaches to Distribution Logistics and Supply Chain Management. XIII, 421 pages. 2002.

Vol. 520: B. Glaser, Efficiency versus Sustainability in Dynamic Decision Making. IX, 252 pages. 2002.

Vol. 521: R. Cowan, N. Jonard (Eds.), Heterogenous Agents, Interactions and Economic Performance. XIV, 339 pages. 2003.

Vol. 522: C. Neff, Corporate Finance, Innovation, and Strategic Competition. IX, 218 pages. 2003.

Vol. 523: W.-B. Zhang, A Theory of Interregional Dynamics. XI, 231 pages. 2003.

Vol. 524: M. Frölich, Programme Evaluation and Treatment Choise. VIII, 191 pages. 2003.

Vol. 525: S. Spinler, Capacity Reservation for Capital-Intensive Technologies. XVI, 139 pages. 2003.

Vol. 526: C. F. Daganzo, A Theory of Supply Chains. VIII, 123 pages. 2003.

Vol. 527: C. E. Metz, Information Dissemination in Currency Crises. XI, 231 pages. 2003.

Vol. 528: R. Stolletz, Performance Analysis and Optimization of Inbound Call Centers. X, 219 pages. 2003.

Vol. 529: W. Krabs, S. W. Pickl, Analysis, Controllability and Optimization of Time-Discrete Systems and Dynamical Games. XII, 187 pages. 2003.

Vol. 530: R. Wapler, Unemployment, Market Structure and Growth. XXVII, 207 pages. 2003.

Vol. 531: M. Gallegati, A. Kirman, M. Marsili (Eds.), The Complex Dynamics of Economic Interaction. XV, 402 pages, 2004.

Vol. 532: K. Marti, Y. Ermoliev, G. Pflug (Eds.), Dynamic Stochastic Optimization. VIII, 336 pages. 2004.

Vol. 533: G. Dudek, Collaborative Planning in Supply Chains. X, 234 pages. 2004.

Vol. 534: M. Runkel, Environmental and Resource Policy for Consumer Durables. X, 197 pages. 2004.

Vol. 535: X. Gandibleux, M. Sevaux, K. Sörensen, V. T'kindt (Eds.), Metaheuristics for Multiobjective Optimisation. IX, 249 pages. 2004.

Vol. 536: R. Brüggemann, Model Reduction Methods for Vector Autoregressive Processes. X, 218 pages. 2004.

Vol. 537: A. Esser, Pricing in (In)Complete Markets. XI, 122 pages, 2004.

Vol. 538: S. Kokot, The Econometrics of Sequential Trade Models. XI, 193 pages. 2004.

Vol. 539: N. Hautsch, Modelling Irregularly Spaced Financial Data. XII, 291 pages. 2004.

Vol. 540: H. Kraft, Optimal Portfolios with Stochastic Interest Rates and Defaultable Assets. X, 173 pages. 2004.

Vol. 541: G.-y. Chen, X. Huang, X. Yang, Vector Optimization. X, 306 pages. 2005.

Vol. 542: J. Lingens, Union Wage Bargaining and Economic Growth. XIII, 199 pages. 2004.

Vol. 543: C. Benkert, Default Risk in Bond and Credit Derivatives Markets. IX, 135 pages. 2004.

Vol. 544: B. Fleischmann, A. Klose, Distribution Logistics. X, 284 pages. 2004.

Vol. 545: R. Hafner, Stochastic Implied Volatility. XI, 229 pages. 2004.

Vol. 546: D. Quadt, Lot-Sizing and Scheduling for Flexible Flow Lines. XVIII, 227 pages. 2004.

Vol. 547: M. Wildi, Signal Extraction. XI, 279 pages. 2005.

Vol. 548: D. Kuhn, Generalized Bounds for Convex Multistage Stochastic Programs. XI, 190 pages. 2005.

Vol. 549: G. N. Krieg, Kanban-Controlled Manufacturing Systems. IX, 236 pages. 2005.

Vol. 550: T. Lux, S. Reitz, E. Samanidou, Nonlinear Dynamics and Heterogeneous Interacting Agents. XIII, 327 pages. 2005.

Vol. 551: J. Leskow, M. Puchet Anyul, L. F. Punzo, New Tools of Economic Dynamics. XIX, 392 pages. 2005.

Vol. 552: C. Suerie, Time Continuity in Discrete Time Models. XVIII, 229 pages. 2005.

Vol. 553: B. Mönch, Strategic Trading in Illiquid Markets. XIII, 116 pages. 2005.

Vol. 554: R. Foellmi, Consumption Structure and Macroeconomics. IX, 152 pages. 2005.

Vol. 555: J. Wenzelburger, Learning in Economic Systems with Expectations Feedback (planned) 2005.

Vol. 556: R. Branzei, D. Dimitrov, S. Tijs, Models in Cooperative Game Theory. VIII, 135 pages. 2005.

Vol. 557: S. Barbaro, Equity and Efficiency Considerations of Public Higer Education. XII, 128 pages. 2005.

Vol. 558: M. Faliva, M. G. Zoia, Topics in Dynamic Model Analysis. X, 144 pages. 2005.

Vol. 559: M. Schulmerich, Real Options Valuation. XVI, 357 pages. 2005.

Vol. 560: A. von Schemde, Index and Stability in Bimatrix Games. X, 151 pages. 2005.

Vol. 561: H. Bobzin, Principles of Network Economics. XX, 390 pages. 2006.

Vol. 562: T. Langenberg, Standardization and Expectations. IX, 132 pages. 2006.

Vol. 563: A. Seeger (Ed.), Recent Advances in Optimization. XI, 455 pages. 2006.

Vol. 564: P. Mathieu, B. Beaufils, O. Brandouy (Eds.), Artificial Economics. XIII, 237 pages. 2005.

Vol. 565: W. Lemke, Term Structure Modeling and Estimation in a State Space Framework. IX, 224 pages. 2006.

Vol. 566: M. Genser, A Structural Framework for the Pricing of Corporate Securities. XIX, 176 pages. 2006.

Vol. 567: A. Namatame, T. Kaizouji, Y. Aruga (Eds.), The Complex Networks of Economic Interactions. XI, 343 pages. 2006.

Vol. 568: M. Caliendo, Microeconometric Evaluation of Labour Market Policies. XVII, 258 pages. 2006.

Vol. 569: L. Neubecker, Strategic Competition in Oligopolies with Fluctuating Demand. IX, 233 pages. 2006.

Vol. 570: J. Woo, The Political Economy of Fiscal Policy. X, 169 pages. 2006.

Vol. 571: T. Herwig, Market-Conform Valuation of Options. VIII, 104 pages. 2006.

Vol. 572: M. F. Jäkel, Pensionomics. XII, 316 pages. 2006

Vol. 573: J. Emami Namini, International Trade and Multinational Activity. X, 159 pages, 2006.

Vol. 574: R. Kleber, Dynamic Inventory Management in Reverse Logistics. XII, 181 pages, 2006.

Vol. 575: R. Hellermann, Capacity Options for Revenue Management. XV, 199 pages, 2006.

Vol. 576: J. Zajac, Economics Dynamics, Information and Equilibnum. X, 284 pages, 2006.

Vol. 577: K. Rudolph, Bargaining Power Effects in Financial Contracting. XVIII, 330 pages, 2006.

Vol. 578: J. Kühn, Optimal Risk-Return Trade-Offs of Commercial Banks. IX, 149 pages, 2006.

Vol. 579: D. Sondermann, Introduction to Stochastic Calculus for Finance. X, 136 pages, 2006.

Vol. 580: S. Seifert, Posted Price Offers in Internet Auction Markets. IX, 186 pages, 2006.

Vol. 581: K. Marti; Y. Ermoliev; M. Makowsk; G. Pflug (Eds.), Coping with Uncertainty. XIII, 330 pages, 2006.

Vol. 582: J. Andritzky, Sovereign Default Risks Valuation: Implications of Debt Crises and Bond Restructurings. VIII, 251 pages, 2006.

Vol. 583: I.V. Konnov, D.T. Luc, A.M. Rubinov† (Eds.), Generalized Convexity and Related Topics. IX, 469 pages, 2006.

Vol. 584: C. Bruun, Adances in Artificial Economics: The Economy as a Complex Dynamic System. XVI, 296 pages, 2006.

Vol. 585: R. Pope, J. Leitner, U. Leopold-Wildburger, The Knowledge Ahead Approach to Risk. XVI, 218 pages, 2007 (planned).

Vol. 586: B.Lebreton, Strategic Closed-Loop Supply Chain Management. X, 150 pages, 2007 (planned).

Vol. 587: P. N. Baecker, Real Options and Intellectual Property: Capital Budgeting Under Imperfect Patent Protection. X, 276 pages , 2007.

Vol. 588: D. Grundel, R. Murphey, P. Panos , O. Prokopyev (Eds.), Cooperative Systems: Control and Optimization. IX, 401 pages , 2007.

Vol. 589: M. Schwind, Dynamic Pricing and Automated Resource Allocation for Information Services: Reinforcement Learning and Combinatorial Auctions. XII, 293 pages , 2007.

Vol. 590: S. H. Oda, Developments on Experimental Economics: New Approaches to Solving Real-World Problems. XVI, 262 pages, 2007.

Vol. 591: M. Lehmann-Waffenschmidt, Economic Evolution and Equilibrium: Bridging the Gap. VIII, 272 pages, 2007.

Vol. 592: A. C.-L. Chian, Complex Systems Approach to Economic Dynamics. X, 95 pages, 2007.

Vol. 593: J. Rubart, The Employment Effects of Technological Change: Heterogenous Labor, Wage Inequality and Unemployment. XII, 209 pages, 2007

Vol. 594: R. Hübner, Strategic Supply Chain Management in Process Industries: An Application to Specialty Chemicals Production Network Design. XII, 243 pages, 2007

Vol. 595: H. Gimpel, Preferences in Negotiations: The Attachment Effect. XIV, 268 pages, 2007

Vol. 596: M. Müller-Bungart, Revenue Management with Flexible Products: Models and Methods for the Broadcasting Industry. XXI, 297 pages, 2007

Vol. 597: C. Barz, Risk-Averse Capacity Control in Revenue Management. XIV, 163 pages, 2007

Vol. 598: A. Ule, Partner Choice and Cooperation in Networks: Theory and Experimental Evidence. Approx. 200 pages, 2007

Vol. 599: A. Consiglio, Artificial Markets Modeling: Methods and Applications. XV, 277 pages, 2007

Vol. 600: M. Hickman, P. Mirchandani, S. Voss (Eds.): Computer-Aided Scheduling of Public Transport. Approx. 424 pages, 2007

Vol. 601: D. Radulescu: CGE Models and Capital Income Tax Reforms: The Case of a Dual Income Tax for Germany. XVI, 168 pages, 2007

Vol. 602: N. Ehrentreich: Agent-Based Modeling: The Santa Fe Institute Artificial Stock Market Model Revisited. XVI, 225 pages, 2008