

Artificial Markets and Intelligent Agents

by

Tung Chan

S.B. Massachusetts Institute of Technology (1995)

M.Eng. Massachusetts Institute of Technology (1995)

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

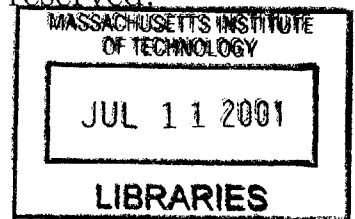
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

© Massachusetts Institute of Technology 2001. All rights reserved.

BARKER



Author
Department of Electrical Engineering and Computer Science
March 1, 2001

Certified by
Tomaso Poggio
Uncas and Helen Whitaker Professor of Brain and Cognitive Sciences
Thesis Supervisor

Certified by
Andrew W. Lo
Harris and Harris Group Professor
Supervisor

Accepted by
Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

Artificial Markets and Intelligent Agents

by

Tung Chan

Submitted to the Department of Electrical Engineering and Computer Science
on March 1, 2001, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In many studies of market microstructure, theoretical analysis quickly becomes intractable for all but the simplest stylized models. This thesis considers two alternative approaches, namely, the use of experiments with human subjects and simulations with intelligent agents, to address some of the limitations of theoretical modeling.

The thesis aims to study the design, development and characterization of artificial markets as well as the behaviors and strategies of intelligent trading and market-making agents. Simulations and experiments are conducted to study information aggregation and dissemination in a market. A number of features of the market dynamics are examined: the price efficiency of the market, the speed at which prices converge to the rational expectations equilibrium price, and the learning dynamics of traders who possess diverse information or preferences. By constructing simple intelligent agents, not only am I able to replicate several findings of human-based experiments, but I also find intriguing differences between agent-based and human-based experiments.

The importance of liquidity in securities markets motivates considerable interests in studying the behaviors of market-makers. A rule-based market-maker, built in with multiple objectives, including maintaining a fair and orderly market, maximizing profit and minimizing inventory risk, is constructed and tested on historical transaction data. Following the same design, an adaptive market-maker is modeled in the framework of reinforcement learning. The agent is shown to be able to adapt its strategies to different noisy market environments.

Thesis Supervisor: Tomaso Poggio

Title: Uncas and Helen Whitaker Professor of Brain and Cognitive Sciences

Thesis Supervisor: Andrew W. Lo

Title: Harris and Harris Group Professor

Acknowledgments

I wish to express my sincere gratitude to all those who helped me during the course of this work. Thanks to Professors Tomaso Poggio and Andrew Lo for their guidance of this research and their advice on both academic and professional matters. Thanks to my thesis readers, Professors Robert Berwick and Patrick Winston, for their helpful comments and suggestions. I am grateful to Prof. Blake LeBaron for his help on constructing my first set of market simulations; to Prof. Ely Dahan for allowing me to conduct trading experiments with his class; to Robert Thau for his help on building the web market software; to Mike Epstein for his insights from a practitioner's point of view. I am also thankful to all the people at CBCL, particularly Federico Giroi, Christian Shelton, Sayan Mukherjee, Tony Ezzat, Gadi Geiger and Adlar Kim, for their friendship and encouragement.

On a more personal note, I am grateful to Olivia Siu, for her effort in proofreading the entire thesis and her patience to listen over all the years. Finally I would like to thank my parents and my brother for their love and support that enabled me to reach this point.

Contents

1	Introduction	14
1.1	Motivations	14
1.2	Background	16
1.2.1	Market Microstructure	16
1.2.2	Experimental Markets	17
1.2.3	Simulated Markets	17
1.3	Major Contributions	18
1.4	Outline	19
2	Markets with Empirical Bayesian Traders	21
2.1	Introduction	21
2.2	Review of the Literature	23
2.2.1	Experimental Markets	23
2.2.2	Simulated Markets	24
2.3	Experimental Design	25
2.3.1	Market Structure and Economic Environment	25
2.3.2	Trading Mechanism	27
2.3.3	Agents	28
2.3.4	Learning Mechanism	30
2.4	Six Experiments	32
2.4.1	Information Aggregation and Identical Preferences	32
2.4.2	Information Dissemination and Identical Preferences	33
2.4.3	Information Aggregation and Heterogenous Preferences	33

2.4.4	Information Dissemination and Heterogenous Preferences . . .	34
2.4.5	Empirical Bayesian and Momentum Traders	34
2.4.6	Empirical Bayesian and Nearest-Neighbor Traders	35
2.5	Results and Discussion	35
2.5.1	Homogeneous Preferences	39
2.5.2	Heterogeneous Preferences	40
2.5.3	Momentum Traders	51
2.5.4	Nearest-Neighbor Traders	56
2.6	Conclusions	59
3	Experimental Markets for Product Concepts	61
3.1	Introduction	62
3.2	Background	64
3.2.1	Opinion-collecting Electronic Markets	64
3.2.2	Virtual Concept Testing	66
3.2.3	Rational Expectation Models and Experimental Markets . . .	67
3.3	Design of Markets and Securities	68
3.4	A Theoretical Market Equilibrium	71
3.5	Possible Trading Strategies	74
3.6	Simulations with Artificial Agents	75
3.7	Experimental Results and Comparison with Survey Study	85
3.8	Conclusions	93
4	A Rule-based Electronic Market-maker	95
4.1	Markets and Market-making	96
4.2	Background	100
4.2.1	Simulation Models with Market-making Rules	100
4.2.2	Empirical Studies	102
4.3	Design of the Market-maker	103
4.3.1	Market Quality (MQ)	105
4.3.2	Inventory Control (IC)	107

4.3.3	Speculative Motive (SM)	108
4.3.4	The Arbitration Mechanism	109
4.4	Simulations with Historical Data	110
4.4.1	The Data Set	110
4.4.2	Generation of the Order Flow	111
4.4.3	Simulations	114
4.5	Performance of the Rule-based Market-maker	117
4.5.1	Static Order Flow	117
4.5.2	Modeled Order Flow	124
4.6	Conclusions	124
5	An Adaptive Electronic Market-maker	128
5.1	Background	129
5.1.1	Market-making Models	129
5.1.2	Reinforcement Learning	132
5.2	A Reinforcement Learning Model of Market-making	136
5.2.1	Environment States	137
5.2.2	Market-maker's actions	138
5.2.3	Reward	139
5.3	The Basic Model	141
5.3.1	Market Structure	141
5.3.2	Strategies and Expected Profit	143
5.3.3	Market-making with Reinforcement learning Algorithms	149
5.3.4	Simulation Results	152
5.4	The Extended Model	158
5.4.1	Simulation Results	160
5.5	Conclusions	165
6	Conclusions	171

List of Figures

- 2-1 The experimental design of the Artificial Markets simulations 26
- 2-2a Prices, bid-ask spreads, and volume the early periods of Experiment
2.4.1 37
- 2-2b Prices, bid-ask spreads, and volume in the later periods of Experiment
2.4.1 37
- 2-2c Absolute price-deviations in Experiment 2.4.1 38
- 2-2d Empirical distribution of moving-average prices in Experiment 2.4.1 . 38
- 2-3a Prices, bid-ask spreads, and volume in the early periods of Experiment
2.4.2 41
- 2-3b Prices, bid-ask spreads, and volume in the later periods of Experiment
2.4.2 41
- 2-3c Absolute price-deviations versus number of momentum traders in Ex-
periment 2.4.2 42
- 2-3d Empirical distribution of moving-average prices in Experiment 2.4.2 . 42
- 2-3e Deciles of percentage wealth differences between insiders and unin-
formed traders in Experiment 2.4.2 43
- 2-4a Prices, bid-ask spreads, and volume in the early periods of Experiment
2.4.4 44
- 2-4b Prices, bid-ask spreads, and volume in the later periods of Experiment
2.4.4 44
- 2-4c Absolute price-deviations for markets with 0, 25 and 50 momentum
traders in Experiment 2.4.4 45
- 2-4d Allocative efficiency in Experiment 2.4.4 46

2-4e	Empirical distribution of moving-average prices in Experiment 2.4.4	46
2-5a	Prices, bid-ask spreads, and volume in the early periods of Experiment 2.4.3	47
2-5b	Prices, bid-ask spreads, and volume in the later periods of Experiment 2.4.3	47
2-5c	Absolute price-deviations of market prices in Experiment 2.4.3	48
2-5d	Allocative efficiency in Experiment 2.4.3	49
2-5e	Empirical distribution of moving-average prices in Experiment 2.4.3	49
2-6a	Absolute price-deviations of market prices in Experiment 2.4.5	52
2-6b	Absolute price-deviations of market prices in Experiment 2.4.5	53
2-6c	Empirical distribution of moving-average prices in Experiment 2.4.5	54
2-6d	Wealth differences between empirical Bayesian and momentum traders in Experiment 2.4.5	55
2-7a	Absolute price-deviations of market prices in Experiment 2.4.6	57
2-7b	Wealth differences between empirical Bayesian and nearest-neighbor traders in Experiment 2.4.6	58
3-1	The profiles of the eleven bike pumps.	70
3-2	Login and trading interface.	72
3-3a	An example of a non-linear demand function.	78
3-3b	Total demand of the market	78
3-4a	Typical realizations of markets in Simulation 1	82
3-4b	Deviations of mean, median and closing prices from the equilibrium price in Simulation 1.	83
3-4c	Standard deviation of agents' beliefs in Simulation 1	83
3-5a	Deviations of mean, median and close prices in Simulation 2	84
3-5b	Standard deviation of agents' beliefs in Simulation 2	85
3-6	Price and volume history for AirStik in Experiments 1 to 3.	87
3-7	Predicted market share based on the closing price, mean price, median price and VWAP	90

4-1	Empirical distribution $P(LPR BUY, p^a - p^b)$ for the stock IBM . . .	115
4-2	Two approaches of the order flow generation mechanism.	117
4-3	The effect of inventory control	119
4-4	The transaction price for IBM on November 1, 1990 by NYSE, MQ and SM+IC	123
5-1	The Markov chain describing Strategy 1 in the basic model	144
5-2	The Markov chain describing strategies 2 in the basic model	148
5-3	Expected profit per time step for Strategies 1, 2, and 3 in the basic model.	149
5-4	Examples of Q-functions for Strategies 1, 2 and 3. The bold values are the maximums for each row showing the resulting greedy policy.	150
5-5	A typical realization of Experiment 1	154
5-6a	End-of-episode profit of the market-maker in Experiment 1	155
5-6b	End-of-episode Inventory of the market-maker in Experiment 1	155
5-6c	Average absolute price deviation in Experiment 1	156
5-7	Percentages of SARSA simulations converges to Strategies 1, 2 and 3 in Experiment 1.	157
5-8	Percentages of the simulations of the Monte-Carlo method converges to Strategies 1, 2 and 3 in Experiment 2.	158
5-9	A typical realization of Experiment 3	162
5-10a	End-of-episode profit in Experiment 3	163
5-10b	Episodic average spread in Experiment 3	163
5-10c	Episodic average absolute price deviation in Experiment 3	164
5-10d	End-of-episode inventory in Experiment 3	164
5-11	Conditional probability distribution of actions given imbalance and spread in Experiment 3	166
5-12	Spread weight versus episodic average spread in Experiments 3 and 3a	167
5-13	Episodic average spread versus end-of-period profit in Experiment 3 .	168
5-14a	Empirical distribution of end-of-episode profit in Experiment 4	168

5-14b Empirical distribution of average episodic spread in Experiment 4 . . . 169

List of Tables

- 2.1 The order-submission algorithm of AI-agents 29
- 2.2 Summary of the six experiments 33
- 3.1 The order submission algorithm of the trading agents 79
- 3.2 Details of the three trading experiments. 86
- 3.3 Market share predictions for nine products concepts by the survey
study, and Experiments 1 to 3 89
- 3.4 Price correlation coefficients between outcomes from the survey study, and
Experiments 1, 2 and 3. 91
- 3.5 Rank correlation coefficients between outcomes from the survey study and
Experiments 1, 2 and 3. 92
- 3.6 Average absolute difference in predicted market share 93
- 4.1 Empirical distribution of *BUY* 113
- 4.2 Description of different modes of simulation. 118
- 4.3 Trading profits, inventory and volume and market quality measures for
the static order flow 121
- 4.4 Trading profits, inventory and volume and market quality measures for
the modeled order flow 125
- 5.1 Details of the experiments for the basic and extended models. 152

Chapter 1

Introduction

One of the most powerful ideas of modern economics is Adam Smith's (1776) Invisible Hand, the fact that agents acting in their own self-interest can reach an optimal allocation of scarce resources. This remarkable feature of perfectly competitive economies is due, of course, to the presence of markets, exchanges where buyers and sellers trade with each other and, in doing so, establish prices and quantities that equate supply and demand. Although these ideas were developed over two centuries ago, it is only within the past two or three decades that economists have begun to explore the specific mechanisms, i.e., the *market microstructure*, by which markets aggregate and disseminate information dynamically in a world of uncertainty and asymmetric information.

1.1 Motivations

In many of these investigations, the theoretical analysis quickly becomes intractable for all but the simplest stylized models, and even the existence of an equilibrium cannot be guaranteed in many cases.¹ An alternative to this theoretical approach is an experimental one in which individuals are placed in a controlled market setting,

¹See Cohen, Maier, Schwartz & Whitcomb (1986), Schwartz (1993), O'Hara (1995), and Campbell, Lo & MacKinlay (1996, Chapter 3) for overviews of the theoretical and empirical market microstructure literature.

given certain endowments of securities or cash or both, and allowed to trade with each other.² By varying the market structure, the design of the securities that can be traded, and the individuals' endowments, rewards, and information set, we can learn a great deal about the actual behavior of economic agents in a simple competitive environment and how markets perform their resource-allocation function so efficiently. Documenting and studying the interactions of optimizing individuals in an experimental setting is an important first step towards understanding their behavior in real markets.

However, the experimental-markets approach has its own limitations. In particular, although the market structure and economic environment are controlled by the experimenter, the motives and information-processing abilities of the economic agents are not. Therefore, it is often difficult to assess the impact of risk aversion, learning abilities, and the degree of individual rationality on prices and quantities in experimental markets. Moreover, there is no simple means to determine *how* agents process information and derive their trading rules in any given experiment, hence no assurance that any single experimental result is not an artifact of the particular subjects in the experiment.

Lastly, a third approach—the use of *artificially intelligent* agents—can be adopted to address some of the limitations of the theoretical and experimental alternatives. AI-agents are computer programs that contain certain heuristics and computational learning algorithms, with the intention of capturing particular aspects of human behavior. Although AI-agents are figments of our (and the computer's) imagination, their preferences and learning algorithms are transparent and, unlike experimental subjects, can be carefully controlled and modified. Using AI-agents, we can conduct a far broader set of experiments involving more complexities than with human agents. Moreover, the outcomes of such experiments are often more readily compared to theoretical models because we have eliminated the human “wildcard.”

This approach, now commonly known as “agent-based models,”³ allows us to

²Davis & Holt (1993) and Kagel & Roth (1995) are excellent surveys of this fast-growing literature.

³Another term that has been proposed is “agent-based computational economics” or ACE. See

explore new areas of economic theory, especially in dynamic markets with asymmetric information, learning, and uncertainty—a combination that poses many insurmountable technical challenges from a theoretical perspective. However, agent-based models also bring with them new and untested algorithms, parameters that must be calibrated, and other ad hoc assumptions that are likely to be controversial. To address these concerns, we propose using data from human experimental markets to validate and calibrate our agent-based models. In particular, we have designed our market structure along the same lines as those in the experimental-markets literature and show that *simple* AI-agents—agents endowed with only rudimentary computational learning abilities—can replicate several features of human-based experimental markets.

1.2 Background

The work presented in this thesis draws on at least three distinct literatures: the market microstructure literature, the experimental markets literature, and the simulated markets literature.

1.2.1 Market Microstructure

The literature of market microstructure provides important background and context for the experiments and simulations studied in this thesis. Many of the questions and issues that we focus on are those that the market microstructure literature has considered theoretically and empirically. Although our approach takes a decidedly different tack from the recent market microstructure literature, nevertheless, there are several important papers that provide motivation and inspiration for the agent-based models. For example, Garman (1976) developed one of the earliest models of dealership and auction markets and went so far as to deduce the statistical properties of prices by simulating the order-arrival process. Cohen, Maier, Schwartz &

<http://www.econ.iastate.edu/tesfatsi/ace.htm> for further discussion.

Whitcomb (1991) and Hakansson, Beja & Kale (1985) propose more complex simulation models of market-making activities. And a number of papers develop optimal market-making behavior in certain theoretical contexts, e.g., Amihud & Mendelson (1980), Ho & Stoll (1981), and Kyle (1985). Cohen et al. (1986), Schwartz (1993), O'Hara (1995), and Campbell et al. (1996, Chapter 3) provide excellent overviews of the market microstructure literature.

1.2.2 Experimental Markets

Another alternative to the theoretical approach is an experimental one in which individuals are placed in a controlled market setting, given certain endowments of securities and cash, and allowed to trade with each other. Davis & Holt (1993) and Kagel & Roth (1995) provide excellent coverage of the recent literature in experimental markets. In much of this literature, the rational expectations (RE) model has been the main benchmark, and has had mixed success in various studies. Studies of the informational efficiency of experimental markets relative to the RE benchmark generally fall into two categories: information dissemination between fully informed agents (“insiders”) and uninformed agents, and information aggregation among many partially informed agents. The former experiments investigate the common intuition that market prices reflect insider information, hence uninformed traders should be able to infer the true price from the market. The latter experiments explore the aggregation of diverse information by partially informed agents, a more challenging objective because none of the agents possesses full information (traders identify the state of nature with certainty only by pooling their private information through the process of trading).

1.2.3 Simulated Markets

Computer simulations of markets populated by software agents extend the experimental approach by allowing the experimenter to test various theories of learning behavior and market microstructure in a controlled environment. Unlike human-based experi-

ments, in which the dynamics of the subjects' behavior over many trading periods are almost never modeled explicitly, agent-based models can easily accommodate complex learning behavior, asymmetric information, heterogeneous preferences, and ad hoc heuristics.

Garman (1976), Cohen et al. (1991), and Hakansson et al. (1985) were early pioneers of agent-based models of financial markets. More recently, Gode & Sunder (1993) uses this framework to demonstrate a remarkable property of competitive markets: even in the absence of any form of learning or intelligence, markets with agents trading randomly eventually converge to the REE as long as budget constraints were continually satisfied.

Additional examples of trading algorithms for the simple double auction can be found in the report on the Santa Fe Institute Double Auction Tournament by Rust, Miller & Palmer (1992). This tournament focuses on the relative performance of various strategies played against each other. One of its key findings is that a very simple "parasite" strategy that feeds off the others performs best.

Finally, more complex computer-simulated asset markets that emphasize the evolution of trading behavior over time have also been created. LeBaron (forthcoming 1999) surveys many of these computational markets.⁴ These simulations attempt to capture long-range market phenomena as well as short-range trading dynamics, and share our emphasis of building behavioral theories starting at the individual level.

1.3 Major Contributions

This section summarizes the unique aspects and important contributions of this thesis. First, an agent-based approach is combined with an experimental approach to study markets and market participants. This provides a bridge between ad hoc learning models and experiments with human subjects. Data collected from two approaches are compared and contrasted. Specifically simulations are conducted to

⁴Examples include Routledge (1994, 1999), Arifovic (1996), Arthur, Holland, LeBaron, Palmer & Tayler (1997), Lettau (1997), Youssefmir & Huberman (1997).

replicate or verify the results from experiments. Second, a novel application of market mechanisms—the use of markets to collect consumer preferences—is proposed and studied in connection with experimental markets. Third, the feasibility of automated market-making is studied through the use of heuristic rules and an adaptive learning model. Historical data is used to calibrate and parameterize computer simulations, whose results can be compared with evidence from real-world markets. The adaptive model, developed in the reinforcement learning framework, is an original contribution. The model is able to generate strategies that work under different noisy market environments.

1.4 Outline

The thesis aims to study (1) the design, development and characterization of artificial financial markets, (2) the behaviors and strategies of artificially trading and market-making agents through computer simulations and market experiments. Simulations and experiments are conducted to study information aggregation and dissemination in a market. A number of features of the market dynamics are examined: the price efficiency of the market, the speed at which prices converge to the rational expectations equilibrium price, and the learning dynamics of traders who possess diverse information or preferences. By constructing simple intelligent agents, not only am I able to replicate several findings of human-based experiments, but I also find intriguing differences between agent-based and human-based experiments.

The importance of liquidity in securities markets motivates considerable interest in studying the behaviors of market-makers. A rule-based market-maker, built in with multiple objectives, including maintaining a fair and orderly market, maximizing profit and minimizing inventory risk, is constructed and tested on historical transaction data. Following the same design, an adaptive market-maker is modeled in the framework of reinforcement learning. The agent is shown to be able to adapt its strategies to different noisy market environments.

The rest of the thesis is organized as follows. Chapter 2 studies a series of com-

puter simulations designed to match those in experimental-market setting with human subjects. Chapter 3 discusses a novel application of the market mechanism to collect consumer preferences and presents the results from three market experiments. Chapter 4 proposes a rule-based market-maker with which simulations are conducted using historical data. Chapter 5 presents a reinforcement learning model for market-making. Finally Chapter 6 summarizes results of this thesis and discusses some ideas for future work in this area.

Chapter 2

Markets with Empirical Bayesian Traders

2.1 Introduction

Experimental asset markets have yielded many results on the properties of financial markets, and their abilities to disseminate and aggregate information. This understanding of the behavior of partially informed agents in experimental settings is a critical step toward understanding behavior in real markets. Various studies in experimental markets have shown that individuals are able to learn, and transmit information through prices in many different market situations. However, these studies are less specific about the actual mechanism that traders use to process information and learn from experience. This kind of generalization requires a deeper understanding of traders' trading strategies and the specification of the underlying learning processes.

We begin to address this question through the use of computational learning agents. These agents take the place of the experimental subjects and trade with each other in a simulated asset market. Unlike experimental subjects, the characteristics of the computer agents can be carefully controlled and modified to study the overall market behaviors with regard to different properties of the trader population. In the design of our trader agents, we strive to keep them as simple as possible in order to give us an idea of the lower bound of intelligence needed to replicate various market

phenomenon. This simplicity also makes the agents more open to detailed analysis on how they are processing market information.

Computational models allow us to explore new areas of economic theory, especially in dynamic market situations with learning. However, computational models bring with them new untested algorithms, and parameters. Questions about where theory ends, and how simple ad hoc mechanisms begin are quite valid. We believe that experimental data provides one useful route for validation. For this reason we design our markets to follow those as used in the experimental literature.

Specifically, we construct a double-auction market for a single stock that pays one liquidating state-contingent dividend at the end of each trading period, and we allow several types of AI-agents—each endowed with its own preferences, information, and learning algorithm—to trade with each other during repeated trading periods. In the course of six different experimental designs, we investigate a number of features of our agent-based model: the price efficiency of the market (how close market prices are to the rational expectations equilibrium (REE) prices), the speed at which prices converge to the REE, the dynamics of the distribution of wealth among the different types of AI-agents, trading volume, bid/ask spreads, and other aspects of market dynamics. In these experiments, we are able to replicate several findings of human-based experimental markets, e.g., the dissemination of information from informed to uninformed traders, the aggregation of information from traders with private information, and convergence to the REE price after a number of trading sessions.

However, we also find significant differences between agent-based and human-based experiments. For example, in one of our experiments in which agents have heterogeneous preferences and heterogeneous information, prices never converge to the REE; the opposite result was reported by Plott and Sunder (1982) in an experimental market with human subjects. Such differences may point to key features of human learning and inference that we have not captured in the design of our AI-agents, and are just as important as for developing a better understanding of how human markets operate as the features that we are able to replicate.

In Section 2.2, we provide a brief review of both the experimental and computa-

tional literatures. We describe our market environment in Section 2.3 and provide the details of the particular experiments we conduct. The results of those experiments are summarized in Section 2.5, and we conclude in Section 2.6.

2.2 Review of the Literature

2.2.1 Experimental Markets

The rational expectations (RE) model has received a considerable amount of attention in research on experimental markets. The RE model has had mixed success in various studies, depending on the complexity and structure of a market. The study of informational efficiency in the context of RE models can be categorized into two major areas. The first studies information dissemination from a group of insiders who have perfect information to a group of uninformed traders. The idea is that market prices reflect insider information so that uninformed traders can infer the true price from the market. The second examines information aggregation of diverse information in a market by a population of partially informed traders. Aggregation of diverse information is in general more difficult because no single agent possesses full information. Traders can identify the state of nature with certainty only by sharing their individual information in the process of trading.

Plott & Sunder (1982) and Forsythe, Palfrey & Plot (1982) investigate markets with insiders and uninformed traders. They show that equilibrium prices do reveal insider information after several trials of the experiments and conclude that the markets disseminate information efficiently. Furthermore, Plott & Sunder (1982) show that even in markets in which traders are paid different dividends (the same security pays one trader a dividend of 3 in state A but pays another trader a dividend of 5 in the same state, proxying for differences in preferences between the two traders), prices still converge to the REE. They attribute the success of the RE model to the fact that traders learn about the equilibrium price and the state of nature simultaneously from market conditions.

On the other hand, results by Plott & Sunder (1988) and Forsythe & Lundholm (1990) show that a market aggregates diverse information efficiently only under certain conditions: identical preferences, common knowledge of the dividend structure, and complete contingent claims. These studies provide examples of the failure of the RE model and suggest that information aggregation is a more complicated situation. In a related study, O'Brien & Srivastava (1991) find that market efficiency—defined as full information aggregation—depends on “complexity” of the market, as measured by market parameters such as the number of stocks and the number of trading periods in the market.

2.2.2 Simulated Markets

Experiments use simple economic theories to test convergence properties, but the dynamics of the subjects' behavior through the rounds is usually not modeled. The computer simulations performed here provide one possible method for testing the dynamics of learning in experimental settings, and developing theories in the form of agent algorithms which can be used to test further hypothesis on market designs and behavior.

Our agent design is based on the zero intelligence (ZI) traders used in Gode & Sunder (1993), where the generation of bids and offers contains a large random component. Gode & Sunder (1993) emphasizes the impact of budget constraints alone on observed prices and market efficiency. Several other authors have added varying degrees of intelligence to Gode and Sunder's “zero-intelligence” (ZI) traders by restricting the range of bids and asks that they generate. Usually these restrictions involve some function of recently observed trades or quotes. Two examples are Jamal & Sunder (1996) and Cliff & Bruten (1997); both implement simple heuristics to try to limit and improve on simple random bidding.

2.3 Experimental Design

Our experimental design consists of four components: the overall market structure and economic environment, the trading mechanism, the types of traders, and the learning algorithms that each type of trader employs. We describe each of these components in Sections 2.3.1–2.3.4, respectively.

2.3.1 Market Structure and Economic Environment

The general structure of our simulations is a double-auction market in which AI-agents trade a single security that pays a single liquidating state-contingent dividend at the end of a *trading period* by submitting orders for the security during the trading period. Each trading period consists of 40 *trading intervals*, and although the security pays no dividends until the last interval, trading occurs and information is revealed through prices and order flow in each interval. An *epoch* is defined to be a sequence of 75 consecutive trading periods, where an independently and identically distributed (IID) draw of the state of nature and private information is realized in each period. The state of nature is IID across periods and all the traders’ endowments are reset at the start of each period, but the traders become more “experienced” as they learn from one period to the next.¹ Each of the six experiments we conduct (see Section 2.4) consists of 100 trials of an epoch, where each epoch begins with the same initial conditions (types of traders, wealth distribution, etc.). This experimental design is summarized in Figure 2-1.

At the start of a period, three quantities are initialized (but not necessarily revealed): (1) the state of nature; (2) the agents’ endowments of cash and stock, which is identical across all agents throughout our experiments; and (3) the private information of each agent. At the end of a period, the predetermined state of nature is revealed and dividends are distributed to the shareholders.

The state of nature is random and exogenously determined, and the underlying

¹This, of course, applies only to those agents endowed with learning heuristics, e.g., empirical Bayesian and nearest-neighbor traders. See Section 2.3.4 for details.

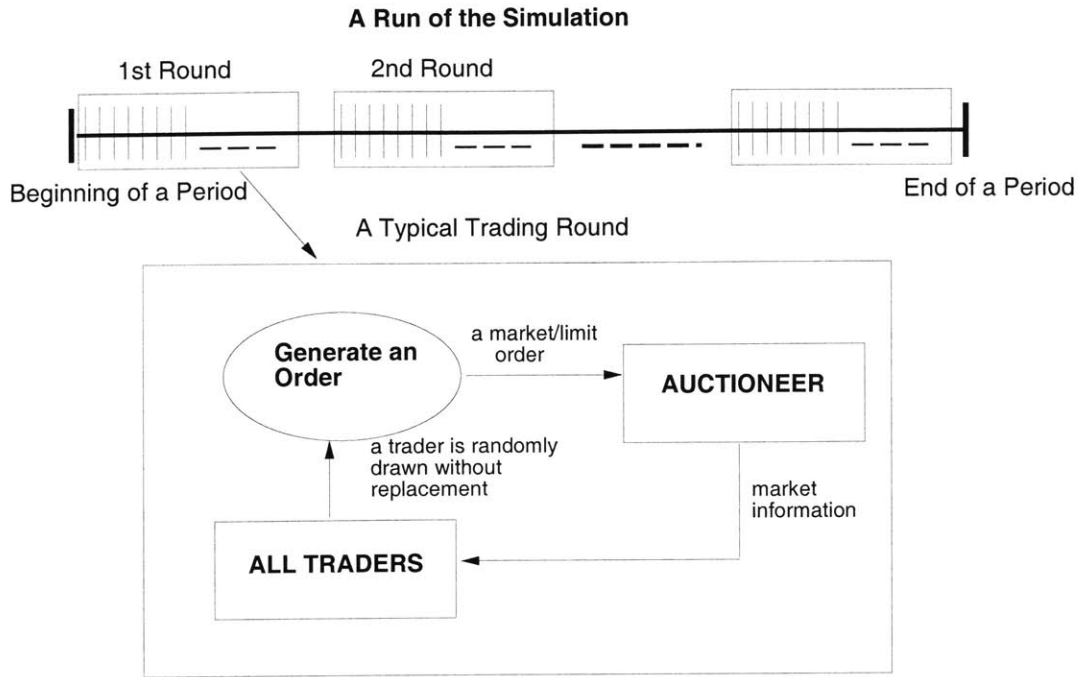


Figure 2-1: The experimental design of the Artificial Markets simulations. An epoch consists of 75 trading periods, and each period contains 40 trading intervals.

distribution of the state is common knowledge. For simplicity, we assume it is discrete and uniform. For example, in an economy with three states, each state has probability $1/3$ of occurring. We denote by $D = (0, 1, 2)$ a stock that pays a dividend of 0 in state 1, 1 in state 2, and 2 in state 3.

To model traders with homogeneous preferences, we assume that a security pays the same D regardless of who holds it. In contrast, to model traders with heterogeneous preferences, we assume that a security pays a *different* vector of dividends to different holders of the security. For example, in a market with two types of agents A and B , suppose the same security pays A a dividend of $D^a = (0, 1, 2)$, but pays B a dividend of $D^b = (2, 0, 1)$. This is a convenient device for capturing the fact that A may value a payoff in a particular state of nature more highly than B (in this example, A values a payoff in state 3 twice as highly as B). In economic terms, these agent-dependent payoffs may be viewed as marginal-utility-weighted payoffs (agents with different preferences will value identical dollar-payoffs differently). Heteroge-

neous preferences (payoffs) will be one motivation for trade in our market.

Differences in information about the likely state of nature is the other motive for trade. Information that is available to all market participants is public information, whereas information only known to some individuals is considered private information. The support of the distribution of dividends and their unconditional probabilities are public information, but some traders receive private information about the state of the nature. Specifically, traders are categorized into three groups according to their information: *insiders* know exactly which state will occur (for example, state 2 will occur, hence $D = (-, 1, -)$), *partially informed* traders who have imperfect information about the state (for example, state 3 will not occur, hence $D = (0, 1, -)$), and *uninformed* traders who have only public information (that is, $D = (0, 1, 2)$). Insiders and partially informed traders receive their private information at the beginning of each period. The distribution of private information is *not* common knowledge.

2.3.2 Trading Mechanism

The trading mechanism is a simplified double-auction market. Agents can either submit a bid or ask, or accept a posted bid or ask. If there is an existing bid for the stock, any subsequent bid must be higher than the current bid to be posted. Similarly, a subsequent ask following an existing ask must be lower than the current ask to be posted. A transaction occurs when an existing bid or ask is accepted (a market order matches with a limit order), or when the bid and ask cross (in which case the transaction price is set at the middle of the bid and ask).

For each trade, we restrict the quantity traded to be one share. There are two reasons for such a substantial simplification. First, allowing variable quantities complicates the analysis considerably, creating another strategic choice for which heuristics must be developed and then analyzed. Second, because one of the goals of this chapter is to determine the minimal level of intelligence required to replicate certain features of more sophisticated human markets, we wish to keep our model as simple as possible while retaining the most essential features of a securities market, e.g., prices as a medium of information dissemination and aggregation. However, we recognize

the importance of quantity as a choice variable—it is intimately associated with risk aversion, for example—and we hope to extend our analysis to incorporate variable shares traded in the near future.

No borrowing or short selling is permitted, and agents must satisfy their budget constraint at all times. Recall that each trading period consists of 40 trading intervals. At the beginning of each interval, a specific ordering of all the agents is drawn at random (uniformly). Following this randomly selected ordering, each agent submits one limit or market order. We fix the number of agents to be 20 for most of the experiments,² hence a maximum of $20 \times 40 = 800$ transactions can occur in any given period in such cases.

2.3.3 Agents

In designing our agents, we follow the spirit Gode and Sunder’s (1993) “zero-intelligence” (ZI) traders by using the simplest heuristics to give us a sense of the lower bound of intelligence needed to replicate various human-market phenomena. This simplicity also allows us to analyze more easily the interactions among agents and how information is disseminated and aggregated.

Specifically, all traders are assumed to be risk neutral, and they maximize their end-of-period expected wealth by choosing between cash and stock. Agents maximize the end-of-period expected value of their portfolios by forecasting the liquidating dividend, and then buying when market prices are low relative to their forecast and selling when market prices are high. Although we do not explicitly model the utility functions of the agents, we do allow for some basic differences in preferences by allowing the dividend payments to differ across agents (see Section 2.3.1). All agents submit orders according to the procedure described in Table 2.1 but they differ in how they determine the expected value of the stock p^* , which we call the *base price*.³ For

²In Experiment 2.4.6 we hold fixed the number of traders of one type while increasing the number of traders of another type.

³This procedure is inspired by the budget constrained ZI traders of Gode & Sunder (1993). It is also closely related to the heuristic trader mechanisms of Jamal & Sunder (1996) and Cliff & Bruten (1997), both of which suggest other methods for updating floor and ceiling levels which help

Scenario	Action
	<i>existing bid, existing ask</i>
$p^* > a$	buy at market
$p^* < b$	sell at market
$b < p^* < a$ and $a - p^* > p^* - b$	post an ask distributed $U(p^*, p^* + S)$
$b < p^* < a$ and $a - p^* \leq p^* - b$	post a bid distributed $U(p^* - S, p^*)$
	<i>no bid, existing ask</i>
$p^* > a$	buy at market
$p^* \leq a$	post a bid distributed $U(p^* - S, p^*)$
	<i>existing bid, no ask</i>
$p^* < b$	sell at market
$p^* \geq b$	post an ask distributed $U(p^*, p^* + S)$
	<i>no bid, no ask</i>
with probability 1/2	post an ask distributed $U(p^*, p^* + S)$
with probability 1/2	post a bid distributed $U(p^* - S, p^*)$

Table 2.1: The order-submission algorithm of AI-agents in the Artificial Markets simulations, where a denotes the best ask price, b the best bid price, p^* the agent’s base price, S the maximum spread from the base price, and $U(x_1, x_2)$ the uniform distribution on the open interval from x_1 to x_2 .

example, if there exists only an ask (no outstanding bid) and the agent’s base price is lower than the ask price, the agent posts a bid price that is uniformly distributed on the interval $(p^* - S, p^*)$, where p^* is the base price and S is a preset maximum spread.

Agents are of three possible types, depending on how they construct their forecasts: empirical Bayesian traders, momentum traders, and nearest-neighbor traders. Empirical Bayesian traders use market information to update their beliefs about the state of the economy.⁴ They form their base price using these beliefs, and attempt to buy (sell) if the base price is higher (lower) than the market price, in which case the stock is under-valued (over-valued) from their perspective. Empirical Bayesian traders continuously observe market activities, update their beliefs, and adjust their positions accordingly. They stop trading when either the market price approaches their base price, or they run out of cash or stock.

to constrain bid and ask ranges.

⁴We use the term “empirical Bayesian” loosely—our traders will not actually be correctly updating their priors using all available time series data since this would be too complicated. They simplify past prices using a moving average and this is used as a proxy for the complete history of observed data, which is then used to update their priors.

Momentum traders are simple technical analysis traders whose forecast of tomorrow’s return is today’s return. Specifically, if at time t the two most recent transaction prices are p_t and p_{t-1} , then a momentum trader’s forecast of the next transaction price is simply $p_t \times (p_t/p_{t-1})$. These traders reinforce and magnify the ups and downs of price movements, introducing extra volatility and irrational valuations of the security which make information aggregation and dissemination more difficult.

Nearest-neighbor traders attempt to exploit any patterns in historical prices to predict market prices by using a nearest-neighbor learning heuristic (see Section 2.3.4). If the empirical Bayesian traders are the “fundamental investors” of the market, the nearest-neighbor traders can be viewed as sophisticated “technicians”. Like the momentum traders, nearest-neighbor traders ignore any information regarding dividends and their associated probabilities. But instead of following a fixed strategy, they learn and adapt to changing market conditions.

2.3.4 Learning Mechanism

Empirical Bayesian traders condition their beliefs on market information. Specifically, the agents want to compute the expected dividend $E[D|p_0, p_1, \dots, p_t]$. For simplicity, we only consider transaction prices and ignore other market variables such as bid/ask prices and spreads and volume. We also assume that most of the relevant information is embedded in the transaction prices of the last k trades, hence a k -period moving average of prices m_t is used to summarize market information at time t ,

$$m_t = \frac{1}{k} \sum_{\tau=t-k+1}^t p_\tau . \tag{2.1}$$

We set $k = 10$ in our simulations. Given the series of moving-average prices m_k, m_{k+1}, \dots, m_t and the realized dividend D_i , the conditional distribution $P(m|D_i)$ can be estimated empirically, and using Bayes Theorem, $P(D_i|m)$ can be determined:

$$P(D_i|m) = \frac{P(m|D_i)P(D_i)}{\sum_{j=1}^N P(m|D_j)P(D_j)} \tag{2.2}$$

where $P(D_i)$ is the prior probability of dividend state i given by a trader's private information set, and N is the number of possible states. Consequently, for $D = (D_0, D_1, \dots, D_n)$ and given a moving-average price m , the conditional expectation of the dividend is

$$E[D|m] = \sum_{i=1}^N P(D_i|m)D_i \quad (2.3)$$

This conditional expectation is taken as the base price p^* for the empirical Bayesian traders. The order submission procedure, described in Table 2.1, is then followed.

In the actual implementation, the empirical Bayesian traders estimate the conditional density functions by constructing histograms with series of moving-average prices. Each histogram corresponds to a dividend state. A series is appended and the corresponding histogram is updated with the new moving-average prices after each period of an experiment. By participating in more periods, the empirical Bayesian traders attain more accurate estimates of the conditional probability. Intuitively, the empirical Bayesian traders learn the state by associating relevant market conditions with the realized state. They memorize these associations in form of histograms. These histograms give a picture of how well the agents discern different states given market data.

As for the nearest-neighbor traders, instead of observing the k -period moving-average prices, in each period i they form a sequence of n -tuples from the price series: $\mathbf{x}_n^i, \mathbf{x}_{n+1}^i, \dots, \mathbf{x}_{T_i}^i$ where:

$$\mathbf{x}_t = (p_{t-n+1}, p_{t-n+2}, \dots, p_t) \quad , \quad t = k, k+1, \dots, T, \quad (2.4)$$

p_t is the market at time t , and T_i is the number of transactions in the period. Similar to the empirical Bayesians, the nearest-neighbor traders believe that all relevant information is embedded in the prices of the last n transactions. We set $n = 5$ in our experiments. Each of the n -tuples, \mathbf{x}_t^i , is associated with the end-of-period REE price, or dividend D_i , depending on the state of the economy. The pairs

$(\mathbf{x}_n^i, D_i), (\mathbf{x}_{n+1}^i, D_i), \dots, (\mathbf{x}_{T_i}^i, D_i), (\mathbf{x}_n^{i+1}, D_{i+1}), \dots$ and so on represent the “memory” of a nearest-neighbor trader. The nearest-neighbor traders predict the dividend by first observing the most recent n -tuple in the current market, \mathbf{x}_t^j , then finding its r nearest neighbors in terms of Euclidean distance from memory. The forecast is defined to be the mean of the associated dividends of the r nearest neighbors.

The parameter r controls the robustness of the prediction by governing the trade-off between bias and variance of the estimate. If r is too large, the bias becomes large and the estimate is inaccurate. If r is too small, the variance is high and the estimate is noisy and sensitive to individual data points. Through simple trial-and-error, we settled on $r = 10$ as the best compromise between mean-squared-error and computational speed, but no formal optimization was performed.

2.4 Six Experiments

We conduct six distinct experiments, each consisting of 100 trials of an epoch (recall that an epoch is comprised of 75 consecutive trading periods). The market and information structures are identical across the six experiments, but we vary the composition of traders and the diversity of preferences. These differences are described in Sections 2.4.1–2.4.6 and summarized in Table 2.2. In all six experiments, we assume that there are three states of nature that occur with equal probability (unconditionally), and unless indicated otherwise, all agents begin each period with 10 units of cash and 5 shares of stock.

2.4.1 Information Aggregation and Identical Preferences

This experiment contains 20 agents with identical preferences (hence the dividend payoff D is the same for each agent) and all agents are partially informed that one of the three states is impossible. For example, if state 1 is the state that will be realized at the end of the period, at the beginning of the period one trader is informed that state 0 will not occur, and another trader is informed that state 2 will not occur. Although none of the traders knows in advance which state will occur, collectively,

Experiment	Information	Preferences	Dividend	Endowment		Number of Traders		
				Cash	Stock	E.B.	Mom.	N.N.
1	Aggregation	Homogeneous	(0, 1, 2)	10	5	20	0	0
2	Dissemination	Homogeneous	(0, 1, 2)	10	5	20	0	0
3a	Aggregation	Heterogeneous	(0, 1, 3) ^a	10	5	10	0	0
			(2, 0, 1) ^b			10	0	0
3b	Aggregation	Heterogeneous	(0, 1, 3) ^a	40	5	10	0	0
			(2, 0, 1) ^b			10		
4a	Dissemination	Heterogeneous	(0, 1, 3) ^a	10	5	10	0	0
			(2, 0, 1) ^b			10		
4b	Dissemination	Heterogeneous	(0, 1, 3) ^a	40	5	10	0	0
			(2, 0, 1) ^b			10		
5	Aggregation	Homogeneous	(0, 1, 2)	10	5	20	0, ..., 150	0
6	Aggregation	Homogeneous	(0, 1, 2)	10	5	15	0	5

Table 2.2: Summary of the six experiments conducted in the Artificial Markets simulations. ‘E.B.’, ‘Mom.’, and ‘N.N.’ denote the number of Empirical Bayesian, momentum, and nearest-neighbor traders in each experiment. Each experiment consists of 100 statistically independent repetitions of 75 trading periods.

the market has perfect information about which state will occur. The REE price is simply the value of D in the realized state of nature, and the dividend payoff is $D = (0, 1, 2)$.

2.4.2 Information Dissemination and Identical Preferences

This experiment contains 20 agents with identical preferences, but there are 10 *insiders* who know what the state of nature is, and 10 uninformed traders who have only public information, i.e., the distribution of D . The REE price is D in the realized state, and the dividend payoff is $D = (0, 1, 3)$.

2.4.3 Information Aggregation and Heterogenous Preferences

This experiment contains 20 agents divided into two groups of 10 according to their preferences. In the three possible states of nature, Group A receives a dividend $D^a = (0, 1, 3)$ and Group B receives $D^b = (2, 0, 1)$. All traders have private information which rules out one of the two states that will not occur. Given the state of nature, the REE price is the higher of D^a and D^b in that particular state. For example, given

that state 2 will occur, the REE price is 3.

We run this experiment twice. In the first run, we set agents' endowments at the usual levels: 10 units of cash and 5 shares of stock. In the second run, we increase each agent's cash endowment to 40 units, relaxing budget constraints considerably.

2.4.4 Information Dissemination and Heterogenous Preferences

There are two groups of traders with diverse preferences. Group A receives dividend $D^a = (0, 1, 3)$ and group B receives $D^b = (2, 0, 1)$. There are 5 insiders and 5 uninformed traders in groups A and B, respectively. The REE price is the higher of D^a and D^b given the state.

As in Experiment 2.4.3, we run this experiment twice. In the first run, we set agents' endowments at the usual levels: 10 units of cash and 5 shares of stock. In the second run, we increase each agent's cash endowment to 40 units, relaxing budget constraints considerably.

2.4.5 Empirical Bayesian and Momentum Traders

In this experiment we test the robustness of our market's price-discovery mechanism by varying the proportion of empirical Bayesian and momentum traders in the population. The empirical Bayesian traders provide the market with information and, by their trading activities, move market prices towards the REE. The momentum traders, on the other hand, introduce a substantial amount of noise and volatility into market prices. How much noise can the market "tolerate" before the price-discovery mechanism breaks down, i.e., prices no longer converge to the REE?

To answer this question, we fix the number of empirical Bayesian traders at 20 and perform a sequence of 14 experiments in which the number of momentum traders is increased incrementally from 0 in the first run to 150 in the 14th run.⁵ By maintaining

⁵ Specifically, the 14 runs correspond to experiments with 0, 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 125, and 150 momentum traders, respectively.

the same number of empirical Bayesian traders across these 14 runs, we keep constant the amount information in the market while successively increasing the amount of noise induced by momentum traders.

2.4.6 Empirical Bayesian and Nearest-Neighbor Traders

In this experiment we have 15 empirical Bayesian traders and 5 nearest-neighbor traders. The nearest-neighbor traders are designed to exploit any predictability in prices, hence their trading performance is a measure of the market’s weak-form efficiency. If the market is weak-form efficient, then the empirical Bayesian traders should perform at least as well as the nearest-neighbor traders (because there is nothing for the “technicians” to pick up). On the other hand, if prices contain predictable components, the nearest-neighbor traders should outperform the empirical Bayesians.

2.5 Results and Discussion

In all six experiments, we focus on the informational efficiency of the market, i.e., do “prices fully reflect all available information”? Specifically, we compare market prices to their REE counterpart by measuring their average absolute price-deviation:

$$\Delta_p = \frac{1}{T} \sum_{t=1}^T |p_t - D| \quad (2.5)$$

where p_t is the transaction price and D is the REE price, and by considering the rate of convergence of p_t to D over the epoch.

In addition, we investigate bid-ask spreads, trading volume, and the wealth distribution across the different types of traders. Narrowing bid-ask spreads show that prices are converging, implying that buyers and sellers are reaching a common price. Diminishing volume, on the other hand, suggests that the market is approaching its equilibrium. This is either because all traders come to the same expected price and therefore have no incentives to trade, or they simply run out of cash or stock to trans-

act further. And the difference in wealth between two types of traders provides an indication of the economic impact of the differences among the traders. For example, in the case of insiders versus uninformed traders (Experiment 2.4.2), the differences in wealth between the two groups provide a measure of the value of insider information. We measure this difference as $\Delta_w(i, j)$ where

$$\Delta_w(i, j) \equiv \frac{W_i - W_j}{W_j} \times 100 \quad (2.6)$$

and W_i and W_j are the total wealth levels of the two types of traders.

We also investigate the expectations formed by the agents by examining their empirical conditional density functions of the moving-average price given the states. This collection of conditional density functions represents the agents' beliefs formed with their prior information and updated continuously with market prices. The agents use these density functions to distinguish one state from another, hence these functions are central to understanding how the agents learn.

In experiments that have a diverse dividend structure, we define *allocative efficiency*, following Smith (1962), as the ratio between total dividends earned by all traders and the total maximum dividends that can possibly be extracted from the market. For example, 100% allocative efficiency implies that all shares are held by traders in the group that receives the highest dividend in the realized state. The REE predicts 100% allocative efficiency in that all shares will be allocated to the traders valuing them most highly.

Recall that each experiment consists of 100 trials of an epoch consisting of 75 consecutive trading periods, and a trading period contains 40 trading intervals. Because of the enormous quantity of data generated from these simulations, it is difficult to provide numerical summaries of the results. Therefore, we summarize our findings in a series of graphs (Figures 2-2a–2-7b) and discuss them in Sections 2.5.1–2.5.4.

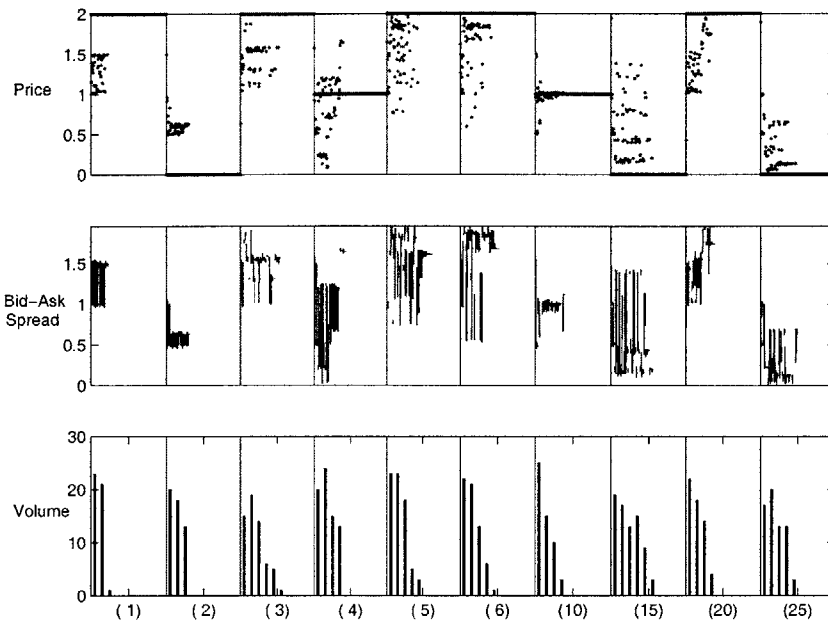


Figure 2-2a: Prices, bid-ask spreads, and volume in the early periods of a typical realization of Artificial Markets Experiment 2.4.1 (information aggregation with identical preferences).

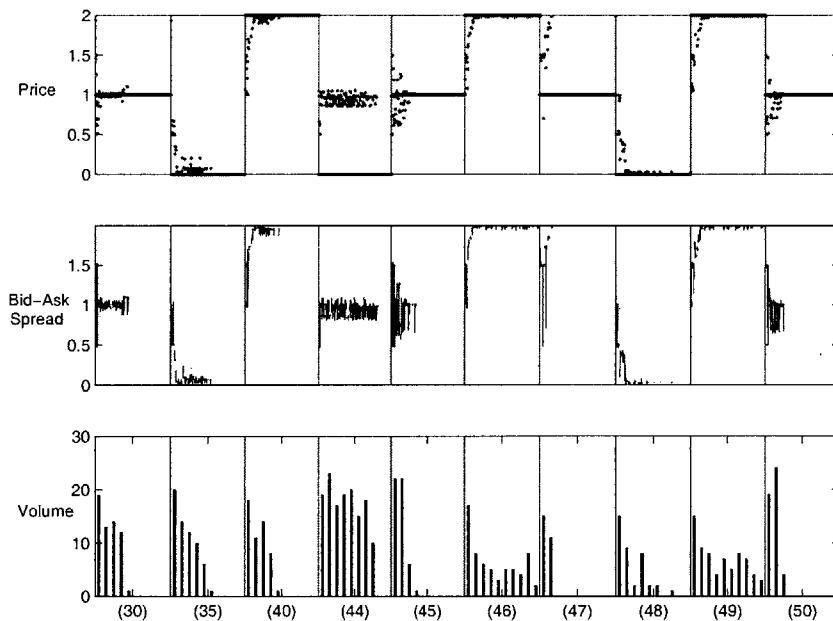


Figure 2-2b: Prices, bid-ask spreads, and volume in the later periods of a typical realization of Artificial Markets Experiment 2.4.1 (information aggregation with identical preferences).

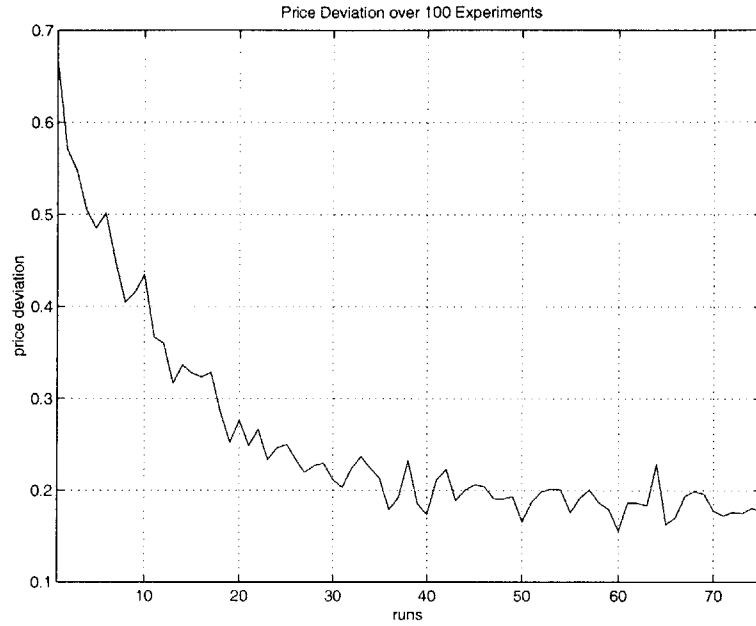


Figure 2-2c: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions of Artificial Markets Experiment 2.4.1 (information aggregation with identical preferences).

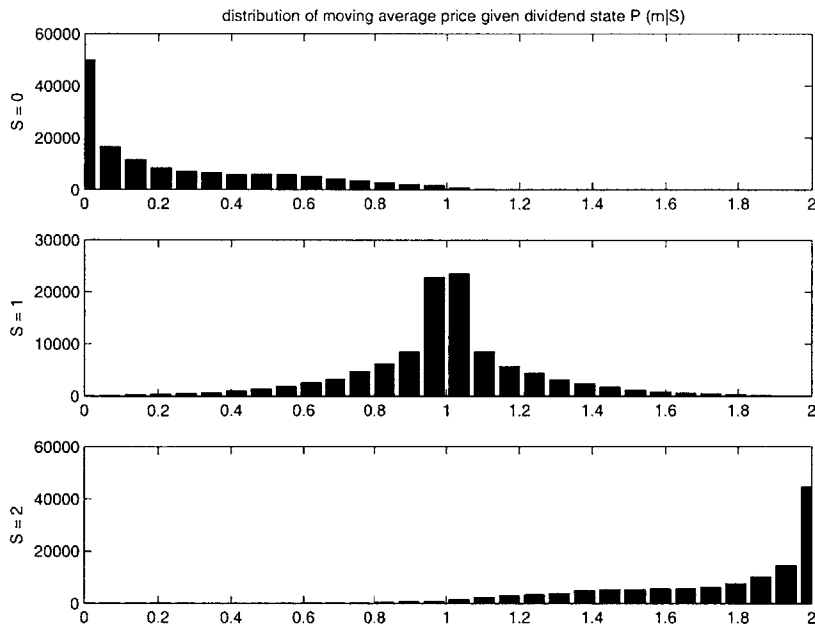


Figure 2-2d: Empirical distribution of moving-average prices, conditioned on the state of nature S , in Artificial Markets Experiment 2.4.1 (information aggregation with identical preferences).

2.5.1 Homogeneous Preferences

With identical preferences, the results from our simulation are similar to those in the human-based experimental markets literature, and are captured in Figures 2-2a and 2-2b which plot the transaction, bid, and ask prices and volume in selected periods of a typical epoch of Experiment 2.4.1.

In this experiment the convergence to the REE is apparent. Figure 2-2a shows market activities in the earlier periods of the market in a typical trial of the experiment. In this stage, agents are actively learning and observing, with little evidence of convergence. However, in the later periods (see Figure 2-2b), after agents have accumulated sufficient knowledge regarding how states and prices are related, convergence becomes more apparent.

Figure 2-2c plots the average price-deviations Δ_p (see (2.5)) for each of the 75 periods of the epoch, averaged over the 100 trials of the experiment. Market efficiency clearly improves substantially over the epoch. Figure 2-2d plots the conditional distribution of the moving-average price (see (2.1)) for each of the three states, obtained by summing up the frequency counts for m_t across the 100 trials and for each state. These histograms show that the three states are clearly distinguishable by the agents.

In Experiment 2.4.2, the evidence of convergence is even more compelling (see Figures 2-3a and 2-3b). In contrast to Experiment 2.4.1, prices converge faster in this experiment and are closer to the REE price (Figure 2-3c), and bid-ask spreads are smaller. There are two reasons for such a difference in the two experiments, despite the fact that both markets have approximately the same amount of information. First, in Experiment 2.4.1 traders must trade with each other to “pool” their information to determine the correct price, whereas in Experiment 2.4.2 the insiders know the correct price. Second, in the former case the distribution of information to the traders is random. For example, there may be many more traders given the information $D = (0, 1, -)$ than those given $D = (-, 1, 3)$, biasing the consensus in one direction or another.

Figure 2-3e plots the cross-sectional distribution of percentage wealth differences

Δ_w (see (2.6)) between the informed and uninformed traders for the 75 trading periods. For each period, we compute the average wealth within the two groups, take the percentage difference, and plot the deciles of these differences over the 100 trials. Not surprisingly, insiders have a substantially higher wealth than the uninformed. The difference in their wealth represents the value of the insider information and may be an estimate of the price traders would be willing to pay if information signals were sold. Observe that the value of insider information is diminishing over the epoch as uninformed traders learn. This is consistent with results from human-based experimental markets such as Sunder (1992) in which information is sold in a sealed bid auction. In such experimental markets, traders lower their bids for information once they learn to infer the states after a few periods of experience.

2.5.2 Heterogeneous Preferences

In contrast to the identical-preference cases, the prices in experiments involving diverse preferences do not seem to converge to the REE price. This can be explained by the fact that our agents attempt to recover the state of nature from market information alone, and not from the preferences of other agents (which is not common knowledge), despite the fact that heterogeneity is an important feature of their world. In fact, they are not even “aware” of the possibility of differences in dividend payoffs across traders.

Figures 2-5a–2-4e summarize the results from Experiments 2.4.3 and 2.4.4. Because our agents must infer the state of nature from market prices alone, we expect the REE model to fail in both experiments. The intuition for this conjecture comes from the fact that market prices are less useful for discriminating among states of nature in the presence of heterogeneity. For example, Figure 2-4e plots the conditional distribution of the moving-average price in Experiment 2.4.4; the probability of such a realization is almost identical in states 1 and 2, making the two virtually indistinguishable. Even if agents were told which state will occur, they would still have trouble reaching a unanimous price because of the heterogeneity in their preferences.

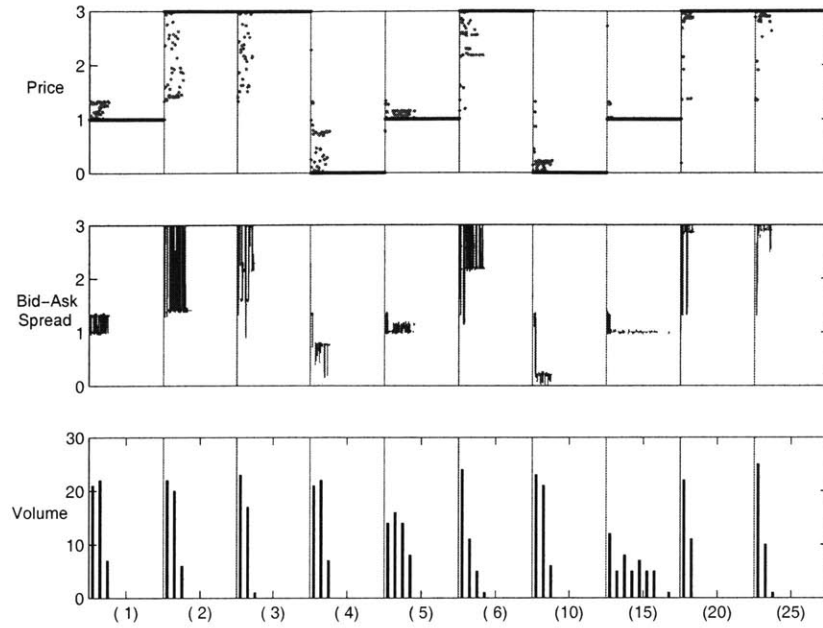


Figure 2-3a: Prices, bid-ask spreads, and volume in the early periods of a typical realization of Artificial Markets Experiment 2.4.2 (information dissemination with identical preferences).

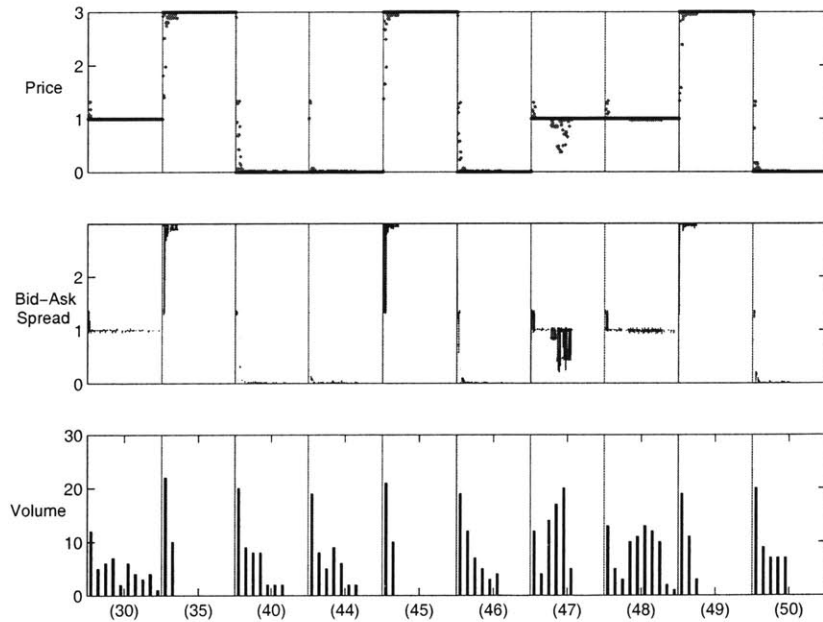


Figure 2-3b: Prices, bid-ask spreads, and volume in the later periods of a typical realization of Artificial Markets Experiment 2.4.2 (information dissemination with identical preferences).

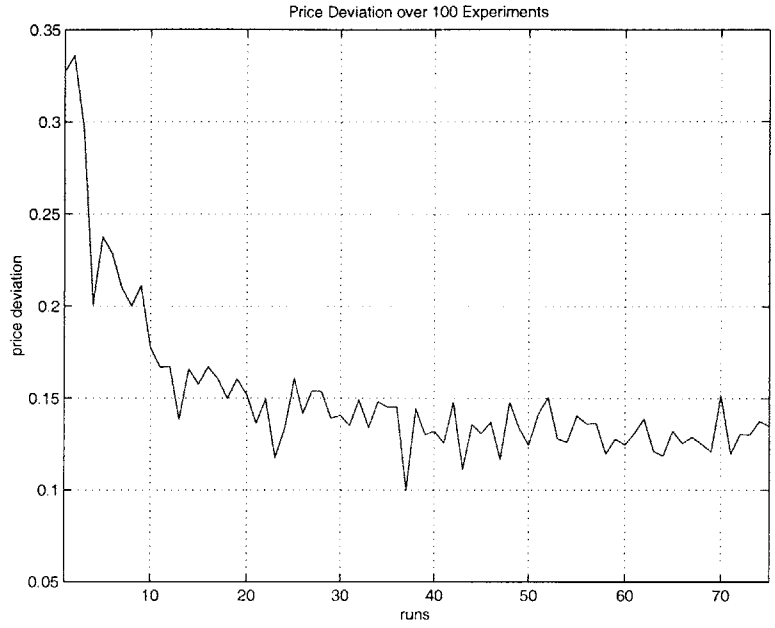


Figure 2-3c: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions of Artificial Markets Experiment 2.4.2 (information dissemination with identical preferences).

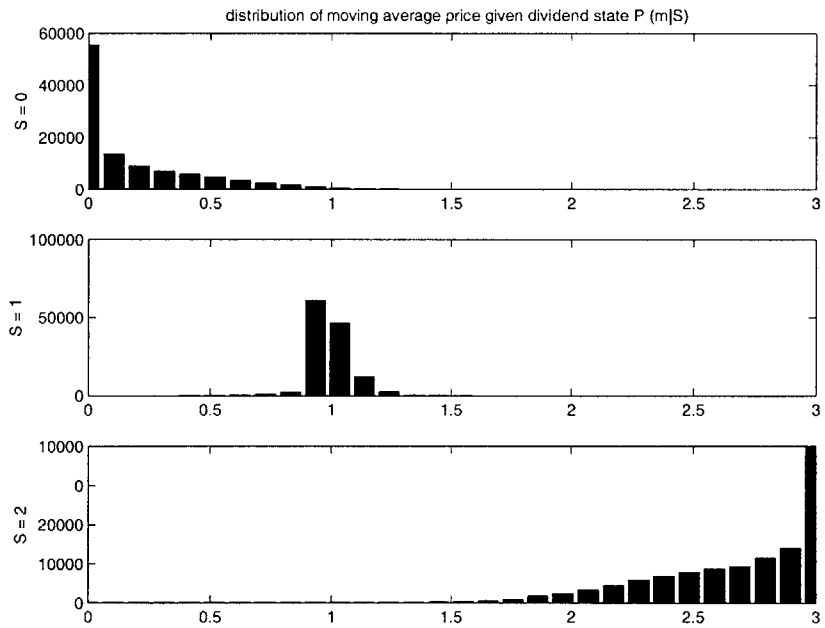


Figure 2-3d: Empirical distribution of moving-average prices, conditioned on the state of nature S , in Artificial Markets Experiment 2.4.2 (information dissemination with identical preferences).

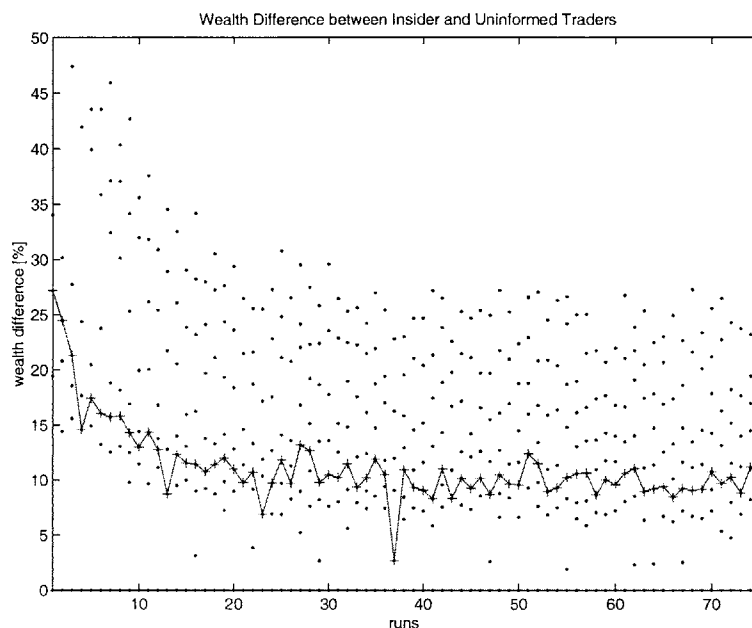


Figure 2-3e: Deciles of percentage wealth differences between insiders and uninformed traders in 100 repetitions of Artificial Markets Experiment 2.4.2 (information dissemination with identical preferences). Medians are indicated by the symbol ‘+’.

However, the degree of market efficiency—as measured by the average absolute price-deviation and allocative efficiency—is influenced by the traders’ initial cash endowments. The outcomes of two experiments, a low-cash (10 units) and a high-cash (40 units) experiment, are summarized in Figures 2-5c and 2-5d. These figures plot average absolute price-deviations and allocative efficiency, respectively, for the two experiments over the 75 periods of an epoch and averaged over 100 trials. Figure 2-5c shows that the standard cash endowment of 10 units does not lead to convergence; average absolute price-deviations and allocative efficiencies do not improve much over the 75 periods. However, an initial cash endowment of 40 units does yield some convergence in Experiments 2.4.3 and 2.4.4.

A concrete example will help to illustrate how the market reaches equilibrium in the high-cash case. In Experiment 2.4.3, type A and type B insiders will receive 3 and 1, respectively, for one share of the stock in state 3. These are their reservation prices. Agents will not buy above or sell below these prices. Between the two groups

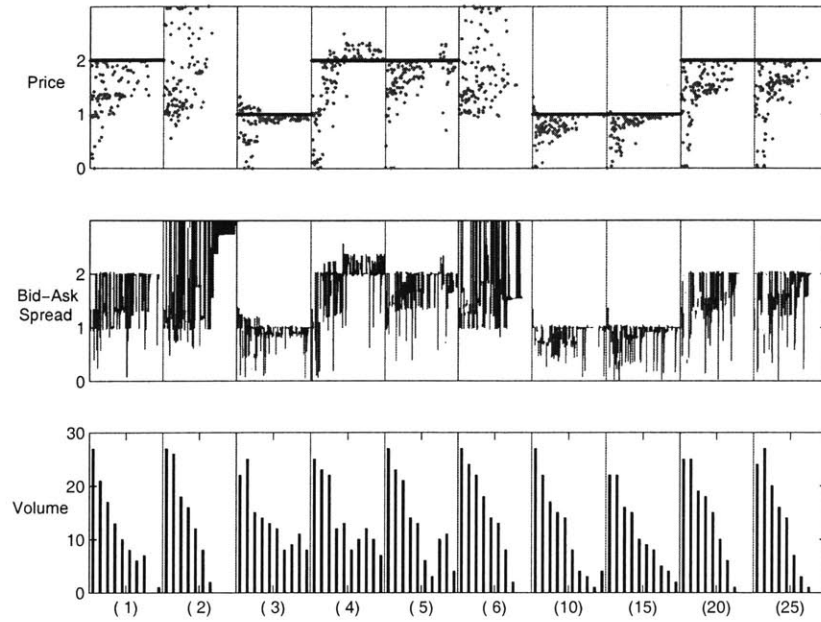


Figure 2-4a: Prices, bid-ask spreads, and volume in the early periods of typical realization of Artificial Markets Experiment 2.4.4 (information dissemination with heterogeneous preferences).

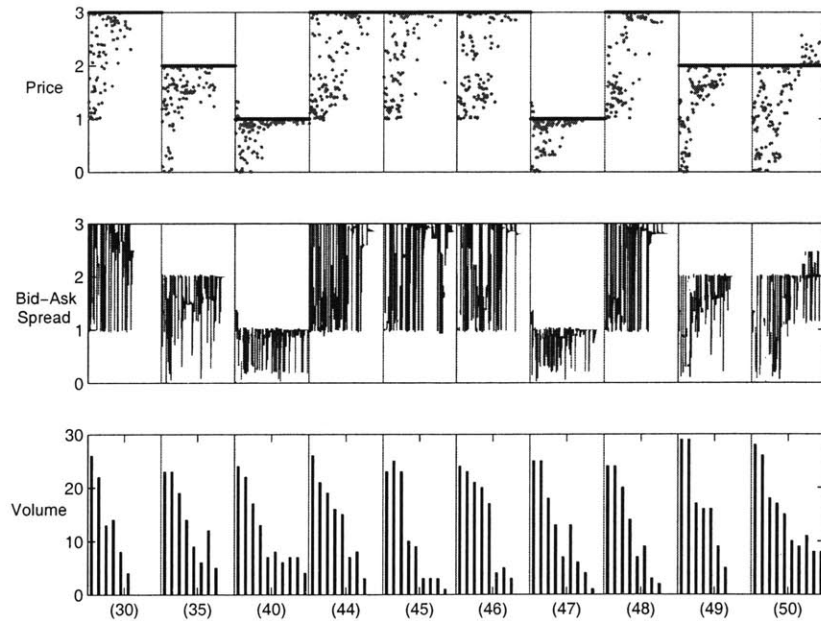


Figure 2-4b: Prices, bid-ask spreads, and volume in the later periods of a typical realization of Artificial Markets Experiment 2.4.4 (information dissemination with heterogeneous preferences).

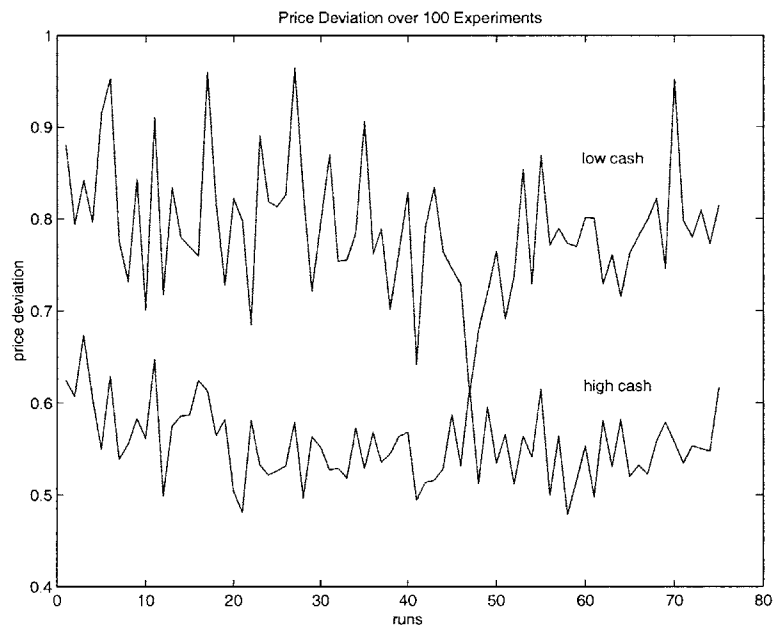


Figure 2-4c: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions of each of two runs of Artificial Markets Experiment 2.4.4 (information dissemination with heterogeneous preferences), the ‘low-cash’ and ‘high-cash’ experiments.

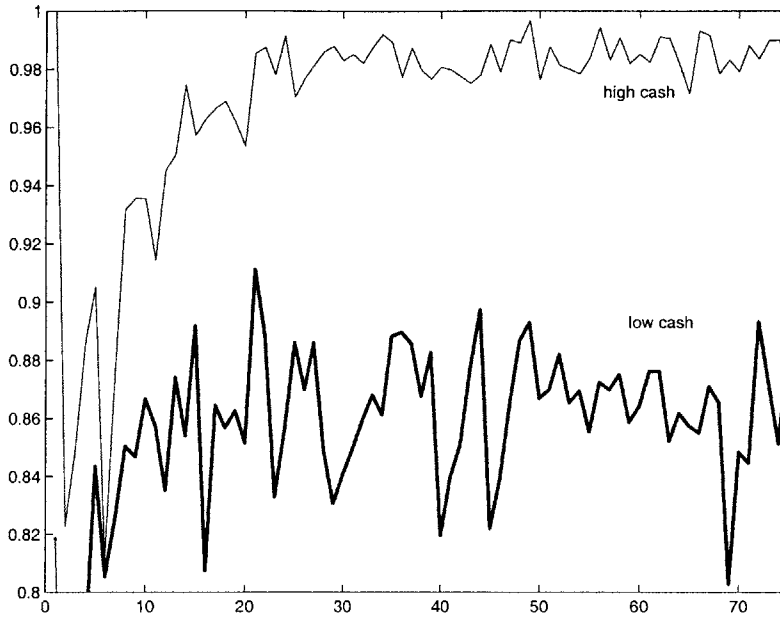


Figure 2-4d: Allocative efficiency, averaged over 100 repetitions of each of two runs of Artificial Markets Experiment 2.4.4 (information dissemination with heterogeneous preferences), the ‘low-cash’ and ‘high-cash’ experiments.

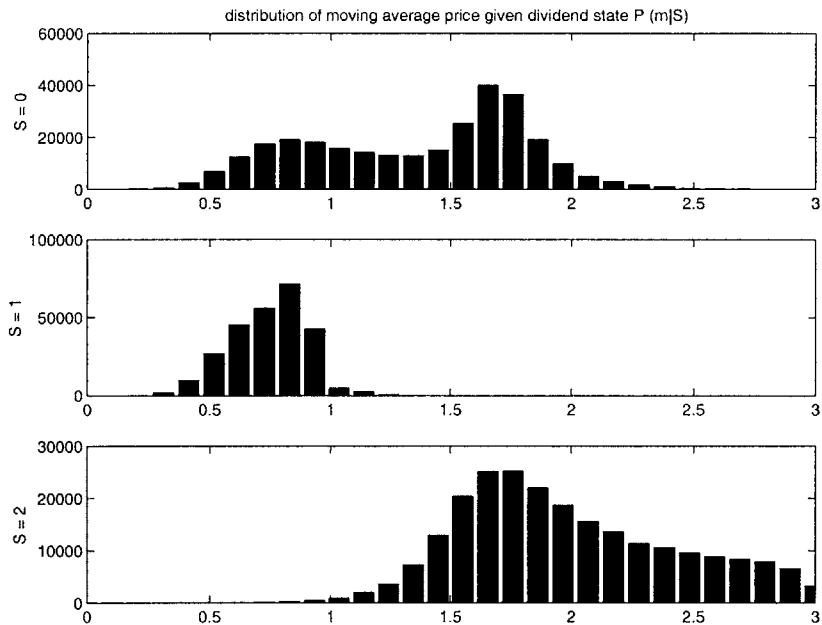


Figure 2-4e: Empirical distribution of moving-average prices, conditioned on the state of nature S , in Artificial Markets Experiment 2.4.4 (information dissemination with heterogeneous preferences).

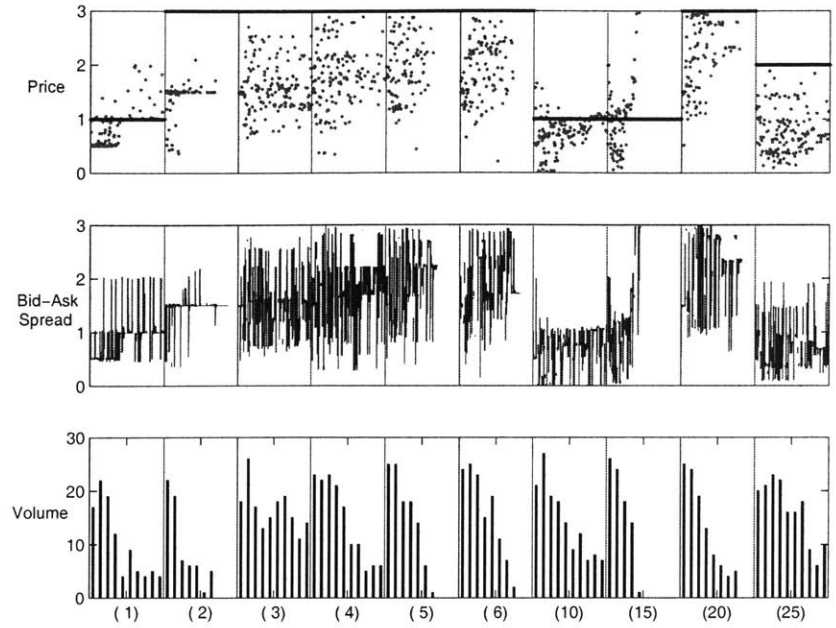


Figure 2-5a: Prices, bid-ask spreads, and volume in the early periods of a typical realization of Artificial Markets Experiment 2.4.3 (information aggregation with heterogeneous preferences).

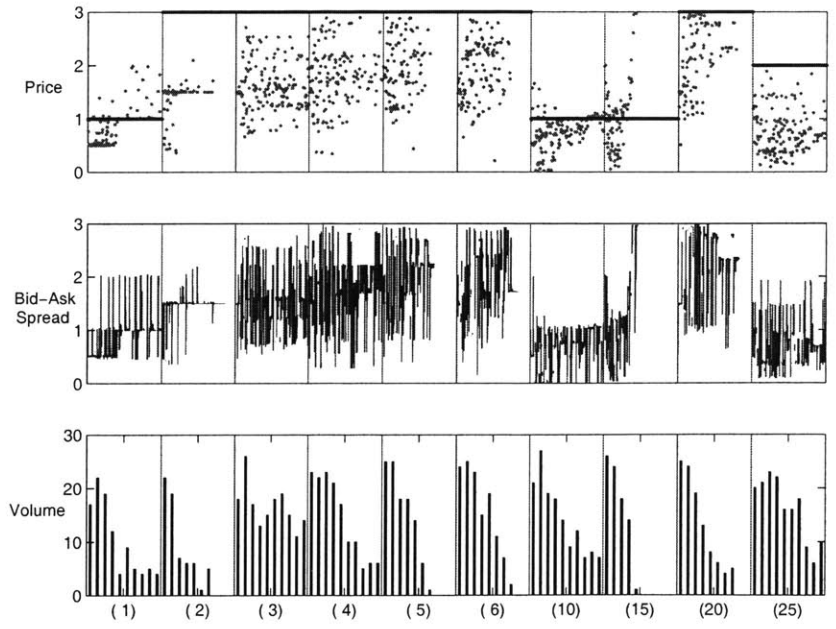


Figure 2-5b: Prices, bid-ask spreads, and volume in the later periods of a typical realization of Artificial Markets Experiment 2.4.3 (information aggregation with heterogeneous preferences).

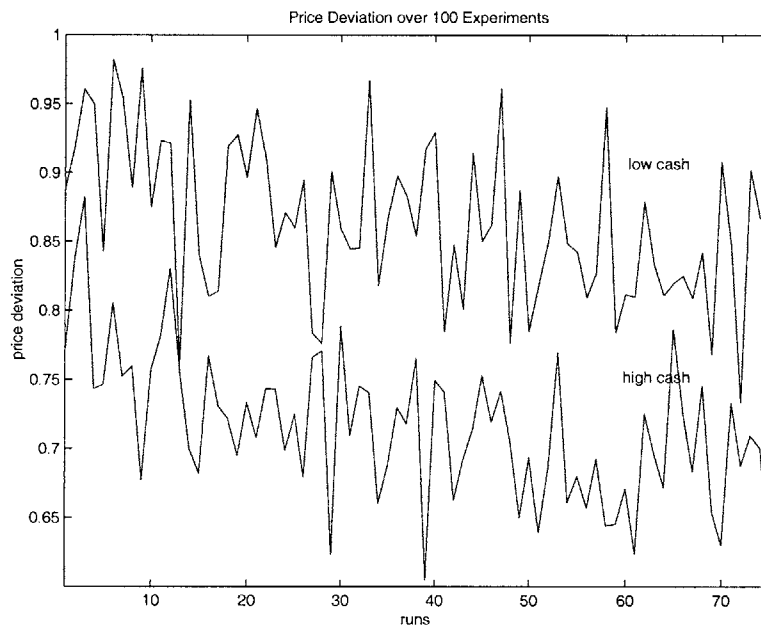


Figure 2-5c: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions of each of two runs of Artificial Markets Experiment 2.4.3 (information aggregation with heterogeneous preferences), the 'low-cash' and 'high-cash' experiments.

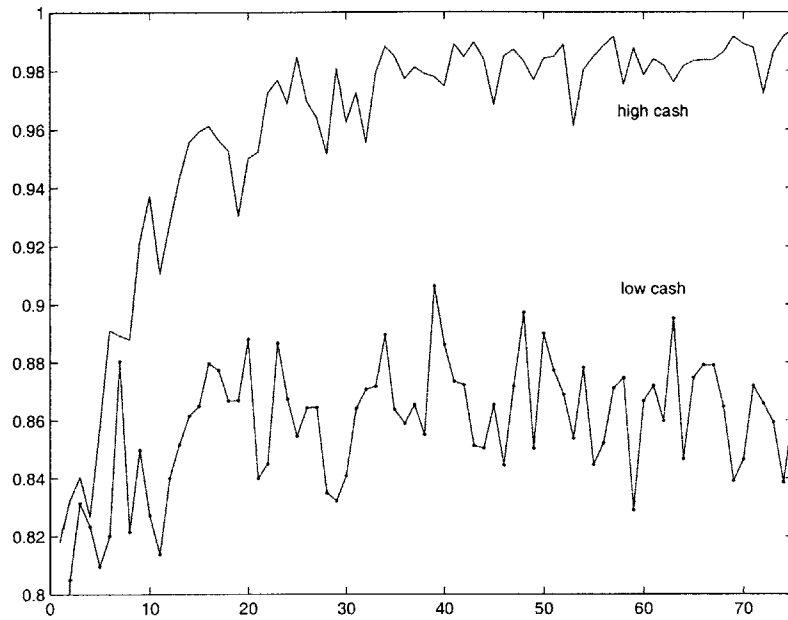


Figure 2-5d: Allocative efficiency, averaged over 100 repetitions of each of two runs of Artificial Markets Experiment 2.4.3 (information aggregation with heterogeneous preferences), the ‘low-cash’ and ‘high-cash’ experiments.

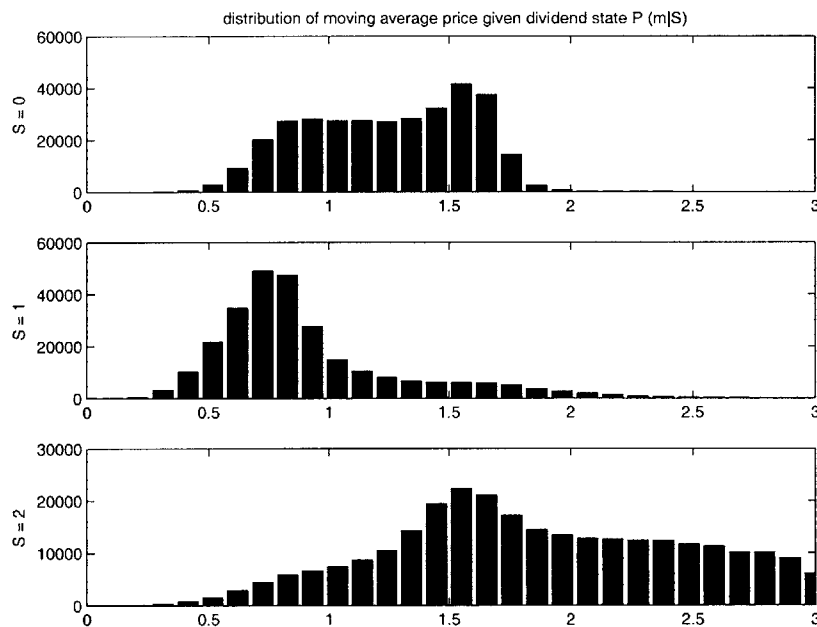


Figure 2-5e: Empirical distribution of moving-average prices, conditioned on the state of nature S , in Artificial Markets Experiment 2.4.3 (information aggregation with heterogeneous preferences).

of insiders, it is only possible for type B to buy from type A. The uninformed agents, without any private information, will have a reservation price approximately equal to 1 regardless of their dividend profile.⁶ Hence, we can conjecture that the transaction prices will range from 1 to 3. Note that type B insiders will bid the highest price—close to 3—and they will never sell the shares. The rest will attempt to buy or sell at roughly 1 but type B insiders will be responsible for most of the buying. Consequently supply diminishes and the price converges gradually to 2.

Not surprisingly, we also observe close to 100% allocative efficiency in the high-cash experiment as Figure 2-5d shows. However, the large bid-ask spreads displayed in Figure 2-5b imply that many traders are still interested in trading at prices far from the REE price, and there is little improvement in this spread across the periods.

Information dissemination in a market with diverse dividends (Experiment 2.4.4) is unsuccessful by our learning agents. This contrasts sharply with the human-based experimental markets studied by Plott and Sunder (1982), where after a few trials, insiders begin to realize that the equilibrium price can be different from what their dividend profiles imply, and they adjust their trading strategy accordingly. Uninformed human traders are also able to infer the equilibrium price from market conditions. The key distinctions between these experimental markets and our simulations are human traders' knowledge of the existence of heterogeneous preferences (diverse dividend payoffs), and their ability to learn the relation between the equilibrium price and the state of nature.

In the market of diverse information and heterogeneous preferences (Experiment 2.4.3), the end-of-period price does not come close to the REE price. We recognize that a market with diverse information is a more difficult scenario than one with insider information. In similar experiments with human subjects, Plott and Sunder (1988) show that information aggregation was unsuccessful in a market with heterogeneous preferences, and they attributed the failure to the complexity involved in inferring the state from market information. In two other sets of experiments, they

⁶This is approximate because their beliefs, conditioned on the market prices, can affect their estimates of the price.

found that the market aggregates information efficiently by having identical dividends across all traders (as in Experiment 2.4.1), or by replacing the single three-state security with three state-contingent claims. In a separate study, Forsythe & Lundholm (1990) confirmed similar results and added that information aggregation can be successful if the information about the heterogeneity in dividend payoffs is made available to all traders. Nevertheless, here our empirical Bayesian traders fail to aggregate information for the same reasons as they fail to disseminate information in Experiment 2.4.4.

2.5.3 Momentum Traders

In Experiment 2.4.5, we add momentum traders to the market to introduce extra noise and volatility to the “signal” perceived by the partially informed empirical Bayesian traders. To quantify the effect that momentum traders have on the market, we plot in Figure 2-6a the average absolute price-deviations in periods 30, 40, 50, and 75, each averaged over 100 trials for each of 14 different runs of this experiment, each run corresponding to a different number of momentum traders, from 0 in run 1 to 150 in run 14 (the number of empirical Bayesian traders is fixed at 20 for all runs).⁷ As expected, the absolute price-deviation curve is highest for the period-30 plot and lowest for the period-75 plot—the market becomes more efficient over time as agents learn.

Figure 2-6a also shows that in all four periods, the absolute price-deviations decrease initially as momentum traders are introduced, but generally increase after the number of momentum traders exceeds 5. Momentum traders add not only noise but also liquidity to the market, and with a small population of these irrational agents in the market, the empirical Bayesians manage to take advantage of the additional liquidity in making the market more efficient. However, when the number of momentum traders reaches 25 or more, the average price-deviation exceeds that of the benchmark case where no momentum traders are present.

⁷See footnote 5.

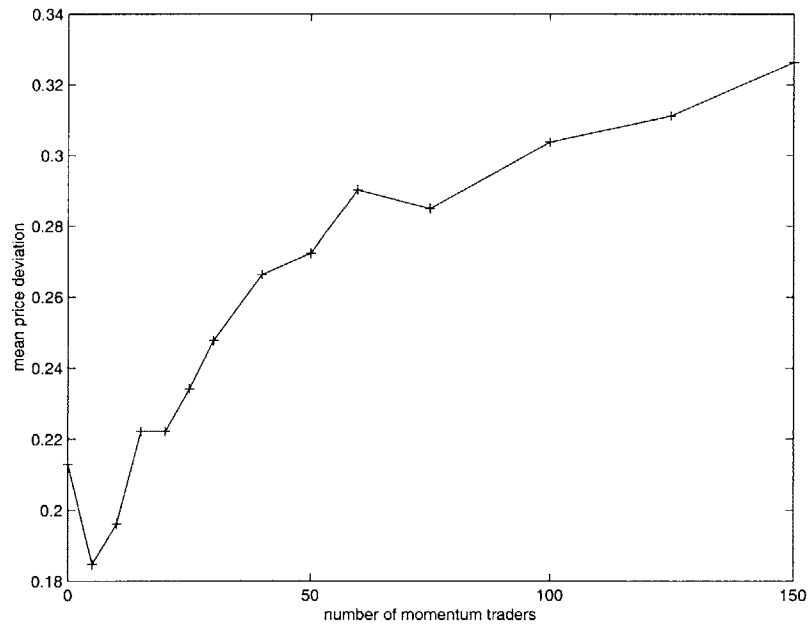


Figure 2-6a: Absolute price-deviations of market prices from the rational expectations equilibrium price in periods 30, 40, 50 and 75, averaged over 100 repetitions, as a function of the number of momentum traders present in Artificial Markets Experiment 2.4.5 (information aggregation with empirical Bayesian and momentum traders).

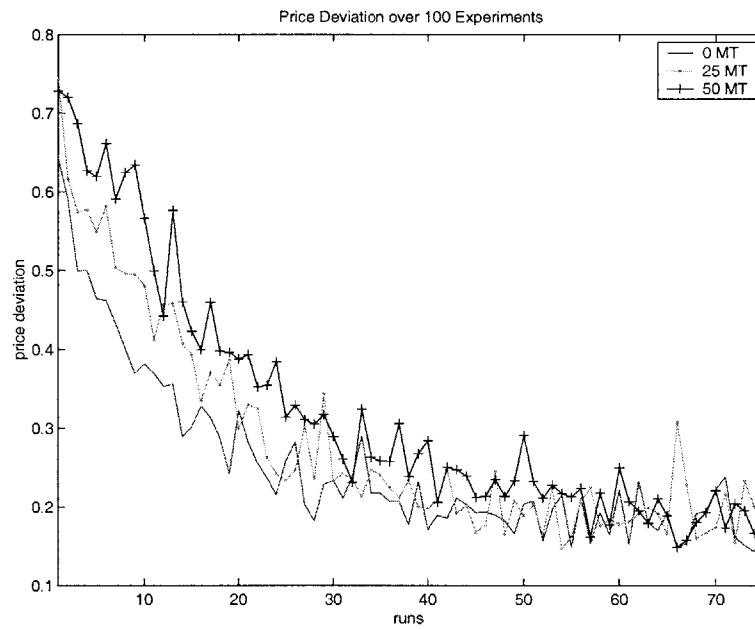


Figure 2-6b: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions, over the epoch for 0, 25, and 50 momentum traders in Artificial Markets Experiment 2.4.5 (information aggregation with empirical Bayesian and momentum traders).

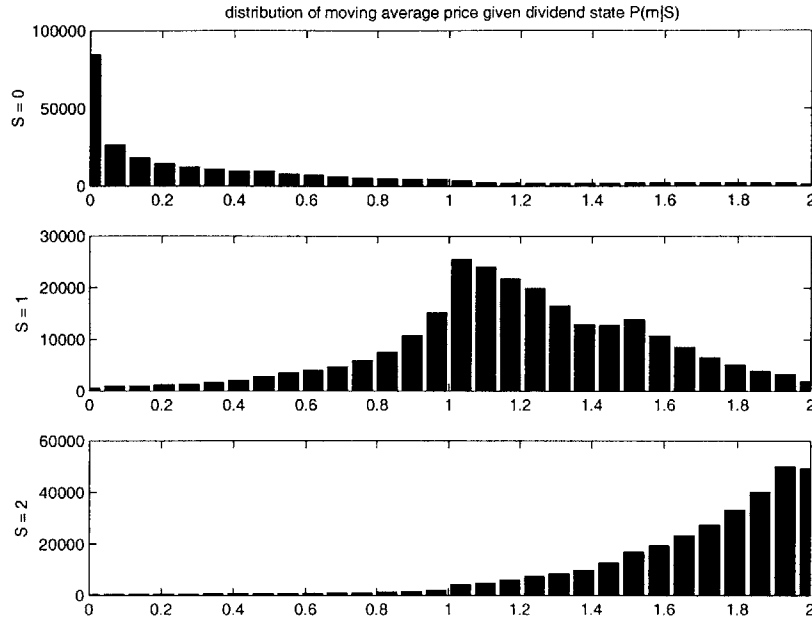


Figure 2-6c: Empirical distribution of moving-average prices, conditioned on the state of nature S , in Artificial Markets Experiment 2.4.5 (information aggregation with 20 empirical Bayesian and 20 momentum traders).

Figure 2-6b provides a more detailed look at the impact of momentum traders on market prices through plots of the average absolute price-deviation over three different runs of Experiment 2.4.5: runs with 0, 25, and 50 momentum traders (each plot is the average over 100 trials). Not surprisingly, the average absolute price-deviations increase with the number of momentum traders. The irrational trading of the momentum traders adversely affects the price convergence at early stage of the markets (roughly from periods 1 to 40). However, as the empirical Bayesian traders learn from and adapt to the strategies of the momentum traders, they are eventually able to overcome the noise from the irrational trading. From periods 65 to 75, the three markets are about equally efficient as measured by price deviation. Both the learning of the empirical Bayesian traders and the liquidity provided by the momentum traders contribute to efficiency of the markets.

Figure 2-6c plots the empirical conditional distributions of the moving-average prices in an experiment with 20 empirical Bayesians and 20 momentum traders. De-

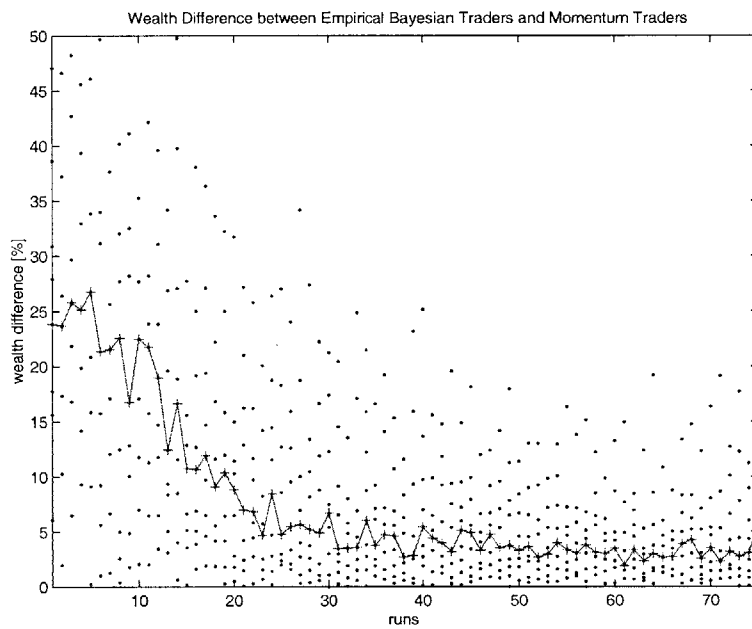


Figure 2-6d: Deciles of percentage wealth differences between empirical Bayesian and momentum traders in 100 repetitions of Artificial Markets Experiment 2.4.5 (information aggregation with 20 empirical Bayesian and 20 momentum traders). Medians are indicated by the symbol '+'.

spite the fact that these distributions have high dispersion, the three states are still distinguishable. In such circumstances, we expect the empirical Bayesians to exploit the irrational momentum traders and end up with a much higher level of end-of-period wealth. After all, the farther the price deviates from the RE price, the higher is the gain of the empirical Bayesians. However, Figure 2-6d shows that this intuition is not complete. Although median wealth differences between empirical Bayesian and momentum traders increase initially (from periods 1 to 5), they generally decline afterwards. The initial increase can be attributed to the empirical Bayesians' learning about the influence of momentum traders. But after some point, the market becomes more efficient, i.e., prices become more informative and closer to the REE. This is consistent with the patterns documented in Figure 2-6b—the initial advantages of the empirical Bayesians diminish through time as profit opportunities are bid away.

2.5.4 Nearest-Neighbor Traders

In the previous experiments, we have shown that the empirical Bayesian traders are successful in disseminating and aggregating information in homogeneous-preferences cases. However, we have not investigated the weak-form efficiency of these markets, i.e., how predictable are price changes? In Experiment 2.4.6, the empirical Bayesian traders are combined with nearest-neighbor traders, traders that attempt to uncover and exploit predictabilities in past prices. Our hypothesis is that if market prices are informationally efficient and do fully reveal all available information, then nearest-neighbor traders will perform poorly against empirical Bayesians.

Figure 2-7a shows the price convergence of this market. The price deviations reach the same levels as those in Experiment 2.4.1, converging rapidly after 50 periods.⁸ Unlike momentum traders, nearest-neighbor traders do not appear to hinder the process of information aggregation.

With respect to the relative performance of the two types of traders, Figure 2-7b shows that in terms of median percentage wealth differences, nearest-neighbor

⁸For this experiment, we extend the epoch to include 100 periods to ensure that prices were converging to the REE instead of cycling.

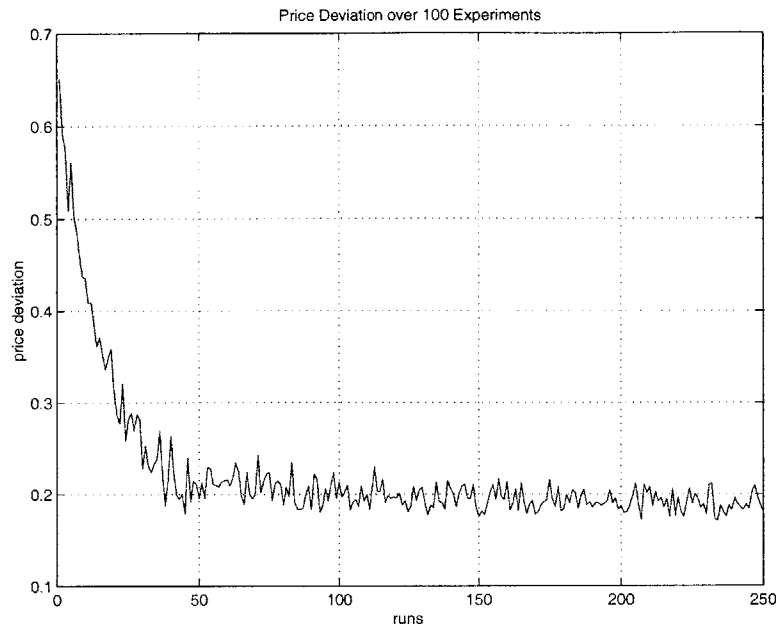


Figure 2-7a: Absolute price-deviations of market prices from the rational expectations equilibrium price, averaged over 100 repetitions, in Artificial Markets Experiment 2.4.6 (information aggregation with 15 empirical Bayesian and 5 nearest-neighbor traders).

traders outperform empirical Bayesian traders in all but the first 4 periods, implying that market prices do have some predictability to be exploited. In fact, the nearest-neighbor traders significantly outperform the empirical Bayesians roughly from period 5 to 40, after which the median wealth difference between the two groups becomes less significant.

This suggests that the predictability in prices is temporary (but more than just a few periods), and that nearest-neighbor traders learn faster than the empirical Bayesians. The first implication is consistent with our observation of the decreasing price deviations (in Figure 2-7a), or equivalently increasing price efficiency, from periods 1 to 40. Nearest-neighbor traders help make the market more efficient.

With respect to the second implication, the two types of traders start learning at the same time and compete with each other to discover the REE price. Evidently, the nearest-neighbor traders are able to exploit predictabilities more quickly hence they outperform empirical Bayesians initially. Eventually, empirical Bayesians are

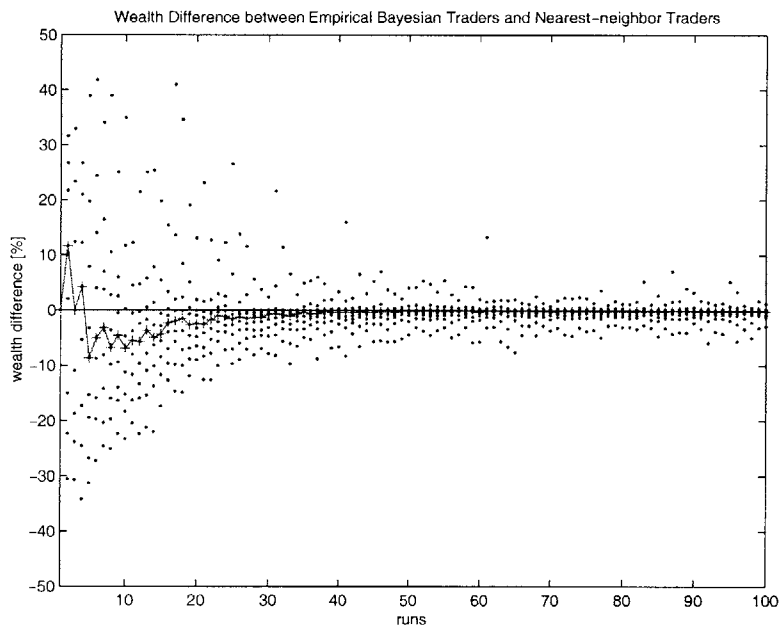


Figure 2-7b: Deciles of percentage wealth differences between empirical Bayesian and nearest-neighbor traders in 100 repetitions of Artificial Markets Experiment 2.4.6 (information aggregation with 15 empirical Bayesian and 5 momentum traders). Medians are indicated by the symbol ‘+’.

able to adapt to the strategy of the nearest-neighbor traders and more accurately infer the state of nature from market data. Consequently, as price becomes more efficient, the advantage enjoyed by the nearest-neighbor traders diminishes. However, the distribution of wealth differences (Figure 2-7b) show that even in the later periods, there are some realizations in which nearest-neighbor traders exhibit small gains over empirical Bayesians. These gains are not caused by price inefficiency, but are due to the fact that empirical Bayesians trade on rather inaccurate unconditional expected dividend at the beginning of each period.⁹

2.6 Conclusions

The rich implications of our agent-based model of financial markets underscore the potential for this new approach to shed light on challenging financial issues that currently cannot be addressed in any other way. Our simulation results accord well with human-based experimental market studies in many cases. Our simple AI-agents can accurately infer and aggregate diverse pieces of information in many circumstances, and they have difficulties in cases where human traders are also unable to determine the rational expectations equilibrium.

In a small number of cases our markets behave differently from human-based experimental markets. In our view, these discrepancies are just as significant as the concordances. For example, the sharp contradiction between Plott and Sunder (1982) and our experimental results in the case of information dissemination under heterogeneous preferences points to several important issues that warrant further investigation (more sophisticated learning algorithms for our agents, non-price learning and communication by human subjects, the dynamics created by heterogeneous preferences, etc.).

The use of AI-agents with simple heuristic trading rules and learning algorithms allows us to perform many new experiments that are well beyond the capabilities

⁹Recall from Section 2.3.4, empirical Bayesian traders compute their expected dividend condition on a k -period moving average price m_t . At the beginning of each period, before k prices are available, they simply trade on unconditional expected dividend.

of experimental markets with human subjects. For example, we show that adding momentum traders to a population of empirical Bayesians has an adverse impact on market performance and the momentum traders do poorly overall. However, this effect diminishes over time as the market becomes more efficient. But in our final experiments in which nearest-neighbor traders—traders that simply trade on patterns in past prices—are added to a population of empirical Bayesians, they are relatively successful free riders, not only matching the performance of empirical Bayesians in the long run, but outperforming the Bayesians in the short run. We conjecture that this advantage comes from the nearest-neighbor traders' ability to exploit short-term predictabilities more efficiently (that is what they are designed to do), and such predictabilities are more readily available in the early periods of trading.¹⁰ These findings raise interesting possibilities when viewed from an evolutionary point of view. In the early periods, selective pressures favor the nearest-neighbor traders, not the empirical Bayesian. If enough free riders enter the market, then prices might fail to converge to the rational expectations price because the market will contain too many free riders hoping to learn from price patterns alone.

¹⁰These results are closely related to parasite strategies documented in Rust et al. (1992).

Chapter 3

Experimental Markets for Product Concepts

Markets are well-known as an efficient tool to collect and aggregate diverse information regarding the value of commodities and assets. They have been particularly successful in the domain of financial assets. The previous chapter demonstrates how information aggregation and dissemination in financial markets can be modeled in an agent-based framework. This section continues the study of markets through both simulations with artificial agents and experiments with human subjects. Specifically, we explore an alternative application of the market mechanism to marketing research—using markets to collect consumer preferences on virtual product concepts. This application of markets is motivated by the desire to seek alternative reliable and accurate means to collect consumer preferences, and the belief that markets are efficient in aggregating information.

This chapter presents the results of three market experiments in which participants express their preferences on some new product concepts by trading virtual securities. The preferences collected are compared and studied together with results derived from a separate study of the same problem using a survey method. We find that results across different market experiments are consistent with each other, and in addition, highly correlated with those from the independent survey method. To gain a better understanding how markets work in this particular application, we relate our market

experiments with some classic examples from the experimental economics literature. Lastly we propose some conjectures on issues including how equilibrium prices are related to individual participants' preferences, how participants gather and process information, and what possible trading strategies are.

3.1 Introduction

The essence of the markets for product concepts centers around the establishment of virtual stock markets that trade virtual securities, each associated with an underlying product or service. These products or services could be a concept or prototype under evaluation or an existing product that anchors the market to the real world. Upon entering a concept market, each participant receives an initial portfolio of cash (virtual or real) and virtual stocks. Participants are also provided with detailed information on the products (stocks) that includes specifications, images, and multimedia illustrations. The objective of the market game is to maximize the value of the portfolio, evaluated at the market closing price. If participants play with real money, they will have the opportunity to profit from trading and bear the risk of losing money. The financial stakes in the game provide incentives to reveal true preferences, process information and conduct research. If fictitious money is used, prizes can be awarded according to individuals' performance. One can also reward all participants just for their service.

As in financial markets, the prices of the stocks are determined by the demand and supply in the market, which depend on the participants' evaluation of their own *and* others' preferences of the underlying products. Thus at the market equilibrium, prices should reflect all participants' aggregate preference of the products. Traders make trading decisions just as they would in a financial stock market: they assess the values of the stocks, sell the overvalued and buy the undervalued. Traders buy and sell virtual stocks, essentially voting on the worth of the products. In this way, a stock's price becomes a convenient index of a product's consumer value.

Surveys, polls, and focus groups are the traditional methods to collect such infor-

mation. Concept markets can serve as an alternative means to the same tasks or as a compliment to other methods. There are multiple reasons why the market mechanism can add value to the collection of diverse consumer preferences:

- *Accuracy*: Market participants have more incentives to trade according to the best of their knowledge because of their financial stakes in the market. The market can also capture, continuously, the ever changing “consumer impulse” for all participants who can express their opinions multiple times during the course of the market. Furthermore markets allow for dynamic movement of a virtual product’s worth or price along a slide scale, rather than attempting to collect yes/no survey answers when consumer’s real sentiment lies somewhere in between.
- *Learning and Interaction*: A concept market participant is not only evaluating on behalf of himself or herself, but also considering the opinion of the public at large. Furthermore, participants can observe others’ valuations of the virtual products and adjust their own dynamically in the market environment. Learning is an important element in these markets.
- *Scalability*: Unlike surveys, markets are intrinsically highly scalable. In fact, the efficiency of the market, and therefore the quality of data collected, improves with the number of participants.
- *Ambiguous and Intangible Attributes*: The market method is particularly useful over survey methods when a product cannot be naturally described or represented by a set of attributes (for example, a movie script). Market participants would evaluate the concepts directly and market prices would effectively reflect the overall viability of the concepts.

However, we also recognize limitations of the market method. Unlike typical marketing research techniques, in which information is collected from individuals and aggregated in subsequent analysis, the market method focuses on aggregate beliefs

and preferences and neglects those of individuals'. Virtual concepts markets are vulnerable to price manipulations and speculative bubbles because the values of virtual securities hinge on the aggregate beliefs, which are endogenously determined within the same market. Traders may form false beliefs that could cause prices to deviate from their fundamentals. For these reasons, the market method must be applied with cautions and consistency of the results must be checked.

The rest of the chapter is organized as follows. Section 3.2 provides the background for survey markets, which includes an overview of similar opinion-collecting markets on the web, a description of a concept testing project that this study is based on, and relevant research in experimental economics. Section 3.3 presents the designs of the securities and markets. 3.4 through 3.6 provide conjectures on how the survey markets work by considering an equilibrium model and simulations with artificial agents. Lastly, section 3.7 presents results from three market experiments.

3.2 Background

3.2.1 Opinion-collecting Electronic Markets

The application of the market mechanism is not restricted to the pricing of assets in financial markets. Different non-financial markets have been established for opinion polling, forecasts and predictions. The Iowa Electronic Markets (IEM)¹ from the University of Iowa is one of the pioneers of non-financial markets in the polling of opinions (Forsythe, Nelson, Neumann & Wright (1993)). The IEM was founded for research and educational purposes. Trading profits from the market provide incentives for traders to collect and process information of relevant future events. The IEM features real-money futures markets in which contract payoffs depend on the outcome of future political and economic events. Examples of these markets are the U.S. Presidential Election Market and the Computer Industry Returns Markets. In the U.S. Presidential Election Vote Share Market, for example, the contract *Rep VS*

¹The Iowa Electronic Markets <http://www.biz.uiowa.edu/iem/>

would pay \$1.00 times the vote share (percentage of popular vote) received by the Republican Party nominee, George W. Bush, in the November 2000 election. Traders at IEM invest their own funds, buy and sell listed contracts according to their own judgment of the likelihood of the underlying events, which is equivalent to the values of the corresponding contracts. On the election day, the contract *RepVS* (Bush) was liquidated at \$0.497 while the contract *DemVS* (the Democratic Party nominee, Al Gore) was liquidated at \$0.499, indicating that the overall market "thought" that Bush and Gore would receive 49.7% and 49.9% of the popular votes respectively.² IEM predicted the voting results, in terms of popular votes, of the past two Presidential elections within two-tenths of a percentage point, outperforming most national polls.³ The Gallup Poll's predictions, for example, deviated from election results by 1.9% and 5.7% for the Democratic candidates in years 1992 and 1996 respectively.⁴

With a similar idea, the Hollywood Stock Exchange or HSX⁵ establishes virtual markets trading movie stocks and star bonds. Each share of a movie stock pays a percentage of a particular movie's U.S. box office total; a star bond is priced based on a movie star's performance at the box office of his or her recently released movies. Prices of these stocks and bonds are determined by the demand and supply in the market, which in turn depend on players' consensus. Players trade with fictitious money called "Hollywood dollar." Market prices serve as predictions on earnings of movies and consensus of movie stars' popularity. Box office forecast is an invaluable service to film makers and Hollywood marketers. But traditional marketing research techniques have been found notoriously inaccurate and unreliable.⁶ HSX markets provide an alternative means of obtaining such forecasts in an arbitrarily large scale inexpensively.

²Bush and Gore received 47.87% and 48.38% of the popular vote respectively.

³*Business Week*, 11/11/96.

⁴<http://www.gallup.com/Election2000/historicalsummary.htm>

⁵The Hollywood Stock Exchange <http://www.hsx.com>

⁶*Wall Street Journal* July 21, 2000. Box Office Muddle: How Will Films Fare From Week to Week.

3.2.2 Virtual Concept Testing

Our virtual markets are set up to collect consumer preferences to facilitate product concept testing, a critical step in new product development. Concept testing is a procedure to narrow down multiple design concepts to the "optimal" design according to the responses collected from potential users of the product or service being studied. Dahan & Srinivansan (2000) presents a methodology that conducts product concept testing over the Internet using virtual prototypes in place of real, physical ones. The authors consider the World Wide Web as an attractive environment for conducting marketing research because of its interactive nature, instantaneous access to respondents, and availability of new technologies to deliver rich multimedia contents. The authors claim that the use of virtual product reduces costs in new product development so that a larger number of concepts can be explored. In their Internet-based approach, virtual prototypes are presented in the form of visual static illustrations and animations, plus a virtual shopping experience. Through an interactive Web page, respondents are able to rank different products by specifying the prices they are willing pay for individual products. After data are collected from respondents, conjoint analysis is conducted to obtain market share predictions of the concept products. Conjoint analysis is a technique used to decompose respondents' preference on individual attributes of a product based on their overall preferences. Green & Wind (1981) provides a tutorial of the technique. The goal of the study is to choose the best design of a bicycle bump among nine concept products and two commercially available products. The authors find that the virtual prototype tests produce market share predictions that closely resemble to those given by tests in which real physical prototypes are used.

Following this successful use of virtual prototypes, we adopt the virtual illustrations and animations from Dahan & Srinivansan (2000) to study the same problem. But instead of conducting surveys, virtual markets are set up.

3.2.3 Rational Expectation Models and Experimental Markets

Our trading experiments are closely related to the literatures in rational expectations model and experimental markets. A typical rational expectations (RE) model (Grossman & Stiglitz (1980)) studies a market of agents with diverse information. Under certain conditions, the competitive equilibrium prices will reveal all relevant information. The most important criterion for convergence is that agents condition their beliefs on market information. In particular, agents make inferences from market information about other agents' private information. The RE model has received considerable attention in the study of experimental markets. Studies of the informational efficiency of a market relative to the RE benchmark fall into two categories: *information dissemination* between fully informed agents ("insiders") and uninformed agents, and *information aggregation* among many partially informed agents. Plott & Sunder (1982), Plott & Sunder (1988), Forsythe et al. (1982) and O'Brien & Srivastava (1991) show that both information aggregation and dissemination occur successfully—markets attain the rational expectations equilibrium—in various experimental markets with human subjects.

Our trading experiments share some common characteristics with experimental markets. Both information aggregation and dissemination can offer explanations on the underlying activities in our virtual markets. One possible scenario is that there may be traders who possess superior information about the products or have high confidence on their beliefs. These can be considered as "insiders." On the other hand, traders who have little knowledge or opinion of the products can be regarded as the "uninformed." The interaction between the insider and uninformed constitutes a scenario of information dissemination. What is intriguing about this scenario is that even with the presence of traders, who may ignore the underlying product information and only focus on market information, the market could still converge to efficient prices that reflect all the relevant information or beliefs. Another possibility is that individual traders form their own beliefs about the products, and realize that market

prices would depend on the aggregate beliefs. This is similar to the information aggregation scenario where traders are partially informed.

However, there are also significant differences between our product concept markets and those in the experimental markets literature. In a typical experimental market, subjects' preferences and their information set are fixed and assigned by the researchers. Therefore, even before trading begins, theoretical equilibrium prices can be calculated. In contrast, in concept market experiments, neither subjects' preferences nor their information sets are known. In fact, these are what experiments aim to discover.

3.3 Design of Markets and Securities

The market method is applied to the same product concept testing problem presented in Dahan & Srinivansan (2000). Three trading experiments were conducted to predict the market share of nine concept bike pumps. A market is set up with eleven securities—nine concept products and two commercially available products—corresponding to the eleven bike pumps in Dahan & Srinivansan (2000). Each of the securities is the stock of the virtual company that manufactures and sells a particular pump as its only product. These pump-manufacturing companies will go public, and their initial public offering (IPO) prices are to be determined in our virtual market. All companies will offer the same number of shares of common stock to the public. The objective of the game is to maximize the value of one's portfolio at market close. The value of a portfolio is calculated as the sum of the cash and total worth of the stocks, which are valued at the market *closing prices*. Participants should strive to maximize their profits by trading the stocks using their personal valuation of the companies, as well as any information they can observe from the trading dynamics. Fictitious money was used in the markets, but we rewarded top players with prizes. This provides the participants an incentive to perform in the experiments. It is assumed that all the eleven companies have identical production cost structures, manufacturing capacity, distribution channels, financial structures, management ex-

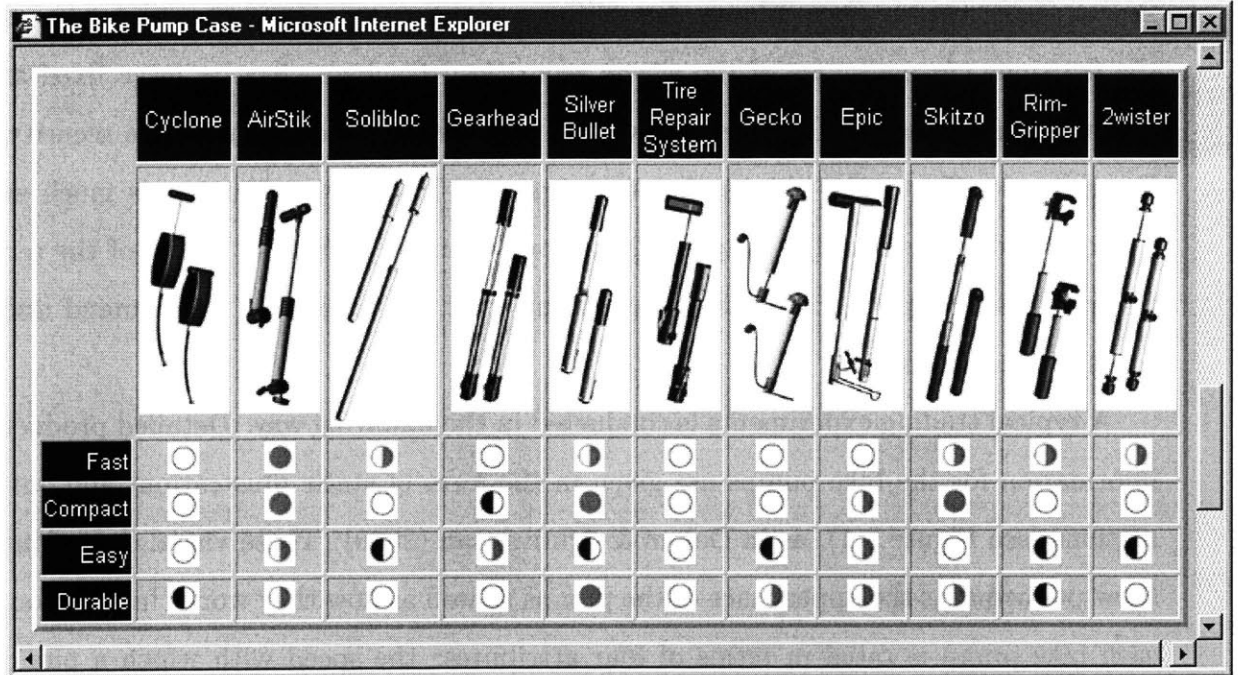
pertise and all other factors that may affect their profitability. In other words, all factors other than the quality and desirability of the pumps can be ignored in valuing the stocks.

The eleven bike pump companies are Cyclone, AirStik, Soliboc, Gearhead, Silver Bullet, TRS, Gecko, Epic, Skitzo, RimGripper and 2wister. To anchor the value of the fictitious currency, one of the eleven securities—Cyclone—has its price fixed at \$10 and is not traded. Cyclone is served as a reference price or numeraire security. For example, if a trader thinks that the company TRS is worth twice as much as Cyclone, he or she would pay up to \$20 for one share of TRS. The stocks of the ten freely traded companies may be priced at any level, depending on the demand and supply in the market.

A typical trading experiments is conducted in the following way. Detailed product information for the bike pumps are given in the form of static illustrations and animations (see Figure 3-1), as in Dahan & Srinivansan (2000). These visual depictions show participants the appearance of the pumps as well as how they work. In addition, each bike pump is rated in terms of four attributes: the speed with which a pump inflates a tire, compactness, the ease of operation and durability. The participants are presented with a single Web page with the visual depictions and profiles of the bike pumps, trading instructions, and the objective of the trading game. They have ten minutes to study the material before trading begins.

All participants are provided with an identical portfolio that consists of \$10,000 of cash and 100 shares for each of the securities. No borrowing or short-selling is allowed in the market. At the end of the experiments, we reward the top three players with highest portfolio values with Amazon.com gift certificates of \$50, \$30 and \$20 respectively. Participants first log in with self-chosen user-names and passwords (see Figure 3-2a), then trade the securities through a graphical user interface (see Figure 3-2b). Market information available to the traders includes the last transaction price and size, current bid/ask prices and sizes, and a historical price and volume chart for each security. A trader can submit either a limit or market order to trade, or cancel an outstanding order that has not been executed. The markets are typical double

(a)



(b)

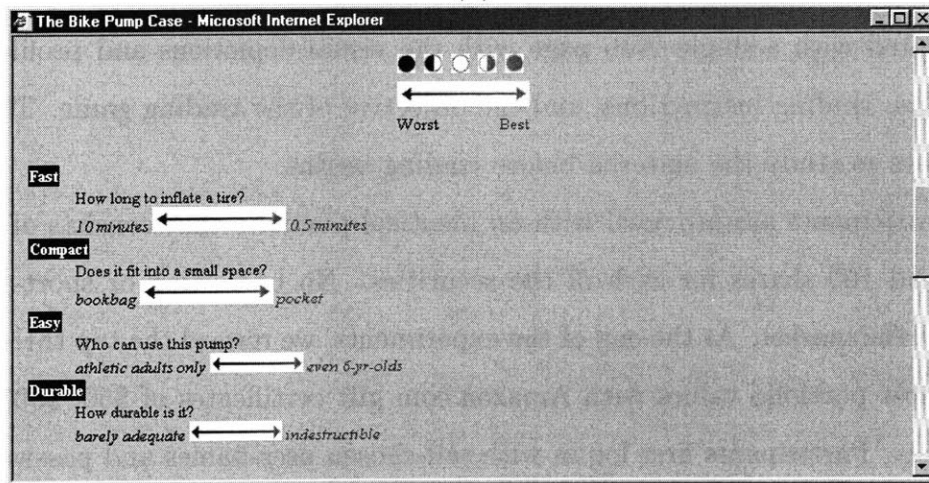


Figure 3-1: The profiles of the eleven bike pumps.

auction markets with no market-makers. A transaction occurs when a market or limit order matches with a limit order on the opposite side of the market. All prices are specified in one-sixteenth of a dollar.

3.4 A Theoretical Market Equilibrium

In this section, we attempt to provide an example of an equilibrium for a market that aggregates diverse beliefs of a group of agents. Consider a portfolio selection problem faced by an agent in a two-period economy with one risk-free and one risky security. The risk-free security yields no return. It serves as a numeraire security with price equal to 1. The payoff of the risky security depends on the state of the economy in period two. Let us consider the risk-free security as cash and the risky security as a stock. There are two possible states of the economy, a bad state and a good state. The stock pays 1 and 0 in the good state and bad state respectively. The futures contracts traded in the Iowa Electronic Market are examples of such a security.⁷ All agents in the economy have identical preferences and are provided with identical initial endowment. This is a typical portfolio selection problem in which agents choose the optimal allocation of cash and stock in period one in order to maximize their expected utility in period two. The essence of the problem is that they may have different beliefs about what will happen in period two.

Specifically, there are N agents in the economy, all possess an initial endowment of w unit of cash in period 1. Agent i believes that the probability of the good state is θ_i and that of the bad state is $1 - \theta_i$. It must choose x_i , the number of shares of the risky asset to own. The stock is traded in the market with a price p . The agent's budget constraint implies that it holds $(w - px_i)$ unit of cash. Consequently its period-two wealth is $\tilde{w}_i = (w - px_i) + \tilde{p}x_i$, where \tilde{p} is the period-two price for the risky security, which equals to 1 in the good state and 0 in the bad state. The agent

⁷For example, the Republican President contract will pay \$1 if the Republican nominee is elected president.

(a)

User login

User ID: MbMTrader password: *

New user? Please register here

User ID: MbMTrader password: * password again: *

(b)

MIT WebMarket: MbMTrader

File

Market News
22:32:38: Market is Open

Symbol: AirStik Size: Price:

Open Orders

Buy 75 shares of AirStik at \$27 1/8.
Buy 30 shares of AirStik at \$27 1/2.
Buy 10 shares of AirStik at \$27 3/4.

Transactions

You bought 35 shares of AirStik for \$27 1/2 per share.
You bought 5 shares of AirStik for \$27 3/4 per share.
You bought 20 shares of AirStik for \$26 per share.
You bought 50 shares of AirStik for \$26 1/2 per share.
You bought 50 shares of AirStik for \$26 7/8 per share.
You bought 10 shares of AirStik for \$26 1/2 per share.

Quote / Portfolio

	Last	Low	High	Bid (size)	Ask (size)	Volume	Quantity	Market Value
AirStik	27 3/4	26	28	27 3/4 (10)	28 (915)	475	455	12626.25
Solibloc	-	-	-	5 1/2 (50)	6 1/4 (60)	0	100	0.00
Gearhead	16 1/8	14 1/2	16 1/8	14 1/2 (450)	16 1/8 (50)	100	100	1612.50
SilverBullet	15 1/2	10	15 1/2	10 (400)	15 1/2 (1200)	300	100	1550.00
TRS	-	-	-	1/8 (70)	- (0)	0	100	0.00
Gecko	-	-	-	1 1/8 (1000)	5 (400)	0	100	0.00
Epic	-	-	-	11 1/8 (200)	12 (1000)	0	100	0.00
Skitzo	-	-	-	- (0)	- (0)	0	100	0.00
RimGripper	-	-	-	17 1/2 (800)	17 5/8 (1200)	0	100	0.00
Zwister	10 3/4	10 3/4	11	10 3/4 (100)	11 (300)	1000	100	1075.00
Cash								344.37
Total								17208.11

Warning: Applet Window

Figure 3-2: Login and trading interface.

maximizes its expected period-two utility subject to x_i , according to their beliefs:

$$\begin{aligned} \max_{x_i} \quad & E[U(\tilde{w})] \quad \Rightarrow \\ \max_{x_i} \quad & \theta_i U((w - px_i) + x_i) + (1 - \theta_i)U(w - px_i) . \end{aligned}$$

Now assume that all agents have a log utility function: $U(w) = \log(w)$. The demand of agent i can be derived from the first order condition of its optimization problem:

$$\begin{aligned} \theta_i \frac{1-p}{w+x_i(1-p)} + (1-\theta_i) \frac{-p}{w-x_i p} &= 0, \\ x_i &= \frac{w(\theta_i - p)}{p(1-p)} . \end{aligned}$$

Agents may hold a long or short position in the risky security, i.e. x_i may be positive or negative. At equilibrium, the total demand in the market must be equal to zero. The market price in period one is derived by imposing this market clearing condition:

$$\begin{aligned} \sum_{i=1}^N x_i &= 0 \quad \Leftrightarrow \\ p^* &= \frac{1}{N} \sum_{i=1}^N \theta_i . \end{aligned}$$

With this particular set-up of the problem, the market price is the average of individuals' beliefs. Note that other equilibria may result under different assumptions. Pennock & Wellman (1997) studies a similar problem with agents that have a negative exponential utility function, $U(w) = -e^{-cw}$, where c is the risk aversion coefficient. In its one-security case, the equilibrium price is found to be

$$p = \frac{\prod_{i=1}^N \theta_i^{c'_i}}{\prod_{i=1}^N \theta_i^{c'_i} + \prod_{i=1}^N (1 - \theta_i)^{c'_i}} ,$$

where c'_i is a function of an agent's risk aversion coefficient.

3.5 Possible Trading Strategies

Our market experiments serve to aggregate diverse preferences or beliefs from all participants. One's belief may consist of three independent elements:

- Product information — This is what a participant knows about the underlying products. All participants are provided with the same facts and specifications of the products, but they may have obtained extra product information from their personal experience outside the experiments.
- A participant's personal preference — This is what surveys and polls try to collect. Although the aggregate preferences of the whole market is the object of interest, one's personal view should contribute to his or her trading decisions.
- A participant's assessment of others' preferences — A participant would form an assessment of what others think so as to make profitable trading decisions.

How are beliefs or preferences aggregated in these markets with virtual securities? Not only should the traders form their own assessment of the stocks, but they should also infer the stocks' potential market value from the market. In a typical market in experimental economics, both the preferences of the traders and the state of nature (for example, probability distribution of a security payoff) are known to the researchers (Plott & Sunder (1982); Plott & Sunder (1988); Forsythe & Lundholm (1990); O'Brien & Srivastava (1991)). Traders are assigned with preferences that specify securities payoffs in various possible states. The theoretical equilibrium (rational expectations equilibrium) prices can be derived given full information of the markets. The main focus of these experiments is whether and under what conditions rational expectations equilibria can be attained in double auction markets. Some attempts have been made to understand the convergence of prices and how learning occurs in the market as a whole (see Chapter 2). But it is unclear how individual human traders learn and react to the market. Attempts to model the trading strategies of individual traders from the market data may be overly ambitious. Here we try to shed some light on some possible strategies.

The objective of the trading game is to predict the final prices of the securities. A trader may form an assessment of the fair values of the securities before trading starts. This opinion may take into account her own preference on the underlying products, and perhaps more importantly what she perceives as the preferences of the whole group. The trader may then make trading decisions based on her belief: she buys undervalued stocks and sells over-valued ones. During the course of the market, the trader may either maintain her valuations or update her beliefs in real time conditioning on her observation of the market dynamics. Learning is taking place if the latter approach is taken. But learning is a rather complex process because one's expectations of prices affect prices, prices are used to infer others' assessments, and the inference of others' assessments in turn affects both prices and expectations of prices.

Some traders may take a dramatically different approach by largely ignoring all fundamental information about the underlying products but focusing on market information only. These traders play the roles of speculators or market-makers who try to gain from the market by taking advantage of price volatility, providing liquidity or looking for arbitrage opportunities. Their presence may introduce mixed effects to the market. While they could enhance liquidity on one hand, they may also introduce speculative bubbles and excess volatility into the market.

3.6 Simulations with Artificial Agents

The equilibrium model presented in Section 3.4 exemplifies how a market equilibrium can be derived from individual beliefs. However, questions about how such equilibria are attained remain unanswered. Characterizations of the learning dynamics in an experimental market are typically very challenging. To gain some insights on the price dynamics and agent trading strategies, we turn to computer simulations. Similar to the presentation of the equilibrium model in Section 3.4, we aim to demonstrate one possibility of how a market can successfully aggregate diverse beliefs, but not to offer an explanation for the observations from our market experiments.

One way to specify an agent's behavior is to directly model its demand given the market price. Suppose the belief of agent i can be characterized by a "fair" price, r_i . For example, in the two-security economy in Section 3.4, for a risk-neutral agent⁸, the expected price given its belief can be considered as a fair price to the agent:

$$r_i = E[\tilde{p}|\theta_i] = \theta_i(1) + (1 - \theta_i)(0) = \theta_i \quad .$$

Given the fair price and market price, p , one can construct a linear demand function,

$$x_i(p) = \beta_i(r_i - p),$$

where β_i is the price elasticity controlling the sensitivity of demand to price deviation from the fair value. With this linear demand function, an agent would buy or sell β_i shares of the security for each unit of the price difference, $(r_i - p)$. Assuming all agents follow this linear demand function, imposing the market clearing condition, the equilibrium price is the average of all fair prices:

$$\begin{aligned} \sum_{i=1}^N x_i &= 0 \quad \Rightarrow \\ p^* &= \frac{1}{N} \sum_{i=1}^N r_i \quad . \end{aligned} \tag{3.1}$$

Depending on the form of the demand function, some other equilibria may exist. Figure 3-3a shows an example of a piece-wise linear demand function that describes the short term demand of a agent. The upward-sloping portions of the curve seem to defy the law of demand. The rationale behind such a demand function is that when the market price is within a certain range from the agent's fair price ($r-1 \leq p \leq r+1$), it buys if the market price is lower than its fair price and sells otherwise; when the market price deviates too much from its fair price, the agent may lose faith in its belief and begin to close its position: selling on decreasing price ($r-2 \leq p \leq r-1$)

⁸A risk-neutral agent has a utility function $U(w) = w$.

and buying on increasing price ($r + 1 \leq p \leq r + 2$). Given such a piece-wise demand function, *multiple* equilibrium prices may result. Consider the total demand from two agents with fair prices r_1 and r_2 . As shown in Figure 3-3b, the market clears in multiple price ranges: $p \leq r_1 - 2$, $p = r_1$, $r_1 + 2 \leq p \leq r_2 - 2$, $p = r_2$, and $p \geq r_2 + 2$.

Consider a simulated market in which there is one single security in the market. Each artificial agent forms a belief about the fair price (r_i) of the security at the beginning of a market, then follows a linear demand $x_i(p) = \beta_i(r_i - p)$. Following from Equation 3.1, we consider the equilibrium price $p^* = \frac{1}{N} \sum_{i=1}^N r_i$ as the benchmark price for a market with N artificial agents with fair prices r_i 's.

We further impose that an agent trades only one share in one transaction. From its demand function, one can derive that the agent would buy one share when $p \leq r_i - 1/\beta_i$ and sell one share when $p \geq r_i + 1/\beta_i$. Equivalently, one can consider the prices $b_i = r_i - 1/\beta_i$ and $a_i = r_i + 1/\beta_i$ as the agent's bid and ask prices respectively. All agents submit orders according to the procedure described in Table 3.1, given their bid and ask prices. All agents are provided with an equal amount of cash. No borrowing is permitted so the total demand for the stock is bounded. The trading mechanism is a simplified double-auction market. Agents can submit a bid and/or an ask, or accept a posted bid or ask. If there is an existing bid for the stock, any subsequent bid must be higher than the current bid. Similarly, on the sell side, a subsequent ask must be lower than the current ask. A transaction occurs when an existing bid or ask is accepted (a market order matches with a limit order).

The computer simulations feature two types of agents. One firmly believes in their beliefs and maintains them throughout the course of the market. Let us call these static agents. The other (adaptive agents) attempts to update their beliefs by making observations of the market. We consider a simple heuristic with which the adaptive agents "learn" from the consensus of the population. In particular, it updates its fair price using the observed market price p :

$$r_i \leftarrow r_i + \alpha_i(p - r_i) \quad ,$$

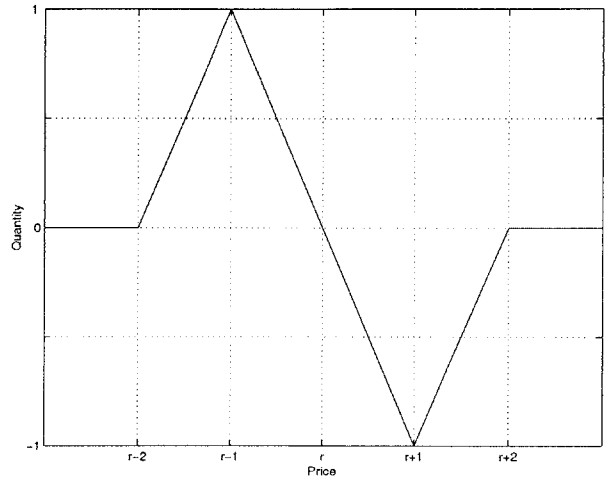


Figure 3-3a: An example of a non-linear demand function.

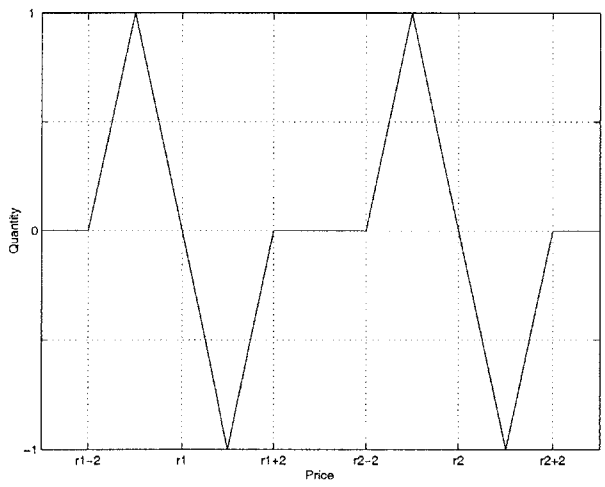


Figure 3-3b: The total demand in the market with two agents with fair prices r_1 and r_2 . the market clears in multiple price ranges: $p \leq r_1 - 2$, $p = r_1$, $r_1 + 2 \leq p \leq r_2 - 2$, $p = r_2$, and $p \geq r_2 + 2$.

Scenario	Action
	<i>existing bid, existing ask</i>
$b_i \geq a$	buy at market
$a_i \leq b$	sell at market
$b_i > b$	post a bid b_i
$a_i < a$	post an ask a_i
	<i>no bid, existing ask</i>
$b_i \geq a$	buy at market
$a_i > a$	post a bid b_i
otherwise	post a bid b_i and an ask a_i
	<i>existing bid, no ask</i>
$a_i \leq b$	sell at market
$b_i < b$	post an ask a_i
otherwise	post a bid b_i and an ask b_i
	<i>no bid, no ask</i>
all cases	post a bid b_i and an ask a_i

Table 3.1: The order submission algorithm of the trading agents, where $a_i = r_i + 1/\beta_i$ denotes the ask price by agent i , $b_i = r_i - 1/\beta_i$ the bid price by agent i , a the best ask price and b the best bid price in the market.

where α_i is the learning rate of the agent, dictating how rapidly belief updates take place. The rationale behind this updating rule is that if a large number of transactions occur at price p , r_i will converge to the market price p . If p is the population average price, this update rule will eventually make all fair prices converge to the population mean.

Do prices fully reflect all available information—individuals’ diverse beliefs on the values of the securities? In all simulations, we focus on the efficiency of the market in terms of whether or how close the price converges to the benchmark equilibrium as described in Equation 3.1. The role and impact of learning are the other issues that we are interested in.

To quantify the performance of the markets, we focus on three statistics of the price: the mean price, median price and closing price. Given a price statistic s_t^p for trial t , we compute its average price deviation from the benchmark equilibrium price in T trials of a market:

$$\Delta_{sp} = \frac{1}{T} \sum_{t=1}^T |s_t^p - \frac{1}{N} \sum_{i=1}^N r_i| .$$

In addition, we study the standard deviation of agents’ beliefs or fair prices at the end of the market session. This variable measures how much disagreement exists among all agents.

We conduct two computer simulations, each consisting of a series of markets. Each market is characterized by parameters such as the composition of the agents (static versus adaptive) and the learning rate of the agent population. We run each market 1000 times or epochs to get average measures of the market performance. A typical epoch of a simulated market consists of 2000 trading intervals. At each trading interval, an agent is chosen randomly from the population to submit an order. Before the beginning of a market, all agents are provided with 5 units of cash and 10 shares of stock. Agents’ beliefs are assigned in the form of a fair price r_i . The fair prices r_i ’s are drawn randomly from a uniform $[0, 1]$ distribution. The following describes the details of the simulations.

- **Simulation 1:** In this simulation, we study the effect of the learning rate on the performance of a market. There are 20 adaptive learning agents, each with an identical learning rate, in all markets. Each market differs by the learning rate of the agent population.
- **Simulation 2:** Markets in this simulation are characterized by the composition of the agent population. There are 20 static agents and 0 adaptive agent in market 1, 19 static agents and 1 adaptive agent in market 2 and so on. Eventually in market 21, there are 0 static agent and 20 adaptive agents. All adaptive agents have a learning rate of 0.01. By varying the composition of agents, we investigate the performance of the market with populations that have heterogenous learning capabilities.

Figure 3-4a shows typical realizations of four markets in Simulation 1, with the learning rate α equal to 0.00, 0.01, 0.05 and 0.10 respectively. One can observe that market price converges to the neighborhood of the equilibrium price for small learning rates. If the agents learn too fast, the market may converge to an arbitrary price, as in the case when $\alpha = 0.1$. As expected, the speed of convergence is proportional to the learning rate. Figure 3-4b shows the price deviations from the equilibrium price for the mean, median and closing price in markets with different α 's. Measured by the mean price, markets with low α values ($[0.005, 0.015]$) converge close to the equilibrium price. At higher learning rates, the median and close prices give relatively smaller price deviations. When there is no learning or when learning is slow, the mean price is a more robust measure for the consensus of the market. As α keeps increasing, convergence in the markets deteriorates on all three measures, implying that if the agents adapt to market consensus too rapidly, the overall market would diverge from the equilibrium. Another evidence of learning in the markets is shown in Figure 3-4c. The standard deviation of the agents' beliefs decreases with the learning rate, indicating that the difference in agents' beliefs diminishes as they approach a market consensus.

Results from Simulation 2 show that the market efficiency improves with an in-

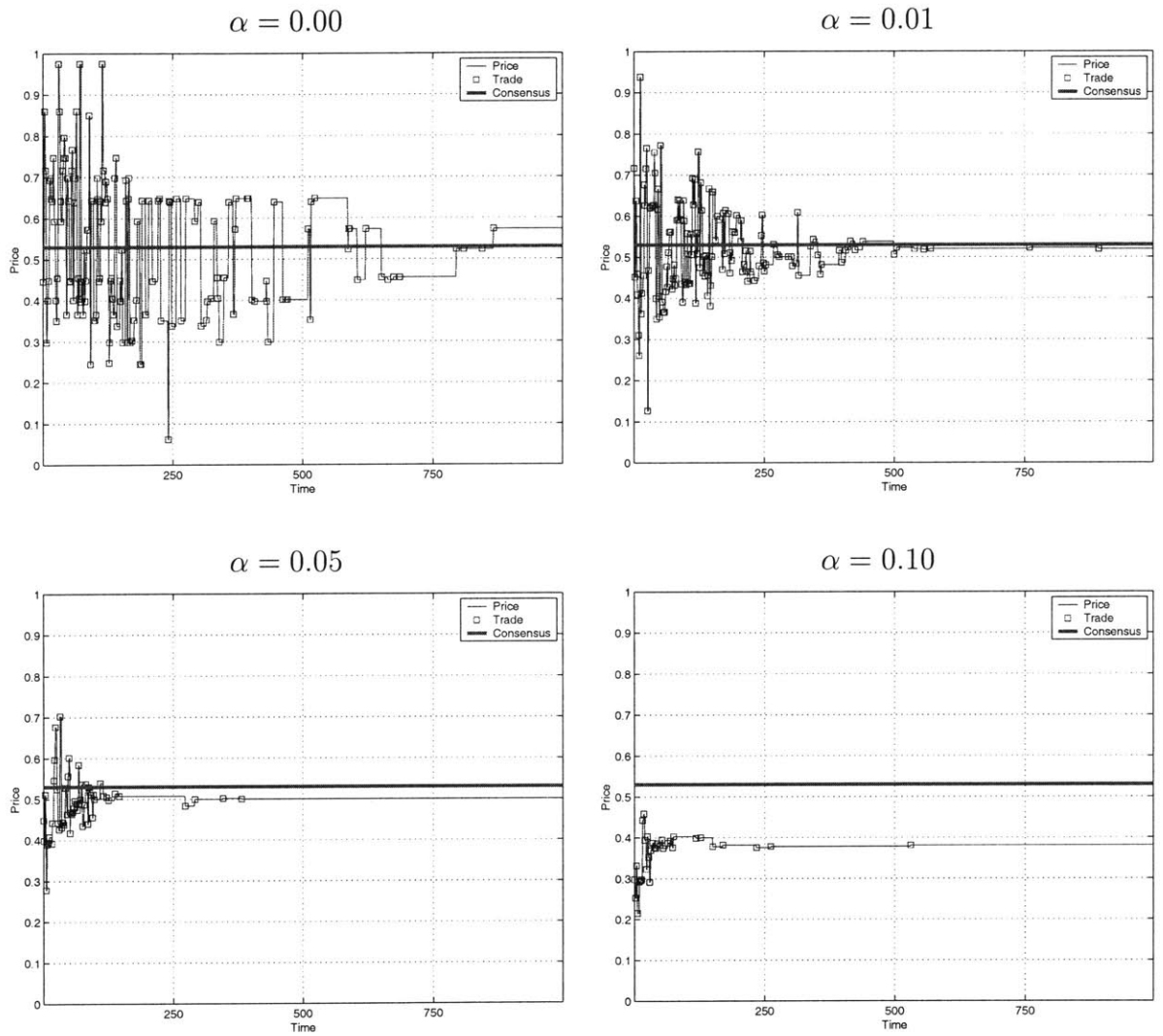


Figure 3-4a: Typical realizations of markets in Simulation 1 with the learning rate $\alpha = 0.00, 0.01, 0.05, 0.1$.

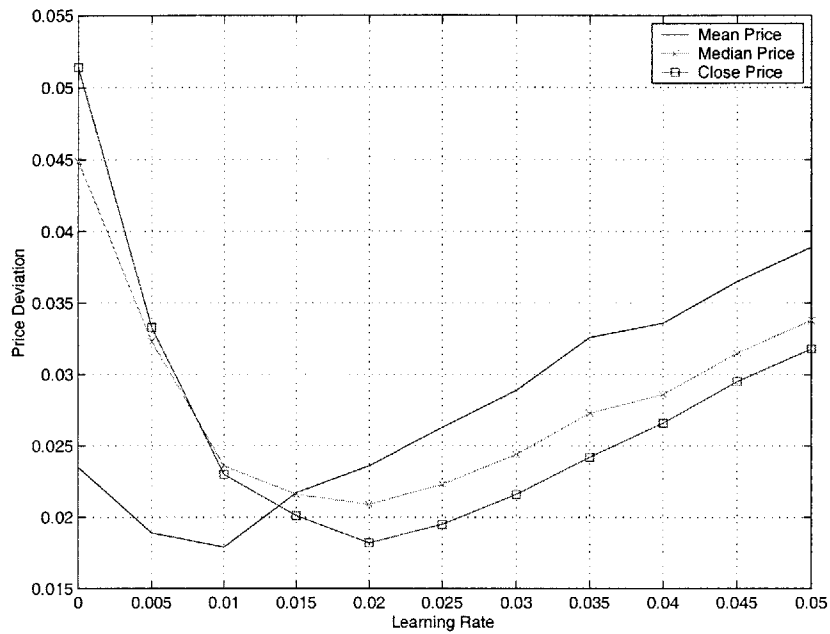


Figure 3-4b: Deviations of mean, median and closing prices from the equilibrium price for markets with agents with different learning rates in Simulation 1.

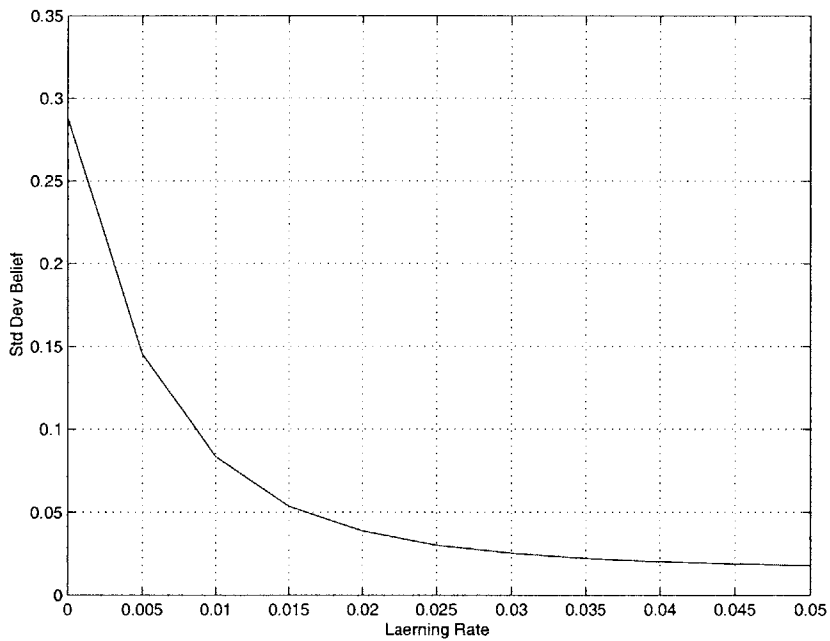


Figure 3-4c: Standard deviation of agents' beliefs at the end of the markets in Simulation 1.

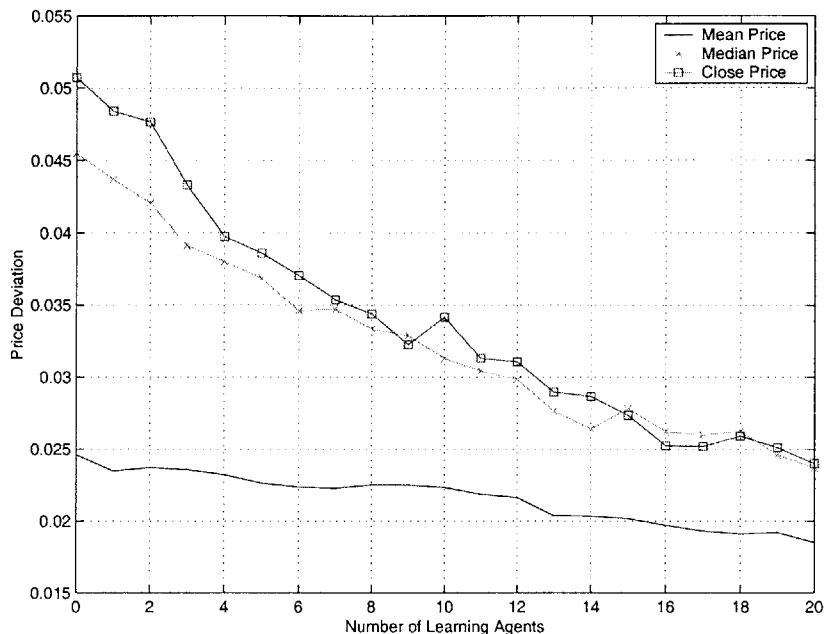


Figure 3-5a: Deviations of mean, median and close prices from the equilibrium price for markets with different numbers of adaptive learning agents in Simulation 2.

creasing number of adaptive agents, given an appropriate learning rate ($\alpha = 0.01$ in this case). Furthermore, the mean price is much more robust in measuring market consensus than the median and closing prices for markets with a heterogeneous agent population (see Figures 3-5a and 3-5b).

The study of agent simulations provides some insights into how a market aggregates diverse beliefs. The implications of these simulations are as follows: (1) learning in a market facilitates its convergence to the market consensus; (2) when learning occurs too rapidly, market price may diverge from the market consensus; (3) mean price is a more robust measure of the market consensus and the closing price does not always converge to an equilibrium price. However, one has to be cautious about the limitation of such studies. Significant assumptions have to be made about the form of beliefs, the functional form of demand functions, the trading strategies, and the updating of beliefs. For example, a demand function may not necessarily be in a simple linear form, and there may exist multiple equilibria in a single market; or agents may not always update their beliefs using a constant learning rate.

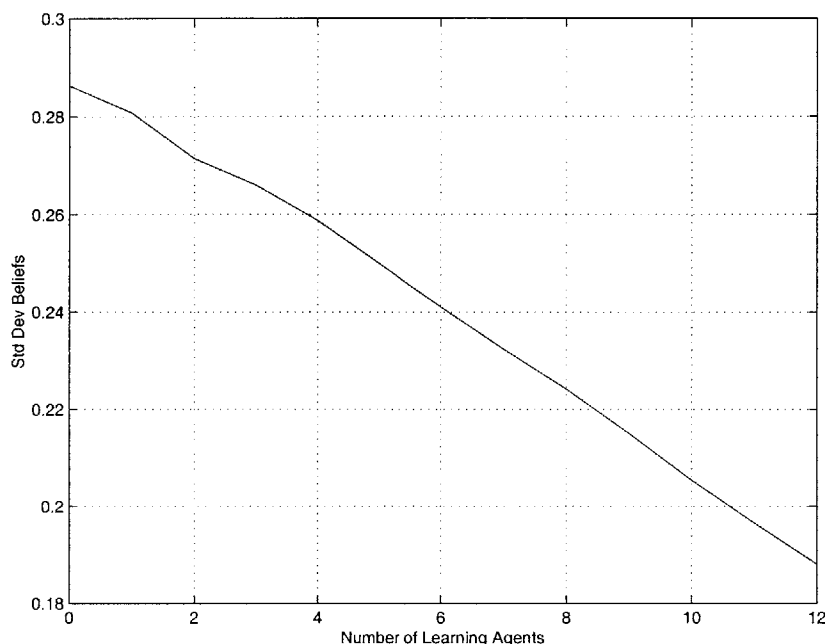


Figure 3-5b: Standard deviation of agents' beliefs at the end of the markets with different numbers of adaptive learning agents in Simulation 2.

3.7 Experimental Results and Comparison with Survey Study

Three trading experiments were conducted from September 1999 to April 2000. Subjects in the trading experiments are MBA students from MIT Sloan School of Management.⁹ Experiments 1 and 2 were conducted at a centralized location, the Sloan school trading laboratory. These two experiments were not timed—we closed the market when trading activities died down. They lasted 10 and 18 minutes. Experiment 3, on the other hand, was conducted over the Internet—participants joined the market remotely from arbitrary locations. The market was open for trading for one hour, during which participants might enter or leave the market as they wished. The longer trading time was aimed to offer more flexibility to the participants and investigate whether time constraints affect the trading activities of a market. There

⁹Three groups of students were recruited from Prof. Ely Dahan's class 15.828 *New Product Development* in fall 1999 and Prof. Andrew W. Lo's class 15.433 *Investment* in Spring 2000.

Experiment Number	Number of Participants	Duration (Minutes)	Volume (Shares)	Volume (Trades)	Venue
1	27	18	10,013	167	Trading Lab
2	26	10	6,150	79	Trading Lab
3	18	60	13,128	233	Internet

Table 3.2: Details of the three trading experiments.

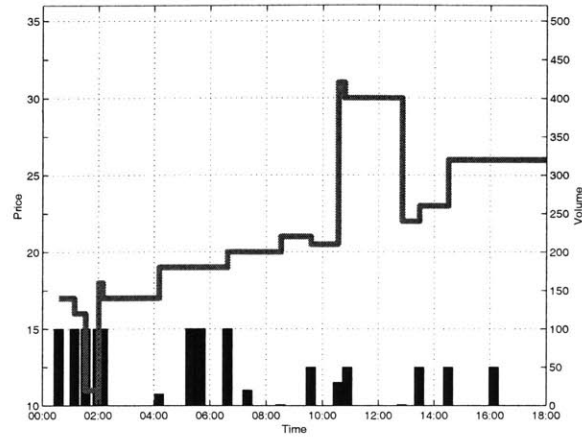
are 18 to 27 participants in each of the experiments. Table 3.2 describes the details of the experiments. Experiment 3 has the highest volume (in share and trade) despite a small number of participants. This seems to suggest that a longer duration does generate more trading interest. Figure 3-6 presents the sample price and volume history of a virtual stock—AirStik—in Experiments 1 to 3. The prices close around \$25 in the three experiments.

For each of the market experiments, trade and quote data is collected. For our analysis, we focus on the trade data, which consists of time series of transaction prices and quantities:

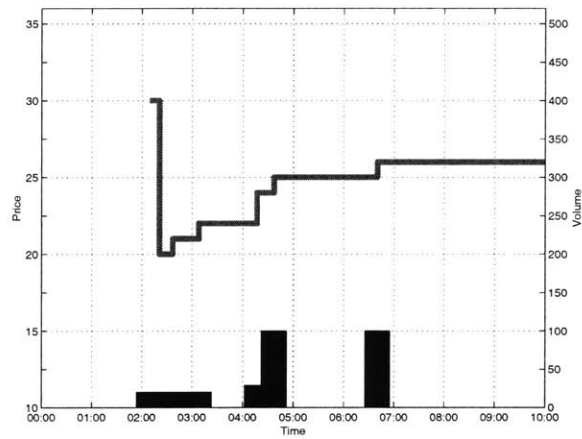
$$(p_{1,i}, q_{1,i}), (p_{2,i}, q_{2,i}), \dots, (p_{T_i,i}, q_{T_i,i}),$$

where i is the index for the i -th product and T_i is the number of trades for the i -th product. Our hypothesis is that prices reveal market consensus of the profitability of the bike pumps. To provide an analogous study to that by Dahan & Srinivansan (2000), we focused on the potential market share of the products. In particular, we propose that a product’s market share can be predicted by its relative market capitalization. The market capitalization of a company, or the total value of its stocks, equals to the product of market price and the number of outstanding shares. The relative market capitalization is defined as the ratio of a company’s market capitalization to the capitalization of the entire market (all the companies). Since all the companies have the same number of outstanding shares, market capitalization is proportional to the market prices. The market closing price is a natural candidate for the valuation of the companies. However, it is observed that the closing price is not

Experiment 1



Experiment 2



Experiment 3

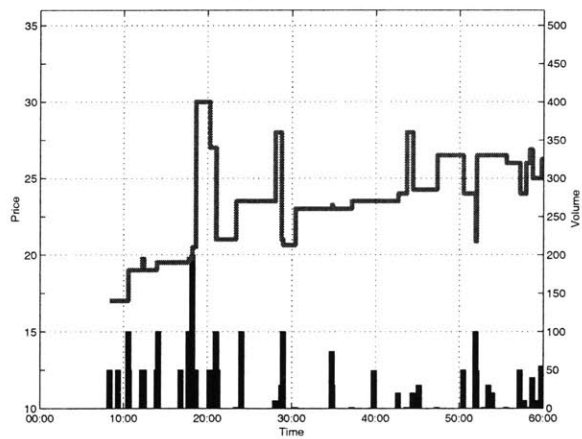


Figure 3-6: Price and volume history for AirStik in Experiments 1 to 3.

a particular robust measure for stock valuation. Because the traders' portfolios are valued at the closing prices, prices tend to become more volatile towards market close. This is especially true for the low-volume stocks. Hence, in addition to the closing price, we consider other price statistics that take into account *all* transactions during the session: the mean, median and volume weighted average price. The mean and median prices are calculated from the time series of trade prices $(p_{1,i}, p_{2,i}, \dots, p_{T_i,i})$; the volume-weighted average price (VWAP) is computed as follows:

$$VWAP_i = \frac{\sum_{t=1}^{T_i} p_{t,i} q_{t,i}}{\sum_{t=1}^{T_i} q_{t,i}}.$$

The mean price is sensitive to outliers—a small number of transactions that occur at extreme prices. Both mean and median prices ignore the volume in a transaction and treat all transactions equally. Volume can be regarded as a measure of the amount of information in a transaction. A trade with higher volume is generally more informative than one with lower volume. In our concept markets, volume is also related to how confident the traders are at the corresponding transaction price. VWAP effectively summarizes the prices by considering the amount of information and confidence behind the trades. In practice, VWAP has been a widely accepted benchmark price in financial markets. It is a commonly used metric for the evaluation of trade executions.

Now given a price statistic s_i^p , which can be the closing, mean, median or volume weighted average prices, we can compute predicted market share as the relative market capitalization.

$$MS_i = \frac{s_i^p n}{\sum_{j=1}^N s_j^p n} = \frac{s_i^p}{\sum_{j=1}^N s_j^p},$$

where N is the number of securities in the market and n is the total number of shares for a security. Among the four price statistics, we expect the median price and VWAP to be particularly robust against potential price volatility. To relate these metrics to those in the computer simulations, it is important to point out that the mean price in Section 3.6 is equivalent to the VWAP because all transactions are restricted one

	Skitzo	Silver Bullet	Epic	Gecko	Cyclone	Rim Gripper	Gear-head	TRS	2wister
Survey	32.1%	25.6%	10.3%	6.4%	3.8%	3.8%	2.6%	2.6%	1.3%
Exp 1	10.4%	13.3%	10.1%	7.5%	6.7%	5.3%	10.0%	7.0%	4.8%
Exp 2	12.2%	14.6%	10.5%	10.2%	5.7%	4.8%	6.4%	6.2%	5.5%
Exp 3	11.0%	11.5%	10.1%	7.5%	6.8%	7.6%	8.2%	7.5%	7.9%

Table 3.3: Market share predictions for nine products concepts by the survey method and Experiments 1 to 3. Market share predictions from experiments are calculated based on VWAP.

share.

To verify the validity of the market method, we ask two questions: (1) whether the results from the market method are consistent across different experiments, and (2) how close the results from the markets are to those from the independent survey study. We focus on the market share predictions derived from the four types of price statistics and those from the survey. Table 3.3 presents the predicted market share based on VWAP for the three experiments. We find that the top three products (Skitzo, Silver Bullet and Epic), in terms of predicted market share, are the same in the three experiments, as well as from the survey study. Furthermore, the rankings among the three are exactly the same across different experiments. In a typical concept testing process, it is more important to be able to identify the best designs because they are more likely to be materialized. Figure 3-7 presents the predicted market share based on all four price statistics for the three experiments. The mean price, median price and VWAP measures are reasonably close, while the closing price measure could be substantially different from the rest in certain stocks (Epic in Experiment 1, Gecko in Experiment 2 and Rim Gripper in Experiment 3).

For consistency across experiments, we calculate the pair-wise sample correlation between market share predictions based on each price statistic from individual experiments. For example, the association between experiments a and b is quantified by the sample correlation between $(MS_1^a, MS_2^a, \dots, MS_N^a)$ and $(MS_1^b, MS_2^b, \dots, MS_N^b)$. For comparison with the survey method, we calculate the sample correlation between the market share predicted by the survey study and those derived from individual

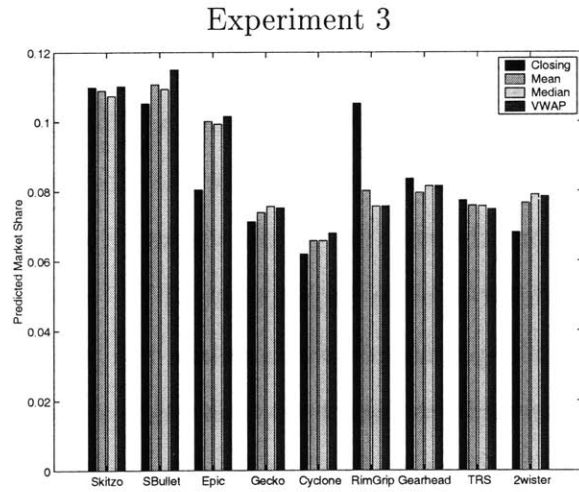
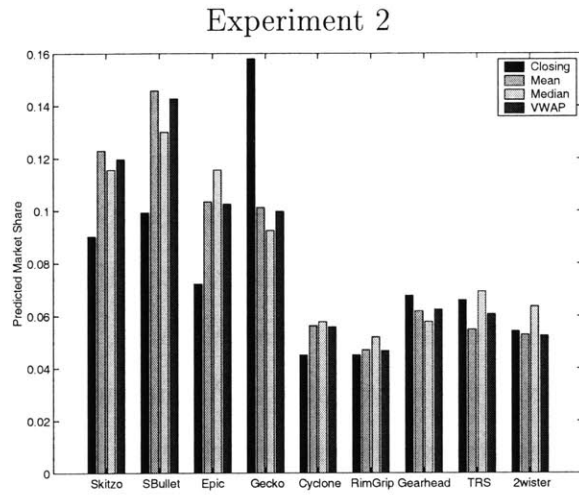
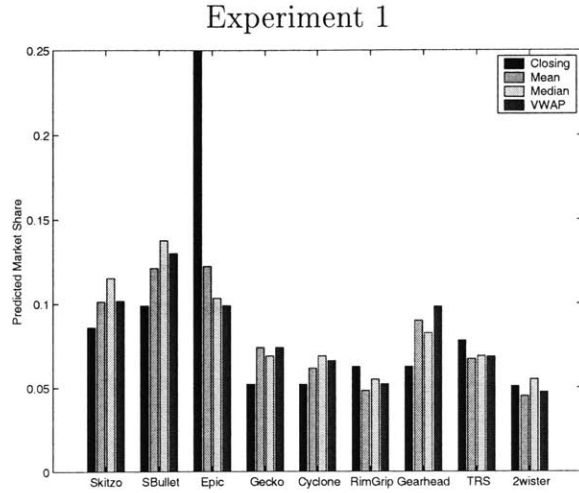


Figure 3-7: Predicted market share based on the closing price, mean price, median price and VWAP for Experiments 1 to 3.

Closing Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.22	0.34	0.72
Exp 3		1.00	-0.02	0.11
Exp 4			1.00	0.05
Exp 5				1.00

Mean Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.68	0.87	0.89
Exp 1		1.00	0.83	0.82
Exp 2			1.00	0.84
Exp 3				1.00

Median Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.87	0.83	0.88
Exp 1		1.00	0.89	0.92
Exp 2			1.00	0.89
Exp 3				1.00

Volume Weighted Average Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.74	0.86	0.89
Exp 1		1.00	0.85	0.82
Exp 2			1.00	0.86
Exp 3				1.00

Table 3.4: Price correlation coefficients between outcomes from the survey study, and Experiments 1, 2 and 3.

experiments. These correlation coefficients are presented in Table 3.4. The results from the three experiments show significant correlation. All correlation coefficients are above 80% for the mean price, median price and VWAP. The results from these three experiments are also highly correlated with the survey data: the correlation coefficients for median price and VWAP measures range from 74% to 89%. As we expected, the closing price is too noisy to give any conclusive results.

To ensure robustness of the results, we present another measure of association between different market share prediction results: Spearman’s rank correlation. We

Closing Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.66	0.63	0.58
Exp 1		1.00	0.42	0.68
Exp 2			1.00	0.34
Exp 3				1.00

Mean Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.73	0.82	0.60
Exp 1		1.00	0.90	0.62
Exp 2			1.00	0.55
Exp 3				1.00

Median Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.75	0.71	0.55
Exp 1		1.00	0.78	0.78
Exp 2			1.00	0.73
Exp 3				1.00

Volume Weighted Average Price				
	Survey	Exp 1	Exp 2	Exp 3
Survey	1.00	0.78	0.78	0.58
Exp 1		1.00	0.97	0.72
Exp 2			1.00	0.68
Exp 3				1.00

Table 3.5: Rank correlation coefficients between outcomes from the survey study and Experiments 1, 2 and 3.

transform all the market share data into ranks, $R_i = \text{rank}(MS_i)$, and calculate the sample correlation (see Table 3.5). A similar conclusion is reached: the results from all the three experiments are significantly correlated among themselves and with the survey data. Numerically, rank correlation coefficients are lower than the price correlation coefficients.

It is also interesting to quantify the association between experiments and the survey by noting their average absolute difference in the predicted market share. The

	Exp 1	Exp 2	Exp 3
Survey	6.1%	5.5%	6.8%
Exp 2		1.3%	1.4%
Exp 3			1.8%

Table 3.6: Average absolute difference in predicted market share between an experiment and the survey, and between two experiments.

difference between experiments a and b is computed as follows:

$$\Delta MS = \frac{1}{N} \sum_{i=1}^N |MS_i^a - MS_i^b|$$

The average absolute differences of predicted market share based on VWAP are presented in Table 3.6. The differences between individual experiments are smaller than those between an experiment and the survey study.

The results from our experiments show a remarkable agreement with those from the survey study despite many fundamental differences between the two methods. These differences include the differences in the data collection mechanism (a virtual security market versus a virtual shopping experience), the modeling of the predicted market share (the use of relative market capitalization of the virtual securities versus conjoint analysis), the questions asked (what you prefer versus what the group prefers), and lastly the subject population (MIT students versus Stanford students).

3.8 Conclusions

In this chapter we study a novel application of the market mechanism: the use of market to aggregate diverse consumer preferences. The idea is tested on a product concept testing study that aims to predict potential market share for some product prototypes. The results from three market experiments show high consistency among themselves and significant correlation with an independent survey study. Furthermore, an equilibrium model and two computer simulations are presented to provide

some insights into how such a virtual market may successfully aggregate the desired information.

Keynes (1958) wrote about the similarities between stock selection and a beauty contest:

“... professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitor as a whole ...”

The analogy is perhaps more accurate for describing what happens in the stock market in the short run. After all, stock prices depend not only on investors' subjective beliefs or expectations, but also on other objective information such as companies earning potential and valuations of assets. On the other hand, the trading experiments presented in this chapter are precisely “beauty contests,” since values of the virtual securities are derived endogenously from the expectations of the market participants, which are largely subjective. To improve the reliability of these virtual markets, one may need to anchor the values of the securities to some objective fundamental variables of the corresponding products. To predict future market shares, for example, one could link security values with the realized market shares of the corresponding products.

Chapter 4

A Rule-based Electronic Market-maker

In designing an efficient market, its institutional structure is the central issue (see Amihud, Ho & Schwartz (1985), Schwartz (1993), Madhavan (1992), O'Hara (1995) and Madhavan (2000)). Intense interest in market microstructure are driven by the rapid structural, technological and regulatory changes. In the U.S. equity markets, a substantial increase in trading volume and competition among exchanges and Electronic Communications Networks (ECNs) result in more rapid technological innovations. A large part of the trading process has been automated, and markets are becoming more “electronic,” particularly in the areas such as order processing and communications among market participants. However, market-making remains largely a human-intensive activity.

On the theoretical side, the trading behavior of market-makers has been one of the focal points in the literature of market microstructure. Various studies investigate specific aspects of the specialist's operation under theoretical settings, for example, market-maker's strategy under asymmetric information (Glosten & Milgrom (1985)) and the role of inventory control in market-making (Garman (1976) and Amihud & Mendelson (1980)).

The increasing availability of detailed market data motivates more empirical studies of the security markets at a level of details never possible before. This chapter aims

to study some heuristic market-making rules through computer simulations using a set of historical transaction data. The market-making rules, based on insights from the microstructure literature, address three primary objectives of a market-maker: (1) maximizing market quality, (2) maximizing profits, and (3) minimizing inventory risk. Two types of simulations, characterized by their order flow generation mechanisms, are conducted. First, historical order data is used directly in the study of the rule-based market-maker. We find that the application of the rules successfully improves the liquidity and maintains market quality that is comparable to that of the real-world markets. However, since the static order flow is not responding to the quotes, market-maker's inventory cannot be effectively controlled. The abnormally high inventory levels in turn lead to unusually high variance in the market-maker's profits. Second, a simulated order flow mechanism is modeled and estimated by combining the trade, quote and order data. In this case, inventory control is found to be more effective, while the objectives pertaining to profits and market quality are also accomplished. Unlike other simulation studies that are confined to simulated market environments, this chapter applies the model to historical data and relates the findings to empirical evidence. Both types of simulations are able to replicate several findings from theoretical models and empirical studies in the literature.

The chapter is organized as follows. Section 4.1 provides an overview of markets and market-making operations. Section 4.2 presents some prior studies of rule-based automated market-makers. The design of the rule-based market-maker is discussed in Section 4.3. The details of the static and modeled order flow simulations and the corresponding results are discussed in Sections 4.4 and 4.5.

4.1 Markets and Market-making

Markets provide trading mechanisms with which traders meet and transact. A trading mechanism is a set of rules governing how trading orders are submitted, how information is disseminated, and how and when specific traders transact. Trading activities may take place at centralized physical locations such as an exchange or

over computer networks and telephone lines. The New York Stock Exchange (NYSE) and the NASDAQ market are typical examples of these markets. The study of market microstructure is related to strategies of traders and market-makers, the trading mechanisms and their effects on price formation, and statistical modeling of different market variables.

In an abstract level, the essence of the trading process can be captured by the representation of a *Walrasian auction* market. A Walrasian auction largely ignores the details of the trading mechanism. A Walrasian auctioneer aggregates all traders' demands and supplies to determine a single market-clearing price, but does not participate as a trading principal in the process. In real-world auction markets, auctioneers are not likely to have the complete knowledge of traders' demand and supply information that is essential to the frictionless operation of Walrasian auctions.

Most securities markets are variations of a *double auction* market. In its simplest form, a double auction market is one in which traders can submit a bid (an order to buy at a specific price or lower) an ask (an order to sell at a specific price or higher), and market buy and sell orders (orders that transact at the corresponding quotation prices at the opposite side of the market). A transaction occurs when there is a match between the buy and sell sides: (1) when the bid price exceeds the offer price or (2) when there are market buy (sell) orders matching existing limit sell (buy) orders. When the auctioneer receives the orders, he either places them in a limit order book (a record of all limit orders), or executes them with the matched orders in the book. He is also responsible in disseminating quote information (best bid and offer prices and sizes) and trade information (the price and size of executions) to the public. The auctioneer maintains a passive role in the sense that he does not trade for his own account. GLOBEX of the Chicago Mercantile Exchange and all Electronic Communication Networks (such as Instinet and Island) are examples of a double auction market. The exact trading rules may vary from one type of market to another. In this thesis, we loosely apply the term double auction to markets without any market-maker's intervention.

Trading in double auctions is limited by the availability of trading interests on the

opposite side of the market. For example, a particular stock may have a much larger number of buy orders than sell orders (or visa versa). This kind of disparity in demand and supply is known as an order imbalance. In a market with an order imbalance, a market-maker may supply *immediacy* and *liquidity* by specifying a bid or an ask price to ensure the presence of a market. He may also improve the depth (quantity of shares to trade given a certain price) of the market by supplying additional bids and asks. Price discovery is another major responsibility of a market-maker. It is a process used to determine the efficient prices of securities, based on demand and supply in the market and other relevant market information. The market-maker is in a privileged position to have access to the information from the order flow that reflects the ever-changing forces of demand and supply. In providing these services, the market-maker must actively participate in the market—trading on his own account and establishing positions. As a result, he may bear the risk of holding unwanted inventory. Lastly market-makers are usually regulated for the quality of market they provide. Measures of market quality include transaction-to-transaction price change, stabilization of the price (price range in a sequence of transactions), the size of the bid-ask spread (difference between bid and ask prices), the depth of the market, etc.

Agency auction markets are auction markets with monopolistic market-makers, that is, one market-maker for one stock. The NYSE and the AMEX are typical agency auction markets. Designated market-makers at these markets are called specialists. The following is how an agency auction market works. Orders arrive in the specialists' posts on the trading floor either electronically or through a floor broker. The orders are then entered into the electronic Display Books which are computer workstations that keep track of all limit and market orders. Specialists manage the auction by interacting with the floor brokers and working with the limit order book (a collection of all the public limit orders). In real time, quote and trade information is reported via electronic networks to all market participants, inside and outside the exchange. The market-making operation is subject to a set of rules and regulations which ensure the quality of the market. At the NYSE, these rules and regulations are referred to as the specialists' *affirmative* and *negative* obligations. From the *Official Floor Manual*

of the NYSE, the affirmative obligation of the specialists can be summarized as

“the maintenance of a fair and orderly market in the stocks in which he is registered, which implies the maintenance of price continuity with reasonable depth and minimizing of temporary disparities between demand and supply ...”

In fulfilling his affirmative obligations, a specialist needs to trade for his own account. But trades that are not necessary for the maintenance of a fair and orderly market are discouraged or even restricted. These criteria constitute the negative obligations. Example of trades that are not to be effected include the purchase at a plus tick or zero plus tick¹, or the purchase of a substantial number of shares offered in the limit order book. The Stock Exchange monitors closely the trading activities of the specialist. Rule violations can result in fines; evaluations of the execution quality are considered in allocating new listings to the specialists.

Agency auction markets ensure efficiency and quality of a market by regulating monopolistic market-makers. Alternatively, competition can be introduced to a market as a driving force for market quality. In a multiple dealer market, there are more than one market-maker or dealer for a single security. Competition among the dealers brings about the efficiency of the market. The NASDAQ market is a typical multiple dealer market. There have been constant debates about which type of market structure is more efficient. The discussions have been focused on the pros and cons of consolidated versus fragmented markets, and the effectiveness of regulation versus competition.

This chapter focuses on the agency auction market, in particular the NYSE. We attempt to model a rule-based electronic market-maker which resembles the specialist at the NYSE. Transaction data from the NYSE is used to verify the market-making rules. A model of the order flow generation mechanism is estimated using the same set of data.

¹A plus tick is resulted when a trade is executed at a price higher than the last sale price. A transaction that follows a plus tick and has the same sale price constitutes a zero-plus tick. Minus and zero-minus ticks are defined in the same way.

4.2 Background

4.2.1 Simulation Models with Market-making Rules

Optimal market-making strategies have been studied by the researchers to capture the dynamics of exchange trading in theoretical environments. In reality, exchange specialists or dealers actually use relatively simple price-adjustment or inventory management rules. There are a few attempts to model specific functions of a market-maker in the rule-based domain. These studies share a common approach that usually consists of three components. First there is a specified market structure that dictates the trading mechanism: the types of orders permitted in the market, whether it is a batch market or continuous market, etc. Second there is a group of traders or investors, who generate orders by simulating their underlying demand functions, which are contingent upon the market-maker's quotes and other market variables. Lastly, a market-maker (or multiple market-makers) will play a central role in these analyses. Upon receiving the orders, it executes or matches them according to its strategies.

Hakansson et al. (1985) studies the feasibility of automated market-making using a set of of rules. The functions of the automated market-maker are limited to “demand smoothing,” which is the balancing of excess demand or supply due to discontinuities of investors' demand functions. A specialist accomplishes demand smoothing, one of his affirmative obligations, by absorbing any excess demand or supply into his own account. Trading orders are generated from individual traders' stochastic demand functions. The market-maker's decision is programmed to follow a collection of rules. In particular, this market-maker considers a downward sloping excess demand function specified by n price and quantity pairs:

$$(p^1, q^1), \dots, (p^k, q^k), (p^{k+1}, q^{k+1}), \dots, (p^n, q^n),$$

such that $q^k \leq 0 \leq q^{k+1}$. The majority of the rules are concerned with inventory control and execution quality aspects of market-making. Two examples of the rules examined in the paper are as follows:

$$\text{R2} \quad \min_{\{p^k, p^{k+1}\}} q_t$$

$$\text{R4} \quad \min_{\{p^k, p^{k+1}\}} |p^k - p_{t-1}| + |p^{k+1} - p_{t-1}|$$

where q_t is market-maker's stock holding time t . Rules R2 chooses a price so as to minimize the inventory holding while Rule R4 attempts to smooth transaction-to-transaction prices. It is shown that R4 will result in the market-maker's holding or short selling of all the shares. This result agrees with what Garman (1976) finds: market-makers must determine their prices by considering their inventory to avoid failure. In general, it is found that rules that minimize inventory performed well, in terms of lower stock holdings and dealer's participation.

Cohen, Maier, Schwartz & Whitcomb (1983) presents another simulation study of market-making and stock exchange trading. The main focus of the study is the effects of stabilization and speculation on the performance of the market and the profitability of the market-maker. The simulation model describes a continuous market where the market-maker acts as a "pure stabilizer" or "speculating stabilizer." A pure stabilizer is an automated market-maker whose sole objective is to stabilize transaction prices; a speculating stabilizer attempts to earn trading profits while maintaining price stability. Orders arrive at the market following a Poisson process. The size of an order is randomly drawn from a gamma distribution; the price of an order is sampled from a distribution that conditions on the market-maker's bid and ask quotations. The policy of the pure stabilizer is constrained by two parameters: a maximum allowable transaction-to-transaction price change, and an inventory limit. If an order would result in a price change that exceeds the allowable limit, the stabilizer will enter a limit order inside the public quote to reduce the price change. But the agent will cease to stabilize price if its inventory reaches the limit. The speculating stabilizer is subject to the same price-change and inventory constraints as the pure stabilizer, but it also attempts to make a profit using order imbalance information in the limit order book.

Simulation results show that the stabilizing rule is effective in improving price continuity. For small maximum allowable price changes, the stabilizing rule is not profitable, which implies that this affirmative obligation is provided at the expense

of the market-maker. The percentage spread is improved with increased inventory limit, which gives the market-maker more freedom in its operation. The standard deviation of the profit, however, goes up with the inventory limit. It is because with a higher inventory limit, the inventory valuation becomes more volatile. The speculating stabilizer, as expected, runs a more profitable operation than the pure stabilizer. It is interesting to note that the speculating policy further reduces the percentage spread compared with that of the pure stabilizing policy. This implies that both the investors and the market-maker are better off under the speculating policy.

4.2.2 Empirical Studies

Most of the academic studies of the trading behavior of dealers have focused on the theoretical aspect due to the difficulty in obtaining detailed relevant data, in particular, dealer positions and trading profits. Many empirical questions regarding dealer behavior in real markets remain unanswered. With data provided by exchanges, a few studies conduct transaction level investigations of dealer trading. Hasbrouck & Sofianos (1993) examines inventory adjustment and profitability of the NYSE specialists from the exchange's specialist data files². The paper focuses on the speed of inventory adjustment (adjustment of inventory to desirable levels) across different stocks, and the decomposition of specialist trading profits by trading horizon. Sofianos (1995) provides detailed statistics on NYSE specialists' trading profits for 2,511 stocks in May 1995. Madhavan & Sofianos (1998) studies the size, direction and timing of the specialists' trades, and suggests that dealers control their inventory by actively buying and selling rather than by adjusting the quotes.

²The specialist equity trade (SPET) file that contains specialist transaction data from November 1990 to January 1991, and the specialist equity trade summary (SPETS) file contains daily summary data covering from November 1988 to August 1990.

4.3 Design of the Market-maker

This chapter is devoted to the study of the behavior of an automated monopolistic market-maker such as the specialists at the New York Stock Exchange and American Stock Exchange. An automated market-maker can be modeled, optimized and calibrated for different types of markets (e.g., stock or commodity markets), different types of securities (e.g., General Electric or Amazon.com), various market conditions (e.g., low versus high volatility), and different objectives (e.g., profit maximization versus market quality maximization). Automation in market-making can add value to financial and non-financial markets in many ways. Improvement in operational efficiency and increased reliability are two obvious benefits. Fairness and transparency in a market can also be enhanced because the logic, parameters and objectives of a market-making program can be publicly disclosed, examined and validated. Market globalization and trading hour extension have been two most recent developments in securities markets. Automated market-making technologies will also play an important role in accomplishing global and 24-hour trading.

There are two views of the market-making operation. First, from the point of view of a market-maker, he would be interested in maximizing his profits while minimizing his risks, especially the inventory-carrying risk. From the investors' point of view, the primary considerations are liquidity of the market, cost of execution, price continuity and other market quality measures. It is important to point out that these two views are conflicting with each other. The upholding of a market-maker's affirmative obligations could affect his profitability. For example, to maintain an orderly market, a market-maker needs to commit his capital to buy or sell against the market trend, and consequently could incur losses and take on unwanted inventory positions. Narrowing bid-ask spread directly reduces a market-maker's trading revenue.

These concerns from both views are addressed by three independent modules of the rule-based market-maker, namely, the *market quality*, *inventory control* and *speculative motive* modules. Each of the modules evaluates the desirability of a quote with respect to the corresponding objective. An arbitration mechanism is used to

combine and conciliate the outputs from the three modules, according to their relative importance specified by some weight parameters. By specifying these weights, one can tune the agent to satisfy any specific objectives of a market or a market-maker. For example, one can build a market maker that maximizes profits subject to some minimum market quality requirements (such as the NYSE specialist), or one that makes a fixed amount of profits while providing the highest possible market quality.

The following describes the quote adjustment mechanism of the electronic market-maker. A quote posted by the market-maker consists of a bid price, bid size, ask price and ask size:

$$\mathbf{q} = (p^b, s^b, p^a, s^a).$$

The electronic market-maker focuses on the quote prices but always fixes the quote sizes s^b and s^a at 1,000 shares³. The minimum price variation for the quote prices is 1/8. The maximum bid-ask spread is limited to 1/4.

At the beginning of a trading day, the electronic market-maker opens the market around the opening price of the stock, p_0 , given in the historical data. In particular, it sets $p^b = p_0 - 1/8$ and $p^a = p_0 + 1/8$. Then public orders arrive at the market. Limit orders are added to the limit order book or crossed with the quote on the other side of the market. A quote revision follows the execution of an order or the addition of a limit order into the limit order book. To revise the quote, the market-maker considers possible quotes in the neighborhood of the last transaction price and *inside* all the public limit orders. That is, the market-maker always quotes its bid and ask prices better or equal to the best public limit prices. Consider the following example.

Example: The stock XYZ is quoted 20 bid and offered at 20 1/8. A limit sell order (denoted as order A) of 1,800 shares at 20 will cross with the bid and result in a trade of 1,000 shares at 20. The market-maker revises the quote by considering all limit orders which include the remaining of order A—an order to sell 800 shares at 20, and

³Quotes that represent public limit orders could exceed this size.

an existing order to buy 400 shares at $19 \frac{3}{4}$. Possible quotes in terms of (p^b, p^a) are $(19 \frac{3}{4}, 19 \frac{7}{8})$, $(19 \frac{3}{4}, 20)$ and $(19 \frac{7}{8}, 20)$. Eventually XYZ is quoted $19 \frac{7}{8}$ to 20, 1,000 shares by 1,000 shares, with 400 shares by 800 shares originated from the public limit orders.

Market orders are executed against the current quote and subsequent revised quotes until they are fully executed. Consider the following examples.

Example: The stock XYZ is quoted 20 bid and offered at $20 \frac{1}{8}$. A market sell order of 1,500 shares will be executed at 20 for the first 1,000 shares. The quote is revised to $19 \frac{7}{8}$ to $20 \frac{1}{8}$. 1,000 shares by 1,000 shares. The remaining 500 shares are executed at $19 \frac{7}{8}$. Another quote revision follows.

The process of quote revision involves the evaluations of all possible quotes by the three independent modules, namely, market quality, inventory control and speculative motive.

4.3.1 Market Quality (MQ)

Some market makers such as the NYSE specialists are regulated by specific rules of the exchange to maintain a minimum execution quality. One of the major motivation in automating the market making process is to ensure a fair and high quality market. To ensure an efficient market, this module focuses on the following measures:

- Price continuity: the transaction-to-transaction price change (PC)
- Price stability: the high-low price difference in a sequence of trades with total volumes of 50, 100 and 250 100-share lots (PR_{50} , PR_{100} and PR_{250})
- Bid-ask spread: the difference between the bid and ask prices (SP);

The value of an action is computed as

$$V_{MQ}(\mathbf{q}) = \alpha_1 PC(\mathbf{q}) + \alpha_2 PR_{50}(\mathbf{q}) + \alpha_3 PR_{100}(\mathbf{q}) + \alpha_4 PR_{250}(\mathbf{q}) + \alpha_5 SP(\mathbf{q}),$$

where α_i 's are the weights specifying relative importance of the market quality measures. The price continuity function is defined as the average price change for trades occurring at the buy and sell side:

$$PC(\mathbf{q}) = -\frac{1}{2}\{|p_t - p^b| + |p_t - p^a|\},$$

where p_t is the last transaction price. The function $PR_n(\mathbf{q})$ calculates the price ranges of the shortest sequence of trades that involves a total volume of $100n$ shares. Suppose a trade occurs at the bid price, the shortest sequence of trades with a total volume equal or larger than $100n$ shares is executed at the prices

$$\mathbf{P}_n^b = (p_{t-t_b+1}, \dots, p_t, p^b),$$

where t_b is the number of the trades that are sufficient to accumulate the specified volume. A price series \mathbf{P}_n^a is defined in a similar way for the trade that occurs at the ask price:

$$\mathbf{P}_n^a = (p_{t-t_a}, \dots, p_t, p^a).$$

The price range $PR_n(\mathbf{q})$ is computed as the average price range on both sides of the market:

$$PR_n(\mathbf{q}) = -\frac{1}{2}\{(\max(\mathbf{P}_n^b) - \min(\mathbf{P}_n^b)) + (\max(\mathbf{P}_n^a) - \min(\mathbf{P}_n^a))\}$$

The spread of a quote is simply defined as the difference between the bid and ask prices:

$$SP(\mathbf{q}) = -(p^a - p^b).$$

4.3.2 Inventory Control (IC)

Theoretical models such as those by Amihud & Mendelson (1980) and Ho & Stoll (1981) show that market-makers who are risk-adverse adjust their quotes as a function of their inventory holdings. The inventory control module aims at keeping an optimum level of inventory. In the process of providing liquidity and maintaining market quality, the market-maker may have purchased or sold shares of the security and resulted in an undesirable position. If his position is higher than what he wants, he would want to reduce it by lowering the offer price and raising the bid price so as to encourage buying and discourage selling from the public.

The solution to this problem comes down to the setting of the bid and offer prices that would satisfy some optimum inventory condition, which is represented by a target inventory, INV^* . We denote INV_t as the current inventory, ΔINV^b as the “expected” change of market-maker’s inventory after a trade occurs at the bid, and ΔINV^a as the “expected” change of inventory after a trade occurs at the ask. The “expected” changes of inventory, depending on the quote prices and sizes, are heuristically determined but not defined on any probability space. Consider the case when the market-maker is holding a long position: $INV_t > 0$. The lower the bid price, the less likely the bid will be hit. Specifically, we define the minimum and maximum bid prices among all the n quotes under consideration: $MIN_{p^b} = \min(p_1^b, p_2^b, \dots, p_n^b)$ and $MAX_{p^b} = \max(p_1^b, p_2^b, \dots, p_n^b)$. The “expected” change of inventory is

$$\Delta INV^b = \frac{s^b}{2} \left(1 + \frac{p^b - MIN_{p^b}}{MAX_{p^b} - MIN_{p^b}} \right),$$

such that $\Delta INV^b = s^b$ for the highest bid price, and $\Delta INV^b = s^b/2$ for the lowest bid price. The variable ΔINV^a is defined in the same way for the sell side of the market.

The value of a quote depends on the absolute difference between the resulting

inventories and the target inventory:

$$V_{IC}(\mathbf{q}) = \begin{cases} |(INV_t + \Delta INV^b) - INV^*| & \text{for } INV_t \leq 0 \\ |(INV_t - \Delta INV^a) - INV^*| & \text{for } INV_t > 0 \end{cases}$$

4.3.3 Speculative Motive (SM)

This module focuses on the unique role of the market-maker who overlooks the order flow and possess privileged access to the limit order book. It aims to exploit the order imbalance information embedded in the limit order book to earn trading profits. Specifically it considers the public limit orders whose prices are within the neighborhood of the last transaction price. If there are more buy orders than sell orders, for example, the market-maker would predict that the price is likely to go up, and therefore try to post higher quotation prices. However, at the same time, the market-maker also attempts to gain from the bid-ask spread by placing the bid at its lowest possible price, which is constrained by the maximum allowable spread and the highest public bid. We consider the buy and sell sides of public orders separately. First consider the set of n public bids that are in the neighborhood of the last transaction price p_t :

$$\{(p_1^b, s_1^b), (p_2^b, s_2^b), \dots, (p_n^b, s_n^b)\}, \quad \text{such that}$$

$$p_t - TH \leq p_i^b \leq p_t + TH, \quad \text{for } 1 \leq i \leq n,$$

where TH is the threshold that specifies the set of limit orders of interest, and n is the number of the corresponding bids. We want to focus on limit orders whose prices are close to the current market price because orders that are far off may not be very informative. The volume weighted average price ($VWAP$) and total size (TS) are

defined as functions of this sets of bids:

$$VWAP^b = \frac{\sum_{i=1}^n p_i^b s_i^b}{\sum_{i=1}^n s_i^b},$$

$$TS^b = \sum_{i=1}^n s_i^b$$

Similarly, variables $VWAP^a$ and TS^a are defined on a set of public asks. Finally, the value of a quote based on the order flow information is defined as

$$V_{SM}(\mathbf{q}) = -\frac{1}{2}\{(p^b - VWAP^b)TS^b + (p^b - VWAP^a)TS^a\}.$$

The volume weighted average prices measure the distance between a quote and corresponding public limit orders, while the total sizes capture the order imbalance information and serve to control the trade-off between the bid and ask distances. Suppose there are more public asks than bids, $TS^a > TS^b$, the module would prefer quotes with asks closer to the public asks.

4.3.4 The Arbitration Mechanism

The three modules are combined linearly representing the overall strategy of the market-maker:

$$V(\mathbf{q}) = w_{MQ}V_{MQ}(\mathbf{q}) + w_{IC}V_{IC}(\mathbf{q}) + w_{OI}V_{OI}(\mathbf{q}),$$

where the weight parameters w_{MQ} , w_{IC} and w_{OI} control the trade-off among the three objectives.

4.4 Simulations with Historical Data

4.4.1 The Data Set

The analysis of the model is based on the TORQ data set from the New York Stock Exchange. The TORQ data set contains transaction, quote, order and audit trail data for a sample of 144 NYSE stocks during the three months period from November 1990 through January 1991. This chapter primarily focuses on the transaction, quote and order files of the database.

There are several important limitations and deficiencies of the data set. The transaction and quote data contain all trades and quote revisions from all trading venues (the NYSE and all the regional exchanges). The order data, however, covers only the orders that enter the market through the Exchange's electronic order routing system (SuperDot system). Other orders, which are usually larger and more difficult to execute, are assisted by floor brokers. These floor orders, representing a significant portion of the order flow, are not available from the data set. Hasbrouck, Sofianos & Sosebee (1993) reports that 75 percent of orders entered the market via SuperDot but only accounted for 28 percent of the executed share volume in 1992.

In studying the data set, one may attempt to relate individual files with each other for a complete picture of the trading activity. However, the time-stamps of the records in the database may pose a problem. First, the time-stamps may not reflect the time of interest. For example, the time associated with a trade is the time at which the trade is recorded but not when it occurs. Second, different time-stamps from different files are recorded by different systems that may not be synchronized to the same clock. For instance, trade records are time-stamped by the Consolidated Trade System (CTS) and quote revisions are time-stamped by the Consolidated Quote System (CQS).

For simplicity, some observations are ignored and some assumptions are made. In the analysis, we only consider limit and market orders, and ignore other order types

that include stop orders⁴, stop limit orders, market-on-close orders⁵, etc. All *good-till-canceled* limit orders⁶ are assumed to be *day-orders*⁷, which are canceled when the market closes each day. Among 144 stocks in the database, 81 stocks that have an average number of daily transactions of more than twenty are selected. This is to ensure that there are enough observations for the estimation of an order flow model.

4.4.2 Generation of the Order Flow

Two approaches are considered in generating an order flow for the trading simulations. First we consider using orders directly from the database, that is, submitting orders to the market-maker according to their time-stamps. The main advantage of this approach is its simplicity. But the order flow in this case is static and not reacting to any market variables, including the market-maker's quotes. This limitation of the static order flow could result in severe distortion in a simulation. For example, a market-maker cannot effectively reduce its inventory by lowering its quote prices in facing a static order flow. This is because the orders were submitted in response to the instantaneous quotes in the market when the data was recorded; the quotes posted by the electronic market-maker are irrelevant.

To establish the dependence of the order flow on the market-maker's quotes, we attempt to model an order flow generation process. One approach is to create a purely simulated environment in which the information structure, investor objectives and trading strategies can be completely specified. This approach is studied in next chapter. In this chapter we attempt to estimate the dependence of the order flow on the market quotes from the historical data. We focus on the following attributes of an order:

1. *side*: buy or sell

⁴A stop order is an order that turns into a market order when the last trade price hits the "Stop" price.

⁵A market-on-close is a market order which is to be executed at the closing price.

⁶A limit order that can be filled any time before the market closes.

⁷A limit order that can be filled any time prior to cancellation.

2. *type*: market or limit orders
3. *price*: limit price (limit orders only)
4. *size*: number of shares

Specifically, an order is represented by a vector (BUY, LPR, SZE) , where BUY is an indicator variable for buy orders,

$$BUY = \begin{cases} 1 & \text{for a buy order} \\ 0 & \text{for a sell order,} \end{cases}$$

LPR denotes an order's price,

$$LPR = \begin{cases} 0 & \text{for a market order} \\ \text{limit price} & \text{for a limit order,} \end{cases}$$

and SZE denotes an order's size. A simulated order is generated by sampling from the empirical distributions of these variables, conditioned on the market-maker's quotes and other relevant variables.

An order is generated by determining the four attributes in a specific sequence. Imagine a trader enters the market and observes the quotes. He may first decide whether to buy or sell, based on his gathered information regarding the fair value of the stock. Trading decisions may also be motivated by other economic objectives such as liquidity needs. Then he needs to decide whether to submit a limit or market order, and at what price if a limit order is to be placed. Again his order placement strategy would depend on the current quote. Finally he needs to specify the size of the order.

The side of an order is determined by considering the current bid and ask prices, p_t^b and p_t^a , and the future stock price. The future stock price, p_t^f , is the volume weighted average price of all transaction price in the next 10 minutes. It serves as a reference to the quote prices. Suppose the sequence of trades from time t to $t + 360$

Estimate of	Total Sample	GE	IBM	SLB	W	UWR
$\Pr(BUY = 1 A^1)$	0.37 (0.14)	0.31	0.31	0.44	0.20	0.47
$\Pr(BUY = 1 A^2)$	0.41 (0.06)	0.44	0.40	0.43	0.37	0.36
$\Pr(BUY = 1 A^3)$	0.54 (0.07)	0.53	0.48	0.56	0.52	0.67
$\Pr(BUY = 1 A^4)$	0.65 (0.09)	0.67	0.61	0.67	0.70	0.85
$\Pr(BUY = 1)$	0.50 (0.05)	0.49	0.46	0.52	0.50	0.63

Table 4.1: Empirical distribution of the indicator variable BUY for the entire sample of 81 stocks and six sample stocks.

is $\{(p_1, q_1), (p_2, q_2), \dots, (p_{K_t}, q_{K_t})\}$, the volume weighted average price is

$$p_t^f = \frac{\sum_{k=1}^{K_t} p_k q_k}{\sum_{k=1}^{K_t} q_k}.$$

The idea is that more buy orders are expected if the price will go up, and more sell orders are expected if the price will go down. Specifically, we are interested in the conditional probability $\Pr(BUY_t = 1|p_t^b, p_t^a, p_t^f)$. To further simplify the problem, we consider the following discrete events

$$\begin{aligned} A_t^1 &= \{p_t^f \leq p_t^b\}, \\ A_t^2 &= \{p_t^b < p_t^f \leq (p_t^b + p_t^a)/2\}, \\ A_t^3 &= \{(p_t^b + p_t^a)/2 < p_t^f < p_t^a\}, \\ A_t^4 &= \{p_t^a \leq p_t^f\}, \end{aligned}$$

and estimate the conditional probabilities $\Pr(BUY_t = 1|A_t^i)$ from the data. These probabilities are estimated by calculating, in terms of shares volume, the proportion of buy orders to all orders given the event A^i . Table 4.1 shows the estimates of $\Pr(BUY_t = 1|A_t^i)$ for six stocks, and the sample mean and standard deviations for the entire sample of 81 stocks.

Given the side of an order, an order placement strategy can be determined depending on the spread, $p^a - p^b$. In estimating the empirical distributions, $P(LPR|BUY, p^a - p^b)$, buy orders with limit prices $LPR < p^b - 4$ are assumed to have $LPR = p^b - 4$; sell orders with limit prices $LPR > p^a$ are assumed to have $LPR = p^a$. Similarly, sell orders with limit prices $LPR > p^a + 4$ are assumed to have $LPR = p^a + 4$; sell orders with limit prices $LPR < p^a$ are assumed to have $LPR = p^a$. For market orders, $LPR = 0$. Figure 4-1 presents the empirical distribution $P(LPR|BUY, p^a - p^b)$ for the stock IBM. Order prices are randomly sampled from these empirical distributions conditioned on the indicator variable BUY and the bid-ask spread.

Lastly, the size of an order is determined by a random draw from the empirical unconditional distribution of the order size $P(SZE)$. Figure 4-2 describes the simulations with the static and modeled order flow.

4.4.3 Simulations

In studying the feasibility of the heuristic rules, we simulate trading with the static and modeled order flow. Three settings of the rule-based market-making agent are considered. In the first settings, denoted as “MQ,” the agent only focuses on maximizing market quality. The second, denoted as “MQ+IC,” extends from the first to further include inventory control. The third, denoted as “SM+IC,” considers the speculative motive and inventory control.

To compare the performance the rule-based agent, two types of “benchmark” analyses are conducted. A natural candidate for such a comparison is the execution of the actual NYSE specialists. Unfortunately the TORQ data does not provide the details of the specialists’ trades, such as the details of the purchases or sales, and stock positions, which are necessary for the analysis of their inventory and profits. Nevertheless, information concerning measures of market quality including price continuity, price stability, and the spread, can be calculated from the trade and quote data. This analysis is denoted as “NYSE.”

To investigate the fulfillment of the affirmative obligations imposed on the rule-based market-maker, we compare its performance with that of an *order matching*

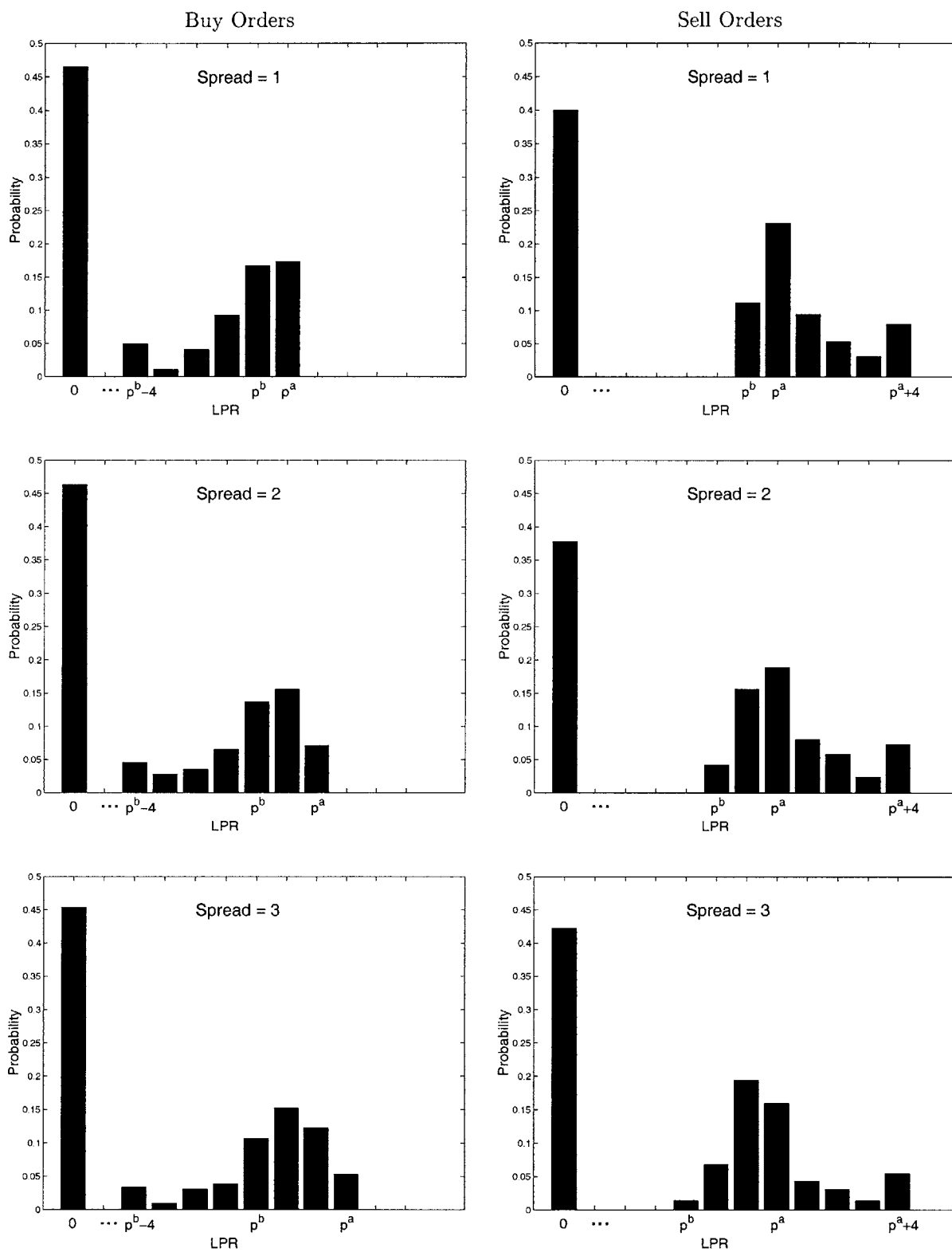


Figure 4-1: Empirical distribution $P(LPR|BUY, p^a - p^b)$ for the stock IBM.

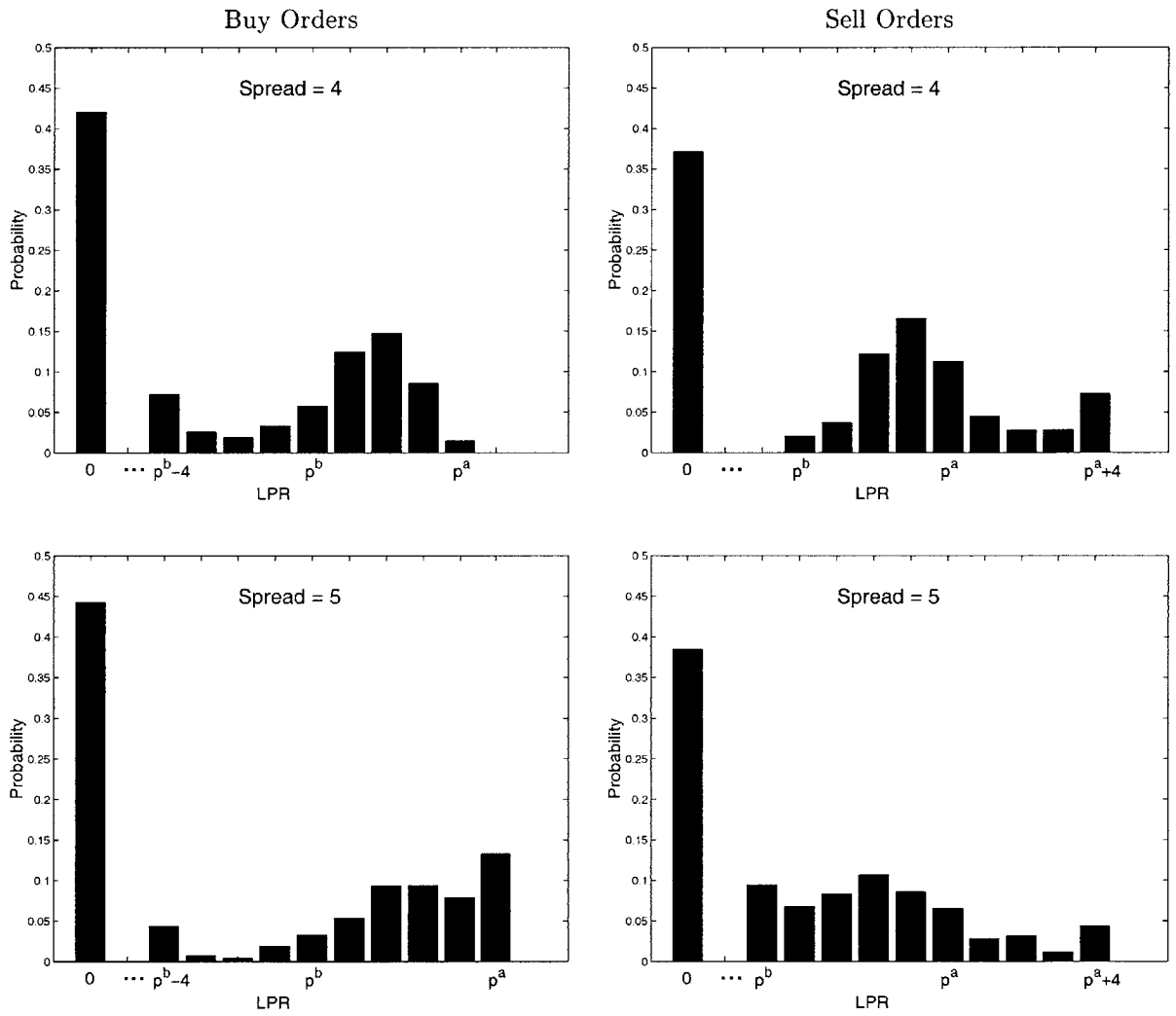


Figure 4-1 (continued)

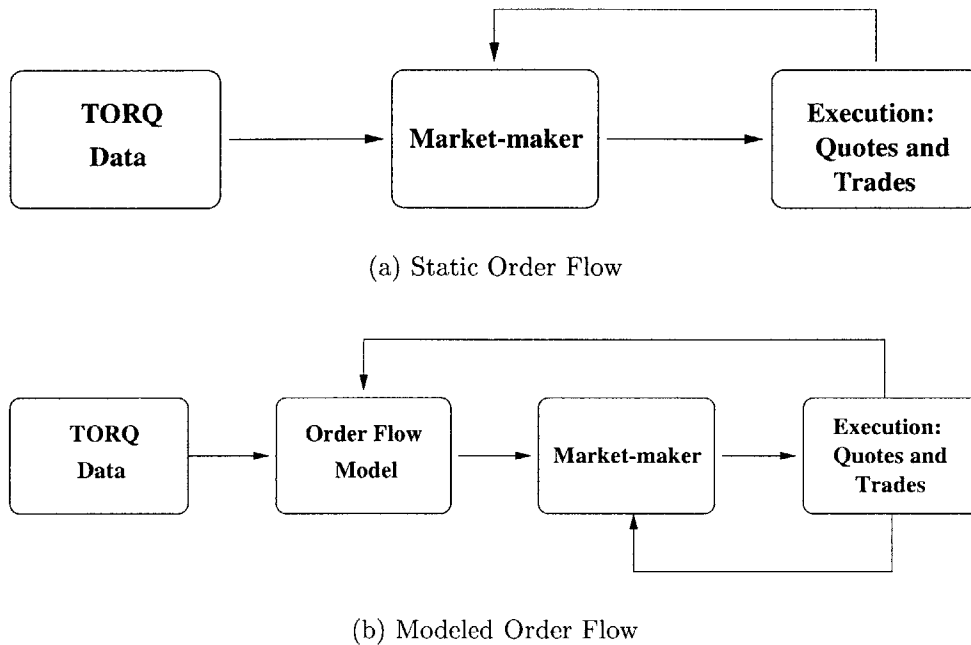


Figure 4-2: Two approaches of the order flow generation mechanism.

mechanism or an auctioneer. This simulates a double auction market in which the auctioneer simply matches orders eligible for execution. Since the auctioneer does not intervene to ensure a market presence, market orders that find no matching limit orders on the other side of the market are canceled. This simulation is denoted as “OM.” The details of all simulations and analyses are summarized in Table 4.2.

4.5 Performance of the Rule-based Market-maker

Details of the results of the simulations with the static and modeled order flows are presented in Tables 4.3 and 4.4. All stocks are ranked by their average daily number of transactions and grouped into four quartiles sub-samples.

4.5.1 Static Order Flow

The static order flow is the stream of orders directly taken from the order data file. These orders were placed given the market quotes and market conditions (such as

Simulation	Description	Data
NYSE	NYSE specialist execution	Trade and quote data
OM	Order matching mechanism	Static order flow
MQ	Electronic market-maker that contains the market quality module	Static and modeled order flow
MQ+IC	Electronic market-maker that contains the market quality and inventory control modules	Static and modeled order flow
SM+IC	Electronic market-maker that contains the speculative motive and inventory control modules	Static and modeled order flow

Table 4.2: Description of different modes of simulation.

news) at the time of submission. For this reason, these orders do not respond to quotes posted by the rule-based market-maker. Consequently, inventory control by adjusting the quote prices becomes ineffective because the underlying order submission dynamics are fixed.

Table 4.3 presents the results from the simulations with the static order flow and the comparison with the NYSE specialist executions. The ineffectiveness in inventory control results in abnormally high closing inventory for the rule-based market-maker. The high inventory holdings are subject to the volatility in market valuations, and in turn lead to high variance in the daily profits. This unfortunately makes the interpretation of the agent's profitability inconclusive. For example, the simulation MQ gives an average daily profit of \$710 but the sample standard deviation is \$577,810. However, the IC modules does reduce inventory holdings to some extent. Figure 4-3 shows the daily closing inventory in simulation MQ and MQ+IC for the stock IBM for the 63-day period, and demonstrates the effect of inventory control. The simulation SM+IC results in a lower average absolute closing inventory than that of MQ+IC, suggesting that in maintaining high market quality, the market-maker has to take on more unwanted inventory.

The electronic market-maker successfully provides additional liquidity to the market. The average trading volume is substantially increased from the order matching

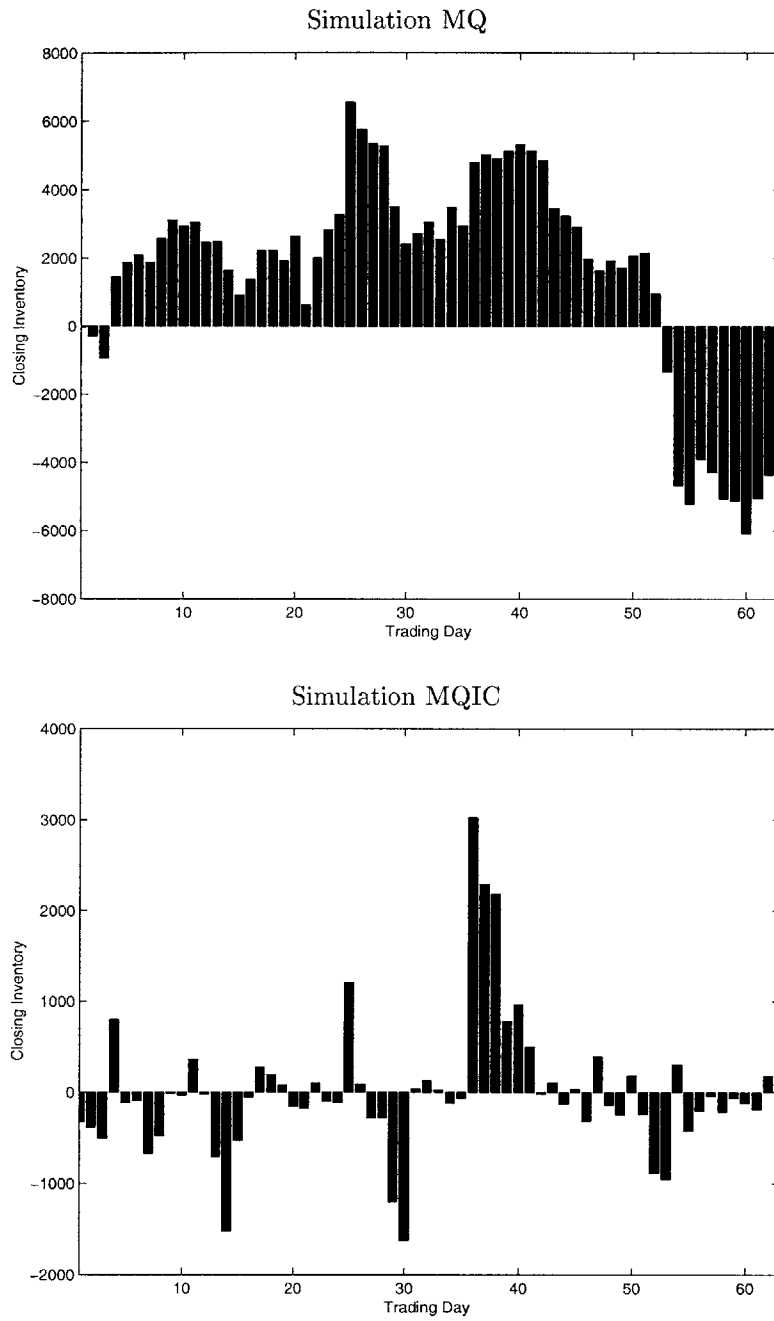


Figure 4-3: The effect of inventory control: the closing inventory positions in simulations MQ and MQIC for the stock IBM in the 63-day period.

mechanism to simulation MQ across all sub-samples. In the OM simulation, market orders that cannot be matched due to the absence of a market are discarded. The difference in trading volume between the simulations OM and MQ indicates the amount of liquidity contributed by the electronic market-maker. The difference in volume is more significant for less liquid stocks (lower quartile sub-samples). For instance, the average volume is increased 84.10 percent from OM to MQ for the stocks in the lowest quartile sub-sample. To measure how actively the market-maker involves in the trading process, we calculate its participation rate, which is defined as the percentage of transactions (in share units) in which the market-maker involves as either a buyer or seller. The participation rate averages around 50 percent, substantially higher than what is reported in Hasbrouck & Sofianos (1993)—13 percent for a sample of 138 stocks. We notice that the participation rate is higher for the simulation MQ, where the market-maker aggressively intervenes to improve market quality. The participation rate is also found higher for less liquid stocks in which the dealer's role as a liquidity provider is more eminent. This result is consistent with the empirical evidence from Madhavan & Sofianos (1998) that the participation rate decreases with trading frequency.

The market-maker significantly improves various measures of market quality over the order matching mechanism. For example, the average spread of OM is 2.46 ticks (or one-eighth) while that of MQ is 1.22 ticks; the average for PC (transaction-to-transaction price change) of OM is 0.63 while that of MQ is 0.37 for the entire sample. Comparing with the NYSE execution, the simulations with the electronic market-maker yield comparable or sometimes better measures of market quality, particularly in the case of MQ. Among the three settings of the electronic market-maker, the one with the speculative motive gives the lowest market quality as the agent actively pursues its self-interest. Figure 4-4 shows sample executions of NYSE, MQ and SM+IC for IBM on November 1, 1990.

Variable	Total Sample	Quartile Sub-samples by Average Number of Daily Transactions			
		1 (lowest)	2	3	4 (highest)
Number of Securities	81	20	20	21	20
Daily profits (\$1,000)					
MQ	0.71 (577.81)	-1.33 (25.41)	-1.11 (45.09)	-1.50 (91.23)	6.87 (1,117.78)
MQ+IC	5.93 (977.72)	-0.51 (24.51)	-0.72 (30.19)	1.80 (95.64)	23.34 (1,965.30)
SM+IC	-8.16 (282.49)	-0.34 (12.20)	-0.32 (14.63)	-0.57 (58.19)	-31.81 (580.82)
Daily closing inventory (100-share lots)					
MQ	1,292.16 (2,216.49)	463.54 (436.70)	892.51 (754.94)	1,120.76 (1,197.83)	2,700.38 (3,844.37)
MQ+IC	891.08 (1,884.66)	452.62 (504.27)	543.17 (481.88)	788.58 (1,102.69)	1,785.09 (3,493.68)
SM+IC	431.09 (1,586.73)	204.62 (484.66)	191.90 (214.07)	407.47 (695.07)	921.57 (3,010.79)
Daily volume (100-share lots)					
OM	953.67 (2,325.75)	85.98 (113.37)	174.19 (161.85)	362.17 (381.88)	3,221.93 (3,853.52)
MQ	1,277.45 (3,024.28)	158.29 (170.71)	302.46 (225.75)	555.87 (468.18)	4,129.27 (5,085.57)
MQ+IC	1,227.48 (2,986.13)	140.06 (154.10)	266.72 (203.36)	512.32 (455.65)	4,026.58 (5,036.50)
SM+IC	1,261.26 (3,114.67)	143.37 (162.57)	271.84 (210.49)	520.15 (465.95)	4,146.73 (5,279.63)
Difference in volume (percent) $(V_{MQ} - V_{OM})/V_{OM}$	33.95	84.10	73.64	53.48	28.13
Avg. (inventory / volume)					
MQ	3.45 (6.33)	5.25 (7.98)	4.38 (8.38)	3.04 (4.07)	1.18 (1.63)
MQ+IC	2.88 (5.48)	5.47 (9.26)	3.11 (3.93)	2.19 (3.06)	0.82 (1.11)
SM+IC	1.15 (2.38)	2.06 (3.66)	1.09 (1.71)	1.22 (2.14)	0.22 (0.43)

Table 4.3: Trading profits, inventory and volume and market quality measures for the static order flow. MQ, IC and SM denote the market quality, inventory control and speculative motive modules, OM denotes the order matching mechanism, NYSE denotes the executions by the NYSE specialists.

		Quartile Sub-samples by Average Number of Daily Transactions				
Variable	Total Sample	1 (lowest)	2	3	4 (highest)	
Participation Rate						
	OM	0.00	0.00	0.00	0.00	0.00
	MQ	0.57	0.68	0.64	0.60	0.38
	MQ+IC	0.46	0.56	0.50	0.48	0.28
	SM+IC	0.48	0.59	0.52	0.51	0.31
Market quality measures (number of 1/8)						
Spread	NYSE	3.21	3.25	3.33	3.23	3.03
	OM	2.46	2.53	2.39	2.49	2.43
	MQ	1.22	1.23	1.23	1.26	1.17
	MQ+IC	1.68	1.79	1.73	1.70	1.49
	SM+IC	1.69	1.80	1.74	1.72	1.51
PC	NYSE	0.43	0.51	0.43	0.39	0.39
	OM	0.63	0.83	0.65	0.57	0.47
	MQ	0.37	0.42	0.38	0.36	0.33
	MQ+IC	0.51	0.60	0.53	0.49	0.42
	SM+IC	0.52	0.61	0.54	0.50	0.43
PR50	NYSE	1.17	1.50	1.26	1.00	0.39
	OM	3.63	3.94	3.37	3.90	3.42
	MQ	1.06	1.10	1.07	1.12	0.93
	MQ+IC	1.55	1.73	1.63	1.63	1.23
	SM+IC	1.57	1.75	1.66	1.66	1.24
PR250	NYSE	2.04	2.21	2.25	1.98	1.85
	OM	6.26	2.61	4.58	6.62	6.67
	MQ	1.74	1.64	1.77	1.91	1.60
	MQ+IC	2.43	2.44	2.65	2.71	2.11
	SM+IC	2.51	2.59	2.77	2.81	2.15

Table 4.3 (continued)

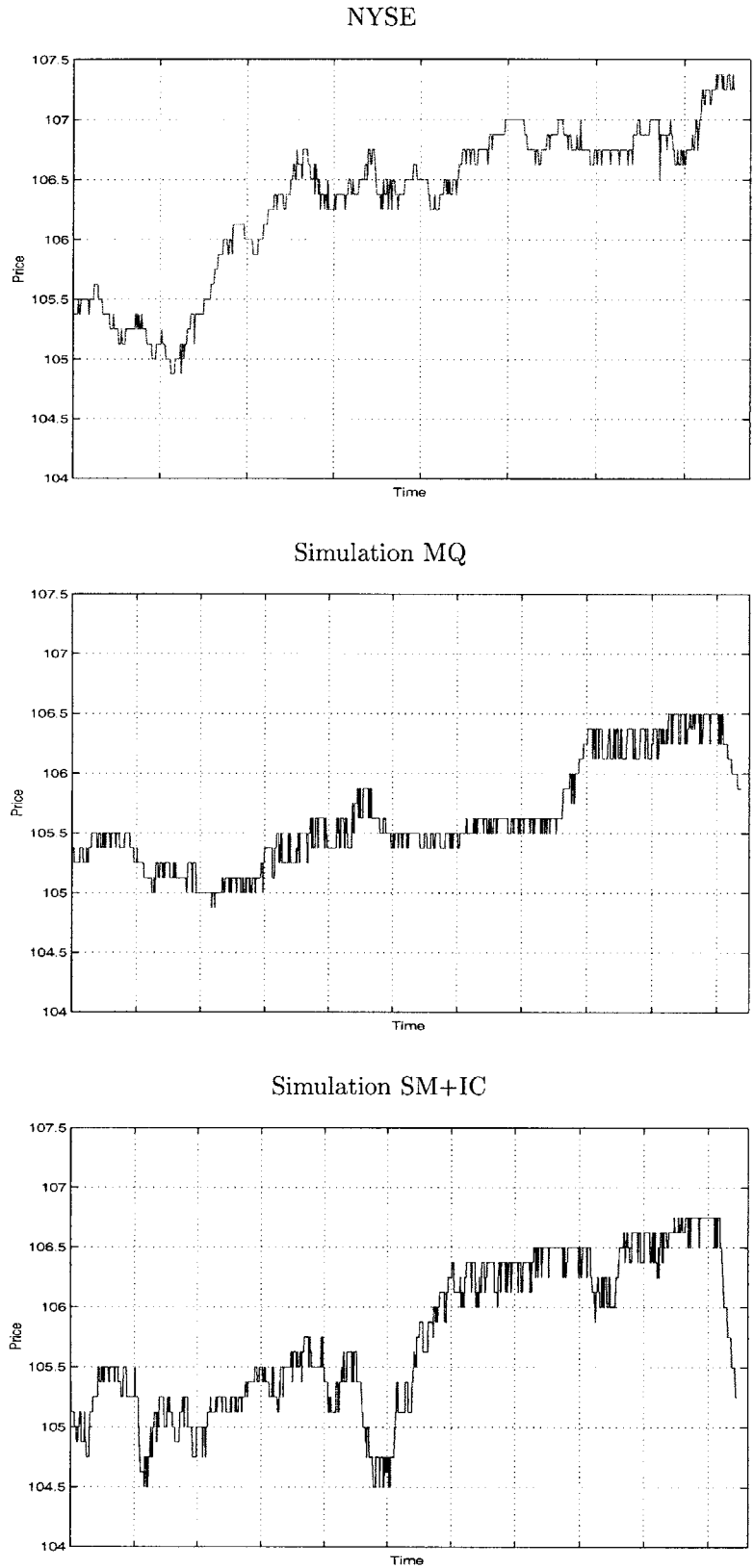


Figure 4-4: The transaction price for IBM on November 1, 1990 by NYSE, MQ and SM+ICC.

4.5.2 Modeled Order Flow

In the simulation of the modeled order flow, inventory control is found to be significantly more effective than in the case of the static order flow. The IC module reduces the variance of profits to reasonable levels: the setting MQ, MQ+IC and SM+IC give standard errors for daily profits of \$159,500, \$41,680 and \$3,240 respectively. Comparing MQ+IC and SM+IC, the latter consistently makes more profits than the former, across all sub-samples. The setting SM+IC yields an average daily profits of \$230. Sofianos (1995) reports a daily average profit of \$552 with a standard error of \$40 over a sample of 2,511 stocks at the NYSE in 1995. Although our results on profitability are statistically insignificant, it does signify the importance of inventory control as suggested by Garman (1976). Without the IC module, the daily fluctuation in profitability would make the market-making business inviable.

Average daily closing inventory drops substantially from 95,785 shares in MQ, to 17,648 shares in MQ+IC, and 2,931 shares in SM+IC. Comparing MQ+IC with SM+IC, the simulation without market quality constraints results in more effective inventory control because in upholding its affirmative obligations, the market-maker is more likely to accumulate unwanted inventory holdings. The ratio of absolute closing inventory to volume is another measure of a dealer's inventory holdings. The ratio averages 0.56 and 3.15 for MQ+IC and SM+IC respectively, compared with a value of 0.84 found in Hasbrouck & Sofianos (1993). The average participation rate ranges from 28 to 42 percent, lower than that found in the static order flow simulations, but higher than the 13 percent found in Hasbrouck & Sofianos (1993). It is still true that the market-maker participates more in less liquid stocks.

As for market quality measures, in general, we find that MQ performs better than MQ+IC, which in turn performs better than SM+IC.

4.6 Conclusions

This chapter presents computer simulations of a rule-based market-maker in market environments that feature static historical order flows and simulated order flows

Variable	Total Sample	Quartile Sub-samples by Average Number of Daily Transactions			
		1 (lowest)	2	3	4 (highest)
Number of Securities	81	20	20	21	20
Daily profits (\$1,000)					
MQ	-0.28 (159.50)	-0.98 (32.06)	-1.81 (55.11)	0.38 (60.61)	1.23 (301.334)
MQ+IC	-9.47 (41.68)	-0.18 (27.05)	-0.89 (30.66)	-0.06 (14.50)	-2.54 (69.93)
SM+IC	0.23 (3.24)	0.08 (4.44)	0.03 (2.82)	0.07 (2.62)	0.73 (2.82)
Daily closing inventory (100-share lots)					
MQ	957.85 (1,716.02)	498.04 (562.39)	627.45 (796.66)	814.73 (785.53)	1,846.72 (2,958.98)
MQ+IC	176.48 (510.71)	259.39 (443.56)	249.79 (588.59)	63.77 (176.11)	135.63 (653.02)
SM+IC	29.31 (59.50)	55.85 (114.40)	25.04 (19.87)	19.30 (14.75)	19.20 (20.23)
Daily volume (100-share lots)					
MQ	590.41 (1,128.01)	103.41 (113.41)	182.32 (142.59)	335.94 (291.93)	1,685.22 (1,775.36)
MQ+IC	604.40 (1,137.06)	104.75 (119.53)	180.77 (140.18)	323.99 (284.69)	1,744.08 (1,758.40)
SM+IC	562.38 (1,078.52)	90.03 (100.93)	154.37 (123.01)	282.12 (249.73)	1,661.76 (1,656.97)
Avg. (inventory / volume)					
MQ	5.61 (25.97)	12.34 (52.05)	5.65 (11.53)	3.62 (5.23)	1.57 (2.34)
MQ+IC	3.15 (13.62)	10.48 (23.21)	2.11 (14.82)	0.33 (1.28)	0.13 (0.56)
SM+IC	0.56 (4.18)	1.60 (8.44)	0.40 (1.65)	0.20 (0.59)	0.03 (0.05)

Table 4.4: Trading profits, inventory and volume and market quality measures for the modeled order flow. MQ, IC and SM denote the market quality, inventory control and speculative motive modules, OM denotes the order matching mechanism.

Variable	Total Sample	Quartile Sub-samples by Average Number of Daily Transactions			
		1 (lowest)	2	3	4 (highest)
Participation Rate					
MQ	0.42	0.55	0.45	0.42	0.28
MQ+IC	0.41	0.55	0.45	0.42	0.24
SM+IC	0.28	0.42	0.31	0.26	0.14
Market quality measures (number of 1/8)					
Spread MQ	1.13	1.15	1.13	1.14	1.09
MQ+IC	1.17	1.17	1.17	1.20	1.15
SM+IC	1.60	1.72	1.65	1.62	1.43
PC MQ	0.53	0.54	0.54	0.53	0.51
MQ+IC	0.57	0.56	0.57	0.59	0.56
SM+IC	0.87	0.97	0.90	0.86	0.74
PR50 MQ	1.09	1.08	1.13	1.16	0.99
MQ+IC	1.34	1.25	1.37	1.54	1.18
SM+IC	2.31	2.80	2.67	2.44	1.64
PR250 MQ	1.53	1.49	1.60	1.68	1.44
MQ+IC	2.42	1.92	2.46	2.89	2.21
SM+IC	3.84	4.39	5.54	4.70	3.11

Table 4.4 (continued)

modeled from transaction data. In the static order flow scenario, the abnormally high inventory levels in turn lead to unusually high variance in the market-maker's profits. Nevertheless, the affirmative obligations of the market-maker are fulfilled to a large extent. In the simulated order flow case, inventory control is more effective while the objectives pertaining to profits and market quality are also accomplished. Both types of simulations are able to replicate several findings from theoretical models and empirical studies in the literature.

The rule-based model presents a simple view of the market-making process, one that focuses on three primary objectives. It is not the goal of this chapter to provide a comprehensive model for market-making. Rather, the focus is on whether and how simple heuristic rules can accomplish certain functions or services provided by a market-maker. We also study the effect of combining different objectives and interactions among them. For instance, inventory control is found to be more effective with a speculative motive than under market quality constraints. The use of historical data, instead of pure simulated models, is another important characteristic of this study. We are able to relate and compare our findings with those from the real-world markets. For example, the profitability statistics and market quality measures are compared with those of the NYSE specialists.

However, many areas of this topic are yet to be addressed. Combining multiple objectives linearly may not be accurate in representing a dealer's overall goal. The modeling of the order flow is over-simplified because the time series properties of the data are largely ignored. But more importantly, a set of fixed and deterministic rules may be inadequate in dealing with markets that typically involve the learning of agents and evolution of trading strategies. A learning model for market-making, the key to this question, is discussed in the next chapter.

Chapter 5

An Adaptive Electronic Market-maker

The heuristic-based market-maker developed in Chapter 4 establishes a basic framework and identifies relevant variables for market-making models. In practice, however, such a static model is generally inadequate in a dynamical market environment. Parameters need to be tuned for different securities, market conditions and user preferences. For example, market-makers' target inventory, execution quality, and the responsiveness to order imbalance (by making quote adjustments) may vary according to their assessment of the current market conditions.

Many theoretical market-making models are developed in the context of stochastic dynamic programming. Bid and ask prices are dynamically determined to maximize some long term objectives such as expected profits or expected utility of profits. Models in this category include those of Ho & Stoll (1981), O'Hara & Oldfield (1986) and Glosten & Milgrom (1985). The main limitation of these models is that specific properties of the underlying processes (price process and order arrival process) have to be assumed in order to obtain a closed-form characterization of strategies.

This paper presents an adaptive learning model for market-making using reinforcement learning under a simulated environment. Reinforcement learning can be considered as a model-free approximation of dynamic programming. The knowledge of the underlying processes is not assumed but learned from experience. The goal of

the paper is to model the market-making problem in a reinforcement learning framework, explicitly develop market-making strategies, and discuss their performance. In the basic model, where the market-maker quotes a single price, we are able to determine the optimum strategies analytically and show that reinforcement algorithms successfully converge to these strategies. The major challenges of the problem are that the environment state is only partially observable and reward signals may not be available at each time step. The basic model is then extended to allow the market-maker to quote bid and ask prices. While the market-maker affects only the direction of a price in the basic model, it has to consider both the direction of the prices as well as the size of the bid-ask spreads in the extended model. The reinforcement algorithm converges to correct policies and effectively control the trade-off between profit and market quality in terms of the spread.

This paper starts with an overview of several important theoretical market-making models and an introduction of the reinforcement learning framework in Section 5.1. Section 5.2 establishes a reinforcement learning market-making model. Section 5.3 presents a basic simulation model of a market with asymmetric information where strategies are studied analytically and through the use of reinforcement learning. Section 5.4 extends the basic model to incorporate additional actions, states, and objectives for more realistic market environments.

5.1 Background

5.1.1 Market-making Models

The understanding of the price formation process in security markets has been one of the focal points of the market microstructure literature. There are two main approaches to the market-making problem. One focuses on the uncertainties of an order flow and the inventory holding risk of a market-maker. In a typical inventory-based model, the market-maker sets the price to balance demand and supply in the market while actively controlling its inventory holdings. The second approach at-

tempts to explain the price setting dynamics employing the role of information. In information-based models, the market-maker faces traders with superior information. The market-maker makes inferences from the orders and sets the quotes. This informational disadvantage is reflected in the bid-ask spread.

Garman (1976) describes a model in which there is a single, monopolistic, and risk neutral market-maker who sets prices, receives all orders, and clears trades. The dealer's objective is to maximize expected profit per unit time. Failure of the market-maker arises when the it runs out of either inventory or cash. Arrivals of buy and sell orders are characterized by two independent Poisson processes whose arrival rates depend on the market-maker's quotes. Essentially the collective activity of the traders is modeled as a stochastic flow of orders. The solution to the problem resembles that of the Gambler's ruin problem. Garman studied several inventory-independent strategies that lead to either a *sure* failure or a possible failure. The conditions to avoid a sure failure imply a positive bid-ask spread. Garman concluded that a market-maker *must* relate its inventory to the price-setting strategy in order to avoid failure. Amihud & Mendelson (1980) extends Garman's model by studying the role of inventory. The problem is solved in a dynamic programming framework with inventory as the state variable. The optimal policy is a pair of bid and ask prices, both as decreasing functions of the inventory position. The model also implies that the spread is positive, and the market-maker has a preferred level of inventory. Ho & Stoll (1981) studies the optimal behavior of a single dealer who is faced with a stochastic demand and return risk of his own portfolio. As in Garman (1976), orders are represented by price-dependent stochastic processes. However, instead of maximizing expected profit, the dealer maximizes the expected utility of terminal wealth which depends on trading profit and returns to other components in its portfolio. Consequently dealer's risks play a significant role in its price-setting strategy. One important implication of this model is that the spread can be decomposed into two components: a risk neutral spread that maximizes the expected profits for a set of given demand functions and a risk premium that depends on the transaction size and return variance of the stock. Ho & Stoll (1983) is a multiple-dealer version of Ho & Stoll (1981). The price-

dependent stochastic order flow mechanism is common in the above studies. All preceding studies only allow market orders traded in the market. O'Hara & Oldfield (1986) attempts to incorporate more realistic features of real markets into its analysis. The paper studies a dynamic pricing policy of a risk-averse market-maker who receives both limit and market orders and faces uncertainty in the inventory valuation. The optimal pricing strategy takes into account the nature of the limit and market orders as well as inventory risk.

Inventory-based models focus on the role of order flow uncertainty and inventory risk in the determination of the bid-ask spread. The information-based approach suggests that the bid-ask spread could be a purely informational phenomenon irrespective of inventory risk. Glosten & Milgrom (1985) studies the market-making problem in a market with asymmetric information. In the Glosten-Milgrom model some traders have superior (insider) information and others do not. Traders consider their information and submit orders to the market sequentially. The specialist, which does not have any information advantage, sets his prices, conditioning on all his available information such that the expected profit on any trade is zero. Specifically, the specialist sets its prices equal to the conditional expectation of the stock value given past transactions. Its main finding is that in the presence of insiders, a positive bid-ask spread would exist even when the market-maker is risk-neutral and make zero expected profit.

Most of these studies have developed conditions for optimality but provided no explicit price adjustment policies. For example, in Amihud & Mendelson (1980), bid and ask prices are shown to relate to inventory but the exact dependence is unavailable. Some analyses do provide functional forms of the bid/ask prices (such as O'Hara & Oldfield (1986)) but the practical applications of the results are limited due to stringent assumptions made in the models. The reinforcement learning models developed in this paper make few assumptions about the market environment and yield explicit price setting strategies.

5.1.2 Reinforcement Learning

Reinforcement learning is a computational approach in which agents learn their strategies through trial-and-error in a dynamic interactive environment. It is different from supervised learning in which examples or learning targets are provided to the learner from an external supervisor.¹ In a typical reinforcement learning problems the learner is not told which actions to take. Rather, it has to find out which actions yield the highest reward through experience. More interestingly, actions taken by an agent affect not only the immediate reward to the agent but also the next state in the environment, and therefore subsequent rewards. In a nutshell, a reinforcement learner interacts with its environment by adaptively choosing its actions in order to achieve some long-term objectives. Kaelbling & Moore (1996) and Sutton & Barto (1998) provide excellent surveys of reinforcement learning. Bertsekas & Tsitsiklis (1996) covers the subject in the context of dynamic programming.

Markov decision processes (MDPs) are the most common model for reinforcement learning. The MDP model of the environment consists of (1) a discrete set of *states* \mathcal{S} , (2) a discrete set of *actions* the agent can take \mathcal{A} , (3) a set of real-valued *rewards* \mathcal{R} or reinforcement signals, (4) a starting probability distribution over \mathcal{S} , (5) a transition probability distribution $p(s'|s, a)$, the probability of a state transition to s' from s when the agent takes action a , and (6) a reward probability distribution $p(r|s, a)$, the probability of issuing reward r from state s when the agent takes action a .

The MDP environment proceeds in discrete time steps. The state of the world for the first time step is drawn according to the starting probability distribution. Thereafter, the agent observes the current state of the environment and selects an action. That action and the current state of the world determine a probability distribution over the state of the world at the next time step (the transition probability distribution). Additionally, they determine a probability distribution over the reward issued to the agent (the reward probability distribution). The next state and a reward are chosen according to these distributions and the process repeats for the next time step.

¹Bishop (1995) gives a good introduction to supervised learning. See also Vapnik (1995), Vapnik (1998), and Evgeniou, Pontil & Poggio (2000).

The dynamics of the system are completely determined except for the action selection (or policy) of the agent. The goal of the agent is to find the policy that maximizes its long-term accumulated rewards, or *return*. The sequence of rewards after time step t is denoted as $r_t, r_{t+1}, r_{t+2}, \dots$; the return at the time t , R_t , can be defined as a function of these rewards, for example,

$$R_t = r_t + r_{t+1} + \dots + r_T;$$

or if rewards are to be discounted by a discount rate γ , $0 \leq \gamma \leq 1$:

$$R_t = r_t + \gamma r_{t+1} + \dots + \gamma^{T-t} r_T,$$

where T is the final time step of a naturally related sequence of the agent-environment interaction, or an *episode*.²

Because the environment is Markovian with respect to the state (*i.e.* the probability of the next state conditioned on the current state and action is independent of the past), the optimal policy for the agent is deterministic and a function solely of the current state.³ For reasons of exploration (explained later), it is useful to consider stochastic policies as well. Thus the policy is represented by $\pi(s, a)$, the probability of picking action a when the world is in state s .

Fixing the agent's policy converts the MDP into a Markov chain. The goal of the agent then becomes to maximize $E_\pi[R_t]$ with respect to π where E_π stands for the expectation over the Markov chain induced by policy π . This expectation can be broken up based on the state to aid in its maximization:

$$\begin{aligned} V^\pi(s) &= E_\pi[R_t | s_t = s], \\ Q^\pi(s, a) &= E_\pi[R_t | s_t = s, a_t = a], \end{aligned}$$

²These definitions and algorithms also extend to the non-episodic, or infinite-time, problems. However, for simplicity this paper will concentrate on the episodic case.

³For episodic tasks for which the stopping time is not fully determined by the state, the optimal policy may also need to depend on the time index. Nevertheless, this paper will consider only reactive policies or policies which only depend on the current state.

These quantities are known as value functions. The first is the expected return of following policy π out of state s . The second is the expected return of executing action a out of state s and thereafter following policy π .

There are two primary methods for estimating these value functions. The first is by Monte Carlo sampling. The agent executes policy π for one or more episodes and uses the resulting trajectories (the histories of states, actions, and rewards) to estimate the value function for π . The second is by temporal difference (TD) updates like SARSA (Sutton (1996)). TD algorithms make use of the fact that $V^\pi(s)$ is related to $V^\pi(s')$ by the transition probabilities between the two states (from which the agent can sample) and the expected rewards from state s (from which the agent can also sample). These algorithms use dynamic-programming-style updates to estimate the value function:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] . \quad (5.1)$$

α is the learning rate that dictates how rapidly the information propagates.⁴ Other popular TD methods include Q-learning (Watkins (1989), Watkins & Dayan (1992)) and TD(λ) (Watkins (1989), Jaakkola, Jordan & Singh (1994)). Sutton & Barto (1998) gives a more complete description of Monte Carlo and TD methods (and their relationship).

Once the value function for a policy is estimated, a new and improved policy can be generated by a policy improvement step. In this step a new policy π_{k+1} is constructed from the old policy π_k in a greedy fashion:

$$\pi_{k+1}(s) = \arg \max_a Q^{\pi_k}(s, a). \quad (5.2)$$

Due to the Markovian property of the environment, the new policy is guaranteed to be no worse than the old policy. In particular it is guaranteed to be no worse at every state individually: $Q^{\pi_{k+1}}(s, \pi_{k+1}(s)) \geq Q^{\pi_k}(s, \pi_k(a))$.⁵ Additionally, the sequence of

⁴The smaller the α the slower the propagation, but the more accurate the values being propagated.

⁵See p. 95 Sutton & Barto (1998)

policies will converge to the optimal policy provided sufficient exploration (*i.e.* that the policies explore every action from every state infinitely often in the limit as the sequence grows arbitrarily long). To insure this, it is sufficient to not exactly follow the greedy policy of Equation 5.2 but instead choose a random action ϵ of the time and otherwise choose the greedy action. This ϵ -greedy policy takes the form

$$\pi_{k+1}(s, a) = \begin{cases} 1 - \epsilon & \text{if } a = \arg \max_{a'} Q^{\pi_k}(s, a'), \\ \frac{\epsilon}{|\mathcal{A}|-1} & \text{otherwise.} \end{cases} \quad (5.3)$$

An alternative to the greedy policy improvement algorithm is to use an actor-critic algorithm. In this method, the value functions are estimated using a TD update as before. However, instead of jumping immediately to the greedy policy, the algorithm adjusts the policy towards the greedy policy by some small step size. Usually (and in this paper), the policy is represented by a Boltzmann distribution:

$$\pi_t(s, a) = \Pr[a_t = a | s_t = s] = \frac{\exp(w_{(s,a)})}{\sum_{a' \in \mathcal{A}} \exp(w_{(s,a')})} \quad (5.4)$$

where $w_{(s,a)}$ is a weight parameter of π corresponding to action a in state s . The weights can be adjusted to produce any stochastic policy which can have some advantages (discussed in the next section).

All three approaches are considered in this paper: a Monte Carlo method, SARSA (a temporal difference method) and an actor-critic method. Each has certain advantages. The Monte Carlo technique can more easily deal with long delays between an action and its associated reward than SARSA. However, it does not make as efficient use of the MDP structure as SARSA does. Therefore, SARSA does better when rewards are presented immediately whereas Monte Carlo methods do better with long delays.

Actor-critic has its own advantage in that it can find explicitly stochastic policies. For MDPs this may not seem to be as much an advantage. However, for most practical applications, the world does not exactly fit the MDP model. In particular, the MDP model assumes that the agent can observe the true state of the environ-

ment. However in cases like market-making that is not the case. While the agent can observe certain aspects (or statistics) of the world, other information (such as the information or beliefs of the other traders) is hidden. If that hidden information can affect the state transition probabilities, the model then becomes a partially observable Markov decision process (POMDP). In POMDPs, the ideal policy can be stochastic (or alternatively depend on all prior observations which is prohibitively large in this case). Jaakkola, Singh & Jordan (1995) discusses the POMDP case in greater details.

While none of these three methods are guaranteed to converge to the ideal policy for a POMDP model (as they are for the MDP model), in practice they have been shown to work well even in the presence of hidden information. Which method is most applicable depends on the problem.

5.2 A Reinforcement Learning Model of Market-making

The market-making problem can be conveniently modeled in the framework of reinforcement learning. In the following market-making problems, an episode can be considered as a trading day. Note that the duration of an episode does not need to be fixed. An episode can last an arbitrary number of time steps and conclude when the certain task is accomplished. The market is a dynamic and interactive environment where investors submit their orders given the bid and ask prices (or quotes) from the market-maker. The market-maker in turn sets the quotes in response to the flow of orders. The job of the market-maker is to observe the order flow, the change of its portfolio, and its execution of orders and set quotes in order to maximize some long-term rewards that depend on the its objectives (*e.g.* profit maximization and inventory risk minimization).

5.2.1 Environment States

The environment state includes market variables that are used to characterize different scenarios in the market. These are variables that are observed by the market-maker from the order flow, its portfolio, the trades and quotes in the market, as well as other market variables:

- Inventory of the market-maker — amount of inventory-holding by the market-maker.
- Order imbalance — excess demand or supply in the market. This can be defined as the share difference between buy and sell market or limit orders received within a period of time.
- Market quality measures — size of the bid-ask spread, price continuity (the amount of transaction-to-transaction price change), depth of a market (the amount of price change given a number of shares being executed), time-to-fill of a limit order, etc.
- Others — Other characteristics of the order flow, information on the limit order book, origin of an order or identity of the trader, market indices, prices of stocks in the same industry group, price volatility, trading volume, time till market close, etc.

In this paper, we focus on three fundamental state variables: inventory, order imbalance and market quality. The state vector is defined as

$$\mathbf{s}_t = (INV_t, IMB_t, QLT_t) ,$$

where INV_t , IMB_t and QLT_t denote the inventory level, the order imbalance, and market quality measures respectively. The market-maker's inventory level is its current holding of the stock. A short position is represented by a negative value and a long position by a positive value. Order imbalance can be defined in many ways. One possibility is to define it as the sum of the buy order sizes minus the sum

of the sell order sizes during a certain period of time. A negative value indicates an excess supply and a positive value indicates an excess demand in the market. The order imbalance measures the total order imbalance during a certain period of time, for example, during the last five minutes or from the last change of market-maker's quotes to the current time. Market qualities measure quantities including the bid-ask spread and price continuity (the amount of price change in a subsequent of trades). The values of INV_t , IMB_t and QLT_t are mapped into discrete values: $INV_t \in \{-M_{inv}, \dots, -1, 0, 1, \dots, M_{inv}\}$, $IMB_t \in \{-M_{imb}, \dots, -1, 0, 1, \dots, M_{imb}\}$, and $QLT_t \in \{-M_{QLT}, \dots, 1, 0, 1, \dots, M_{QLT}\}$. For example, a value of $-M_{inv}$ corresponds to the highest possible short position, -1 corresponds to the smallest short position and 0 represents an even position. Order imbalance and market quality measures are defined similarly.

5.2.2 Market-maker's actions

Given the states of the market, the market-maker reacts by adjusting the quotes, trading with incoming public orders, etc.. Permissible actions by the market-maker include the following:

- Change the bid price
- Change the ask price
- Set the bid size
- Set the ask size
- Others — Buy or sell, provide price improvement (provide better prices than the current market quotes).

The models in this paper focus on the determination of the bid and ask prices and assume fixed bid and ask sizes (*e.g.* one share). The action vector is defined as

$$\mathbf{a}_t = (\Delta BID_t, \Delta ASK_t) ,$$

where $\Delta BID_t = BID_t - BID_{t-1}$ and $\Delta ASK_t = ASK_t - ASK_{t-1}$, representing the change in bid and ask prices respectively. All values are discrete: $\Delta BID_t \in \{-M_{\Delta BID}, \dots, 0, \dots, M_{\Delta BID}\}$ and $\Delta ASK_t \in \{-M_{\Delta ASK}, \dots, 0, \dots, M_{\Delta ASK}\}$, where $M_{\Delta BID}$ and $M_{\Delta ASK}$ are the maximum allowable changes for the bid and ask prices respectively.

5.2.3 Reward

The reward signal is the agent's driving force to attain the optimal strategy. This signal is determined by the agent's objectives. Possible reward signals (and their corresponding objectives) include

- Change in profit (maximization of profit)
- Change in inventory level (minimization of inventory risk)
- Current market quality measures (maximization of market qualities)

The reward at each time step depends on the change of profit, the change of inventory, and the market quality measures at the current time step. The reward can be defined as some aggregate function of individual reward components. In its simplest form, assuming risk neutrality of the market-maker, the aggregate reward can be written as a linear combination of individual reward signals:

$$r_t = w_{pro}\Delta PRO_t + w_{inv}\Delta INV_t + w_{qit}QLT_t, \quad (5.5)$$

where w_{pro} , w_{inv} and w_{qit} are the parameters controlling the trade-off between profit, inventory risk and market quality; $\Delta PRO_t = PRO_t - PRO_{t-1}$, $\Delta INV_t = INV_t - INV_{t-1}$ and QLT_t are the change of profit, the change of inventory, and market quality measure respectively at time t . Note that the market-maker is interested in optimizing the *end-of-day* profit and inventory, but not the instantaneous profit and inventory. However, it is the market quality measures at each time step with which the market-maker is concerned in order to uphold the execution quality for all

transactions. Recall that the agent intends to maximize the total amount of rewards it receives. The total reward for an episode with T time steps is

$$\begin{aligned} R_T &= \sum_{t=1}^T r_t \\ &= w_{pro}PRO_T + w_{inv}INV_T + w_{qlt} \sum_{t=1}^T QLT_t. \end{aligned}$$

Here the market-maker is assumed to start with zero profit and inventory: $PRO_0 = 0$ and $INV_0 = 0$.

The market-maker can observe the variables INV_t and QLT_t at each time t , but not necessarily PRO_t . In most cases, the “true” value or a fair price of a stock may not be known to the market-maker. Using the prices set by the market-maker to compute the reward could incorrectly value the stock. Furthermore the valuation could induce the market-maker to raise the price whenever it has a long position and lower the price whenever it has a short position, so that the value of its position can be maximized. Without a fair value of the stock, calculating the reward as in Equation 5.5 is not feasible. In these cases, some proxies of the fair price can be considered. For example, in a market with multiple market-makers, other dealers’ quotes and execution prices can reasonably reflect the fair value of the stock. Similarly, the fair price may also be reflected in the limit prices from the incoming limit orders. Lastly, the opening and closing prices can be used to estimate the fair price. This approach is motivated by how the market is opened and closed at the NYSE. The NYSE specialists do not open or close the market at prices solely based on their discretion. Instead, they act as auctioneers to set prices that balance demand and supply at these moments. Consequently these prices represent the most informative prices given all information available at that particular time.

In the context of the reinforcement learning algorithm, the total reward for an episode is calculated as the difference between the the end-of-day and the beginning-of-day profit:

$$R_T = PRO_T - PRO_0 = PRO_T.$$

Unfortunately, the profit reward at each time step is still unavailable. One remedy is to assume zero reward at each $t < T$ and distribute all total reward to at $t = T$. An alternative approach is to assign the episodic average reward $r_t = R_T/T$ to each time step.

For this paper two approaches in setting the reward are considered. In the first case, we assume that the reward can be calculated as a function of the true price at each time step. However, the true price is still *not* observable as a state variable. In the second case, we only reveal the true price at the end of a training episode at which point the total return can be calculated.

5.3 The Basic Model

Having developed a framework for the market-maker, the next step is to create a market environment in which the reinforcement learner can acquire experience. The goal here is to develop a simple model that adequately simulates the strategy of a trading crowd given the quotes of a market-maker. Information-based models focusing on information asymmetry provide the basis for our basic model. In a typical information-based model, there is a group of informed traders or insiders who have superior information about the true value of the stock and a group of uninformed traders who possess only public information. The insiders buy whenever the market-maker's prices are too low and sell whenever they are too high given their private information; the uninformed simply trade randomly for liquidity needs. A single market-maker is at the center of trading in the market. It posts the bid and ask prices at which all trades transact. Due to the informational disadvantage, the market-maker always loses to the insiders while he breaks even with the uninformed.

5.3.1 Market Structure

To further illustrate this idea of asymmetric information among different traders, consider the following case. A single security is traded in the market. There are three types of participants: a monopolistic market-maker, insiders, and uninformed

traders. The market-maker sets *one* price, p^m , at which the next arriving trader has the option to either buy or sell *one* share. In other words, it is assumed that the bid price equals the ask price. Traders trade only with market orders. All orders are executed by the market-maker and there are no crossings of orders among traders. After the execution of an order, the market-maker can adjust its quotes given its knowledge of past transactions. In particular it focuses on the order imbalance in the market in determining the new quotes. To further simplify the problem, it is assumed that the stock position is liquidated into cash immediately after a transaction. Hence inventory risk is not a concern for the market-maker. This is a continuous market in which the market-maker executes the orders the moment when they arrive.

For simplicity, events in the market occur at discrete time steps. In particular, events are modeled as independent Poisson processes. These events include the change of the security's true price and the arrival of informed and uninformed orders.

There exists a true price p^* for the security. The idea is that there is an exogenous process that completely determines the value of the stock. The true price is to be distinguished from the market price, which is determined by the interaction between the market-maker and the traders. The price p^* follows a Poisson jump process. In particular, it makes discrete jumps, upward or downward with a probability λ_p at each time step. The size of the discrete jump is a constant 1. The true price, p^* , is given to the insiders but not known to the public or the market-maker.

The insider and uninformed traders arrive at the market with a probability of λ_i and $2\lambda_u$ respectively.⁶ Insiders are the only ones who observe the true price of the security. They can be considered as investors who acquire superior information through research and analysis. They compare the true price with market-maker's price and will buy (sell) one share if the true price is lower (higher) than the market-maker's price, and will submit no orders otherwise. Uninformed traders will place orders to buy and sell a security randomly. The uninformed merely re-adjust their portfolios to meet liquidity needs, which is not modeled in the market. Hence they simply submit buy or sell orders of one share randomly with equal probabilities λ_u .

⁶Buy and sell orders from the uninformed traders arrive at a probability of λ_u respectively.

All independent Poisson processes are combined together to form a new Poisson process. Furthermore, it is assumed that there is one arrival of an event at each time step. Hence, at any particular time step, the probability of a change in the true price is $2\lambda_p$, that of an arrival of an insider is λ_i , and that of an arrival of an uninformed trader is $2\lambda_u$. Since there is a guaranteed arrival of an event, all probabilities sum up to one: $2\lambda_p + 2\lambda_u + \lambda_i = 1$.

This market model resembles the information-based model, such as Glosten & Milgrom (1985), in which information asymmetry plays a major role in the interaction between the market-maker and the traders. The Glosten and Milgrom model studies a market-maker that sets bid and ask prices to earn zero expected profit given available information, while this model examines the quote-adjusting strategies of a market-maker that maximize sample average profit over multiple episodes, given order imbalance information. This model also shares similarities with the work of Garman (1976) and Amihud & Mendelson (1980) where traders submit price-dependent orders and the market-making problem is modeled as discrete Markov processes. But instead of inventory, here the order imbalance is used to characterize the state.

5.3.2 Strategies and Expected Profit

For this basic model, it is possible to compute the ideal strategies. We do this first, before presenting the reinforcement learning results for the basic model.

Closed-form characterization of an optimal market-making strategy in such a stochastic environment can be difficult. However, if one restricts one's attention to order imbalance in the market, it is obvious that any optimum strategy for a market-maker must involve the raising (lowering) of price when facing positive (negative) order imbalance, or excess demand (supply) in the market. Due to the insiders, the order imbalance on average would be positive if the market-maker's quoted price is lower than the true price, zero if both are equal, and negative if the quoted price is higher than the true price.

We now must define order imbalance. We will define it as the total excess demand since the last change of quote by the market-maker. Suppose there are x buy orders

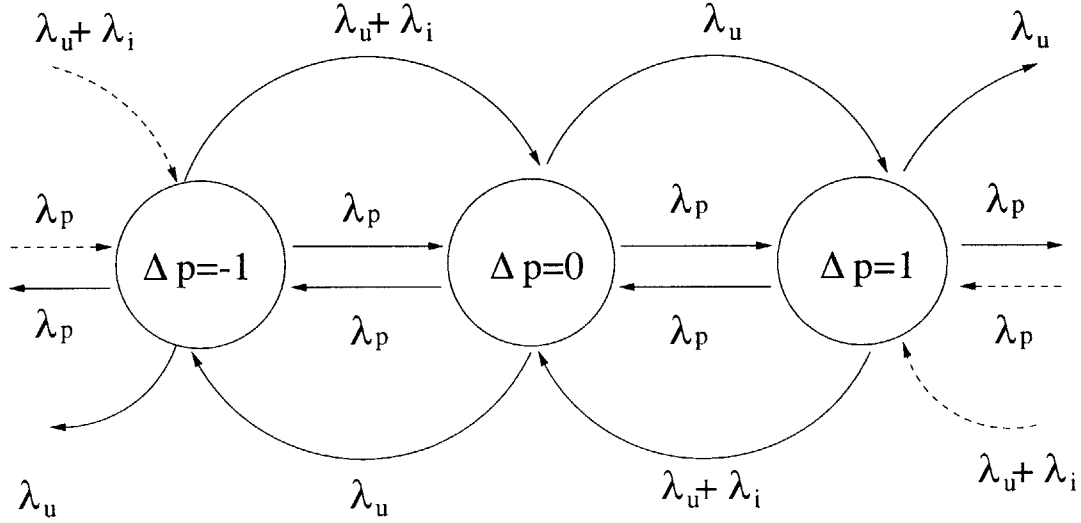


Figure 5-1: The Markov chain describing Strategy 1, imbalance threshold $M_{imb} = 1$, in the basic model.

and y sell orders of one share at the current quoted price; the order imbalance is $x - y$. One viable strategy is to raise or lower the quoted price by 1 whenever the order imbalance becomes positive or negative. Let us denote this as Strategy 1. Note that under Strategy 1, order imbalance can be -1 , 0 and 1 . To study the performance of Strategy 1, one can model the problem as a discrete Markov process.⁷ First we denote $\Delta p = p^m - p^*$ as the deviation of market-maker's price from the true price, and IMB as the order imbalance. A Markov chain describing the problem is shown in Figure 5-1. Suppose $\Delta p = 0$, p^* may jump to $p^* + 1$ or $p^* - 1$ with a probability of λ_p (due to the true price process); at the same time, p may be adjusted to $p + 1$ or $p - 1$ with a probability λ_u (due to the arrival of uninformed traders and the market-maker's policy). Whenever $p \neq p^*$ or $\Delta p \neq 0$, p will move toward p^* at a faster rate than it will move away from p^* . In particular, p always moves toward p^* at a rate of $\lambda_u + \lambda_i$, and moves away from p^* at a rate of λ_u . The restoring force of the market-maker's price to the true price is introduced by the informed trader, who observes the true price. In fact, it is the presence of the informed trader that ensures the existence of the steady-state equilibrium of the Markov chain.

⁷Lutostanski (1982) studies a similar problem.

Let q_k be the steady-state probability that the Markov chain is in the state where $\Delta p = k$. By symmetry of the problem, we observe that

$$q_k = q_{-k}, \quad \text{for } k = 1, 2, \dots \quad (5.6)$$

Focus on all $k > 0$ and consider the transition between the states $\Delta p = k$ and $\Delta p = k + 1$. One can relate the steady-state probabilities as

$$\begin{aligned} q_{k+1}(\lambda_p + \lambda_u + \lambda_i) &= q_k(\lambda_p + \lambda_u) \\ q_{k+1} &= q_k \frac{(\lambda_p + \lambda_u)}{(\lambda_p + \lambda_u + \lambda_i)} \quad \text{for } k = 0, 1, 2, \dots \end{aligned} \quad (5.7)$$

because a transition from $\Delta p = k$ to $\Delta p = k + 1$ is equally likely as a transition from $\Delta p = k + 1$ to $\Delta p = k$ at the steady state. By expanding from Equation 5.8 and considering Equation 5.6, the steady-state probability q_k can be written as

$$q_k = q_0 \left(\frac{\lambda_p + \lambda_u}{\lambda_p + \lambda_u + \lambda_i} \right)^{|k|}, \quad \forall k \neq 0.$$

All steady-state probabilities sum up to one

$$\begin{aligned} \sum_{k=-\infty}^{\infty} q_k &= 1, \\ q_0 + 2 \sum_{i=1}^{\infty} q_k &= 1, \\ q_0 &= \frac{\lambda_i}{2\lambda_p + 2\lambda_u + \lambda_i}. \end{aligned} \quad (5.8)$$

With the steady-state probabilities, one can calculate the expected profit of the strategy. Note that at the state $\Delta p = k$, the expected profit is $-\lambda_i|k|$ due to the informed

traders. Hence, the expected profit can be written as

$$\begin{aligned}
EP &= \sum_{k=-\infty}^{\infty} -q_k \lambda_i |k| & (5.9) \\
&= -2 \sum_{k=1}^{\infty} q_k \lambda_i |k| \\
&= -2q_0 \lambda_i \sum_{k=1}^{\infty} k \left(\frac{\lambda_p + \lambda_u}{\lambda_p + \lambda_u + \lambda_i} \right)^k \\
&= \frac{-2(\lambda_p + \lambda_u)(\lambda_p + \lambda_u + \lambda_i)}{(2\lambda_p + 2\lambda_u + \lambda_i)}.
\end{aligned}$$

The expected profit measures the average profit accrued by the market-maker per unit time. The expected profit is negative because the market-maker breaks even in all uninformed trades while it always loses in informed trades.

By simple differentiation of the expected profit, we find that EP goes down with λ_p , the rate of price jumps, holding λ_u and λ_i constant. The expected profit also decreases with λ_i and λ_u respectively, holding the other λ 's constant. However, it is important to point out that $2\lambda_p + 2\lambda_u + \lambda_i = 1$ since there is a guaranteed arrival of a price jump, an informed or uninformed trade at each time period. Hence changing the value of one λ while holding others constant is impossible. Let us express λ_p and λ_u in terms of λ_i : $\lambda_p = \alpha_p \lambda_i$ and $\lambda_u = \alpha_u \lambda_i$. Now the expected profit can be written as:

$$EP = \frac{-2(\alpha_p + \alpha_u)(\alpha_p + \alpha_u + 1)}{(2\alpha_p + 2\alpha_u + 1)^2}.$$

Differentiating the expression gives

$$\frac{\partial EP}{\partial \alpha_p} = \frac{\partial EP}{\partial \alpha_u} = \frac{-2}{(2\alpha_p + 2\alpha_u + 1)^3} < 0.$$

The expected profit increases with the relative arrival rates of price jumps and uninformed trades.

To compensate for the losses, the market-maker can charge a fee for each transaction. This would relate the expected profit to the bid-ask spread of the market-maker. It is important to notice that the strategy of the informed would be different if a fee

of x unit is charged. In particular, if a fee of x units is charged, the informed will buy only if the $p^* - p^m > x$ and sell only if $p^m - p^* > x$. If the market-maker charges the same fee for buy and sell orders, the sum of the fees is the spread. Let us denote the fee as a half of the spread, $SP/2$. The market-maker will gain $SP/2$ on each uninformed trade, and $|\Delta p| - SP/2$ (given that $|\Delta p| - SP/2 > 0$) on each informed trade. If the spread is constrained to be less than 2, then the informed traders' strategy does not change, and we can use the same Markov chain as before. Given SP and invoking symmetry, the expected profit can be written as

$$EP = \lambda_u SP - 2\lambda_i \sum_{k \geq SP/2}^{\infty} (k - SP/2)q_k.$$

If the market-maker is restricted to making zero profit, one can solve the previous Equation for the corresponding spread. Specifically, if $(1 - \lambda_i)(1 - 2\lambda_i) < 4\lambda_u$, the zero expected profit spread is

$$SP_{EP=0} = \frac{1 - \lambda_i}{2\lambda_u + \lambda_i(1 - \lambda_i)} < 2. \quad (5.10)$$

Although inventory plays no role in the market-making strategy, the symmetry of the problem implies a zero expected inventory position for the market-maker.

Strategy 1 reacts to the market whenever there is an order imbalance. Obviously this strategy may be too sensitive to the uninformed trades, which are considered noise in the market, and therefore would not perform well in high noise markets. This motivates the study of alternative strategies. Instead of adjusting the price when $IMB = 1$ or $IMB = -1$, the market-maker can wait until the absolute value of imbalance reaches a threshold M_{imb} . In particular, the market-maker raises the price by 1 unit when $IMB = M_{imb}$, or lowers the price by 1 unit when $IMB = -M_{imb}$ and resets $IMB = 0$ after that. The threshold equals 1 for Strategy 1. All these strategies can be studied in the same framework of Markov models. Figure 5-2 depicts the Markov chain that represents strategies with $M_{imb} = 2$. Each state is now specified by two state variables Δp and IMB . For example, at the state $(\Delta p = 1, IMB = -1)$,

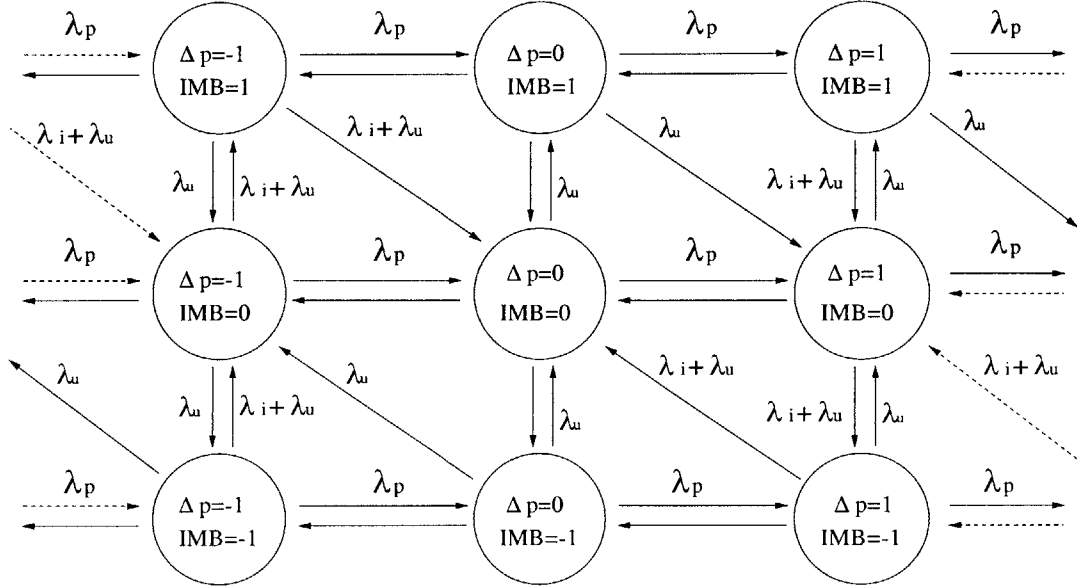


Figure 5-2: The Markov chain describing Strategy 2, with the imbalance threshold $M_{imb} = 2$ in the basic model.

a sell order (a probability of $\lambda_u + \lambda_i$) would move the system to $(\Delta p = 0, IMB = 0)$; a buy order (a probability of λ_u) would move the system to $(\Delta p = 1, IMB = 0)$; a price jump (a probability of λ_u) would move the system to either $(\Delta p = 0, IMB = -1)$ or $(\Delta p = 2, IMB = -1)$.

Intuitively, strategies with higher M_{imb} would perform better in noisier (larger λ_u) markets. Let us introduce two additional strategies: strategies with $M_{imb} = 2$ and $M_{imb} = 3$ and denote them as Strategies 2 and 3 respectively. The expected profit provides a criterion to choose among the strategies. Unfortunately analytical characterization of the expected profit for Strategies 2 and 3 is mathematically challenging. Instead of seeking explicit solutions in these cases, Monte Carlo simulations are used to compute the expected profits for these cases. To compare among the strategies, we set α_p to a constant and vary α_u and obtain the results in Figure 5-3. The expected profit for Strategy 1 decreases with the noise level whereas the expected profit for Strategies 2 and 3 increases with the noise level. Among the three strategies, we observe that Strategy 1 has the highest EP for $\alpha_u < 0.3$, Strategy 2 has the highest EP for $0.3 < \alpha_u < 1.1$ and Strategy 3 has the highest EP for $\alpha_u > 1.1$.

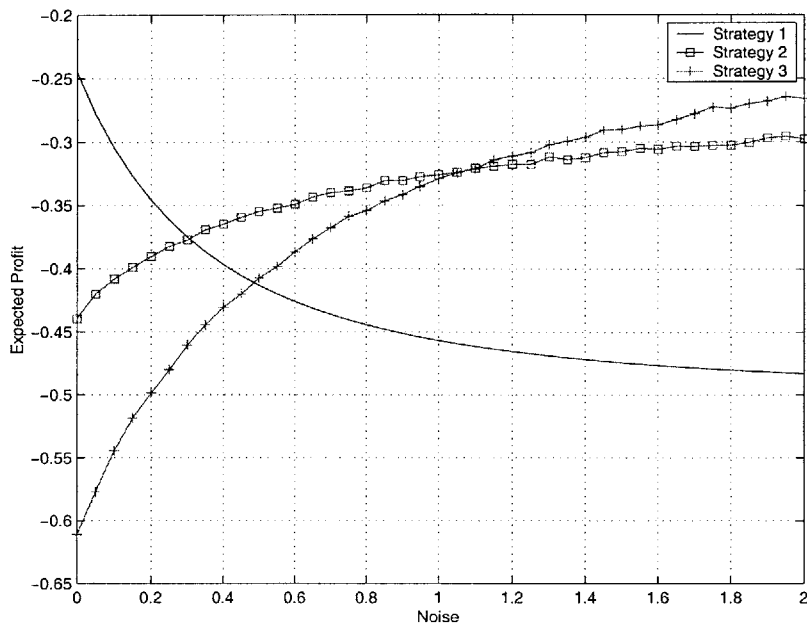


Figure 5-3: Expected profit for Strategies 1, 2, and 3 in the basic model.

5.3.3 Market-making with Reinforcement learning Algorithms

Our goal is to model an optimal market-making strategy in the reinforcement learning framework presented in Section 5.2. In this particular problem, the main focus is on whether reinforcement learning algorithms can choose the optimum strategy in terms of expected profit given the amount of noise in the market, α_u . Noise is introduced to the market by the uninformed traders who arrive at the market with a probability $\lambda_u = \alpha_u \lambda_i$.

For the basic model, we use the Monte Carlo and SARSA algorithms. Both build a value function $Q^\pi(s, a)$ and employ an ϵ -greedy policy with respect to this value function. When the algorithm reaches equilibrium, π is the ϵ -greedy policy of its own Q-function. The order imbalance $IMB \in \{-3, -2, \dots, 2, 3\}$ is the only state variable. Since market-maker quotes only one price, the set of actions is represented by $\Delta p^m \in \{-1, 0, 1\}$. Although the learning algorithms have the ability to represent many different policies (essentially any mapping from imbalance to price changes), in practice they converge to one of the three strategies as described in the previous

		ΔP		
		-1	0	1
<i>IMB</i>	-3	-0.86	-0.86	-0.86
	-2	-1.13	-1.41	-1.62
	-1	-1.16	-1.70	-2.61
	0	-1.62	-0.74	-1.62
	1	-2.63	-1.78	-1.13
	2	-1.67	-1.43	-1.15
	3	-0.91	-0.91	-0.91

(a) Strategy 1

		ΔP		
		-1	0	1
<i>IMB</i>	-3	-1.95	-1.96	-1.96
	-2	-1.97	-2.41	-3.23
	-1	-2.18	-2.03	-2.72
	0	-2.33	-1.79	-2.35
	1	-2.76	-2.05	-2.21
	2	-3.20	-2.37	-2.01
	3	-1.90	-1.89	-1.89

(b) Strategy 2

		ΔP		
		-1	0	1
<i>IMB</i>	-3	-10.10	-19.13	-19.09
	-2	-18.09	-10.67	-18.50
	-1	-12.18	-10.41	-12.82
	0	-11.13	-9.92	-11.56
	1	-11.99	-9.76	-12.11
	2	-18.57	-10.32	-18.14
	3	-19.17	-19.22	-10.54

(c) Strategy 3

Figure 5-4: Examples of Q-functions for Strategies 1, 2 and 3. The bold values are the maximums for each row showing the resulting greedy policy.

section. Figure 5-4 shows three typical Q-functions and their implied policies after SARSA has found an equilibrium. Take Strategy 2 as an example, it adjusts price only when IMB reaches 2 or -2:

Yet, this seemingly simple problem has two important complications from a reinforcement learning point-of-view. First the environment state is only partially observable. The agent observes the order imbalance but not the true price or the price discrepancy Δp . This leads to the violation of the Markov property. The whole history of observed imbalance now becomes relevant in the agent's decision making. For instance, it is more likely that the quoted price is too low when observing positive imbalance in two consecutive time steps than in just one time step. Formally, $Pr[\Delta p | IMB_t, IMB_{t-1}, \dots, IMB_0] \neq Pr[\Delta p | IMB_t]$. Nevertheless the order imbalance, a noisy signal of the true price, provides information about the hidden state variable Δp . Our model simply treats IMB as the state of the environment. However, convergence of deterministic temporal difference methods are not guaranteed for non-Markovian problems. Oscillation from one policy to another may occur. Deterministic policies such as those produced by the Monte Carlo method and SARSA may still yield reasonable results. Stochastic policies, which will be studied in the extended model, may offer some improvement in partially observable environments.

Second, since the true price is unobservable, it is infeasible to give a reward to the market-maker at each time step. As mentioned in Section 5.2.3, two possible remedies are considered. In the first approach, it is assumed that the true price is available for the calculation of the reward, but not as a state variable. Recall that the market-maker's inventory is liquidated at each step. The reward at time t is therefore the change of profit for the time step

$$r_t = \Delta PRO_t = \begin{cases} p_t^* - p_t^m & \text{for a buy order} \\ p_t^m - p_t^* & \text{for a sell order} \end{cases} \quad (5.11)$$

Alternatively, no reward is available during the episode, but only one final reward is given to the agent at the end of the episode. In this case, we choose to apply the Monte Carlo method and assign the end-of-episode profit per unit time, PRO_T/T , to

Experiment Number	Model	Learning Method	State(s) s_t	Actions a_t	Reward r_t
1	basic	SARSA	IMB_t	$\Delta P \in \mathcal{A}$	ΔPRO_t
2	basic	Monte Carlo	IMB_t	$\Delta P \in \mathcal{A}$	PRO_T/T
3	extended	actor-critic	(IMB_t, QLT_t)	$\Delta BID_t \in \mathcal{A}$ $\Delta ASK_t \in \mathcal{A}$	$w_{pro}\Delta PRO_t + w_{qll}QLT_t$
3a	extended	SARSA	(IMB_t, QLT_t)	$\Delta BID_t \in \mathcal{A}$ $\Delta ASK_t \in \mathcal{A}$	$w_{pro}\Delta PRO_t + w_{qll}QLT_t$
4	extended	actor-critic	(IMB_t, QLT_t)	$\Delta BID_t \in \mathcal{A}$ $\Delta ASK_t \in \mathcal{A}$	$- \Delta PRO_t $

Table 5.1: Details of the experiments for the basic and extended models.

all actions during the episode. Specifically, the reward can be written as

$$r_t = \frac{1}{T} \sum_{\tau=1}^T \Delta PRO_{\tau}. \quad (5.12)$$

Table 5.1 shows the options used for each of the experiments in this paper. The first two experiments are conducted using the basic model of this section, whereas the rest are conducted using the extended model of the next section that incorporates a bid-ask spread. Each experiment consists of 15 (10 for the extended model) separate sub-experiments, one for each of 15 (10) different noise levels. Each sub-experiment was repeated for 1000 different learning sessions. Each learning session ran for 2000 (1000 for the extended model) episodes each of 250 time steps.

5.3.4 Simulation Results

In the experiments, the primary focus is whether the market-making algorithm converges to the optimum strategy that maximizes the expected profit. In addition, the performance of the agent is studied in terms of profit and inventory at the end of an episode, PRO_T and INV_T , and average absolute price deviation for the entire

episode, $\overline{\Delta_p} = \frac{1}{T} \sum_{t=1}^T |p_t^m - p_t^*|$. The agent's end-of-period profit is expected to improve with each training episode, though remain negative. Its inventory should be close to zero. The average absolute price deviation measures how closely the agent estimates the true price. Figure 5-5 shows a typical realization of Experiment 1 in episodes 25, 100, 200 and 500. One can observe that the market-maker's price tracks the true price more closely as time progresses. Figures 5-6a and 5-6b show the realized end-of-period profit and inventory of the market-maker and their corresponding theoretical values. The profit, inventory and price deviation results all indicate that the algorithm converges at approximately episodes 500.

With the knowledge of the instantaneous reward as a function of the true price, the SARSA method successfully determines the best strategy under moderate noise level in the market. Figure 5-7 shows the overall results from Experiment 1. The algorithm converges to Strategy 1, 2, or 3, depending on the noise level. For each value of α_u , the percentages of the sub-experiments converging to strategies 1, 2 and 3 are calculated. One important observation is that the algorithm does not always converge to the same strategy, especially under high noise circumstances and around points of policies transitions. The agent's policy depends on its estimates of the Q-values, which are the expected returns of an action given a state. Noisier observations result in estimates with higher variability, which in turn transforms into the variability in the choice of the optimum policy. Noise naturally arising in fully observable environments is handled well by SARSA and Monte Carlo algorithms. However, the mismatch between fully observable modeling assumption and the partially observable world can cause variability in the estimates which the algorithms do not handle as well. This is responsible for the problems seen at the transition points.

The results show that the reinforcement learning algorithm is more likely to converge to Strategy 1 for small values of α ($\alpha < 0.25$) and Strategy 2 for higher values of α ($0.35 < \alpha < 1.00$). There are abrupt and significant points of change at $\alpha \simeq 0.30$ and $\alpha \simeq 1.00$ where the algorithm switches from one strategy to another. These findings are consistent with the theoretical predictions based on the comparison of the expected profits for the strategies (Figure 5-3). When the noise level α exceeds

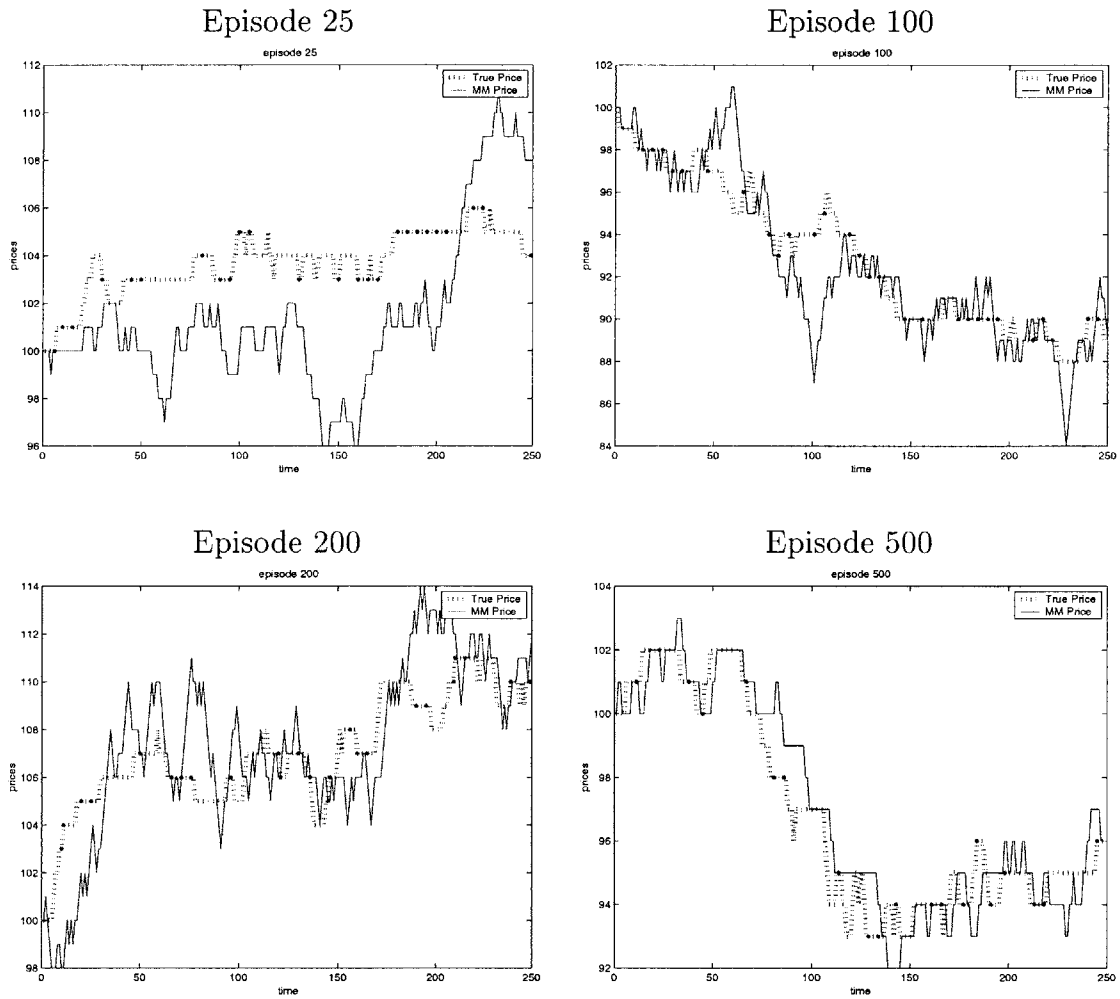


Figure 5-5: Episodes 25, 100, 200 and 500 in a typical realization of Experiment 1. The market-maker's price is shown in the solid line while the true price in dotted line. The maker's price traces the true price more closely over time.

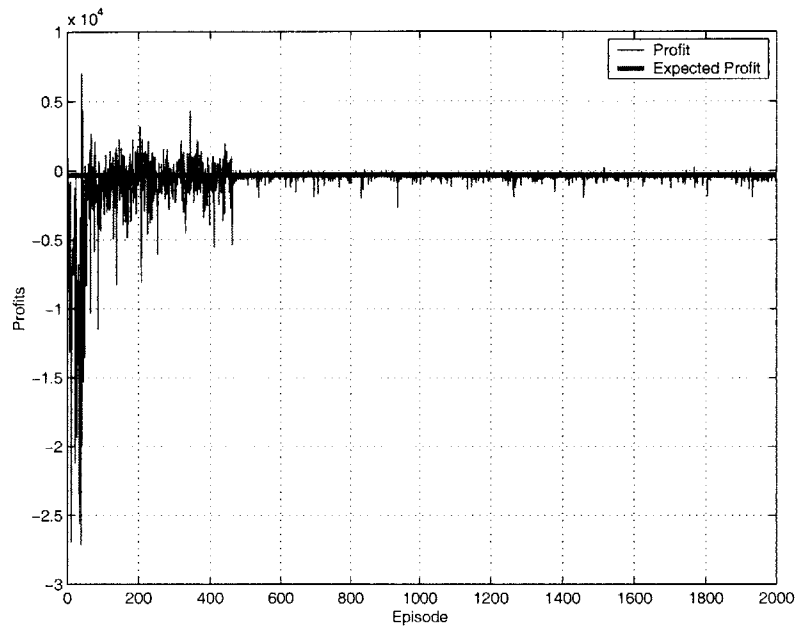


Figure 5-6a: End-of-episode profit and the corresponding theoretical value of the market-maker in Experiment 1 for a typical run with $\lambda_u = 0.25\lambda_i$. The algorithm converges around episode 500 when realized profit goes to its theoretical value.

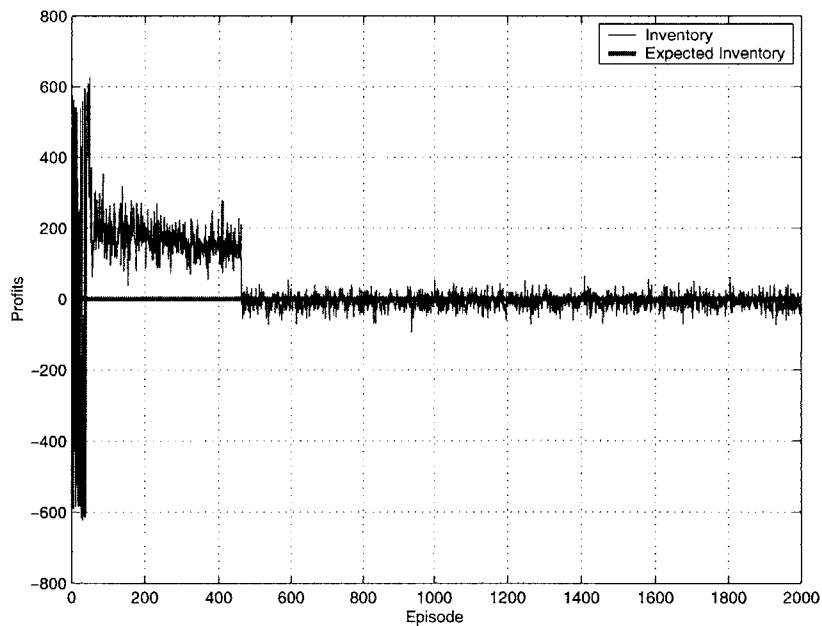


Figure 5-6b: End-of-episode Inventory and the corresponding theoretical value of the market-maker in Experiment 1 for a typical run with $\lambda_u = 0.25\lambda_i$. The algorithm converges around episode 500 when realized inventory goes to zero.

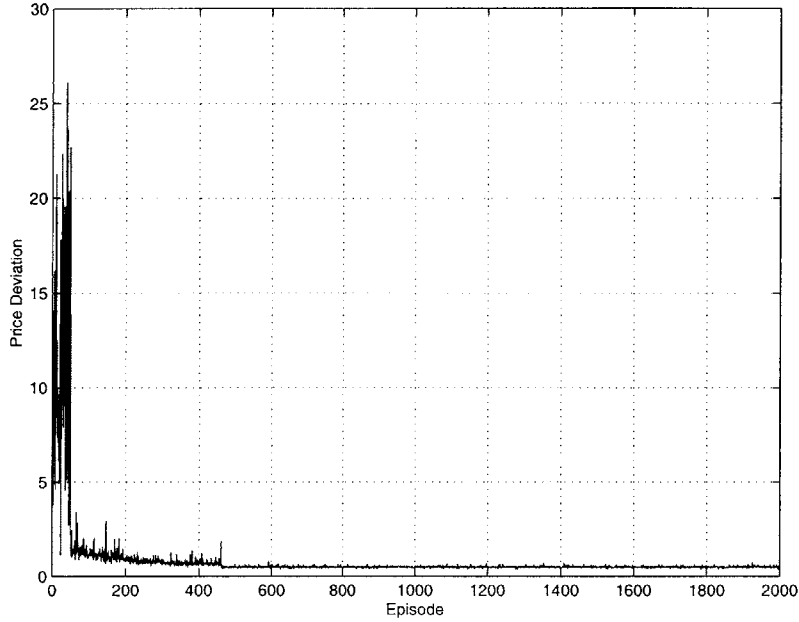


Figure 5-6c: Average absolute price deviation of the market-maker's quotation price from the true price in Experiment 1 for a typical run with $\lambda_u = 0.25\lambda_i$. The algorithm converges around episode 500 when the price deviation settles to its minimum.

the level of 1.0, the algorithm converges to Strategies 2 and 3 with an approximate likelihood of 80 and 20 percent respectively. According to the theoretical prediction, Strategy 3 would dominate the other two strategies when $\alpha_u > 1.1$. Unfortunately, the simulation fails to demonstrate this change of strategy. This is partially due to the inaccuracy in estimating the Q-function with the increasing amount of noise in the market. Furthermore, the convergence to Strategy 3 is intrinsically more difficult than that to Strategies 1 and 2. In order to recommend Strategy 3, the algorithm has to first decide to maintain the price for $IMB \leq 2$, effectively rejecting Strategies 1 and 2, and then estimate the relevant Q-values for $IMB = 3$; more exploration of the state space is necessary to evaluate Strategy 3.

What if no reward is given to the agent during the course of an episode? Experiment 2 is the same as Experiment 1 except for the differences in the learning method and the way reward is calculated. Even without the knowledge of the precise reward at each time step, the Monte Carlo algorithm still manages to shed some light on the choice of the optimum strategy. Figure 5-8 presents the percentages of the strategies

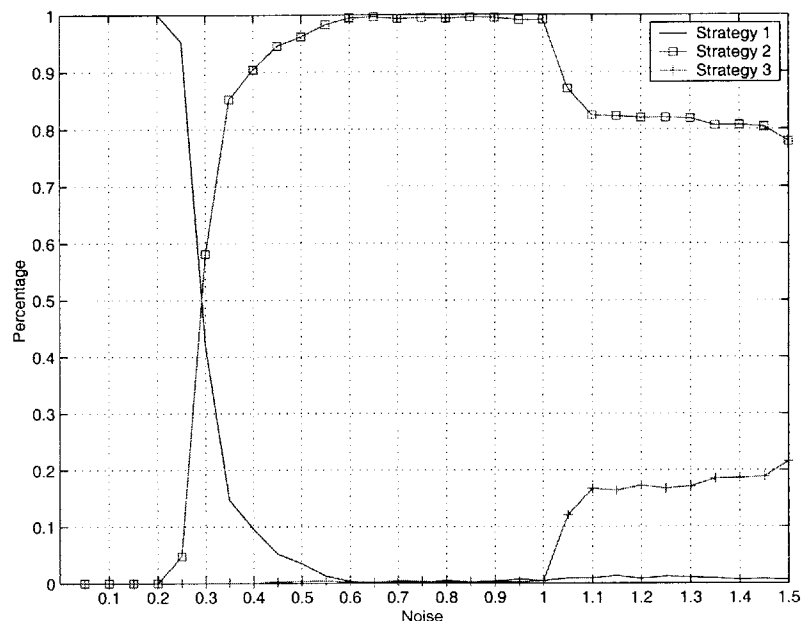


Figure 5-7: Percentages of SARSA simulations converges to Strategies 1, 2 and 3 in Experiment 1.

chosen for different noise levels. The algorithm is more likely to choose Strategy 1 for small values of α_u ($\alpha_u < 0.30$), Strategy 2 for moderate values of α_u ($0.30 < \alpha_u < 1$), and Strategy 3 for large values of α_u ($\alpha_u > 1$). This finding to some extent agrees with what the theory predicts.

Information on how much each action contributes to the total return is missing, unlike in the case of the SARSA method where the value of an action is more immediately realized. This is known as the *credit assignment problem*, first discussed by Minsky (1963). Even without the knowledge of the contribution of individual actions, the Monte Carlo method still works. This is because, on average, “correct actions” yield more reward and episodes with more “correct actions” consequently gather higher total return. But the missing reward information on individual action results in a higher variance in the estimation of values functions.

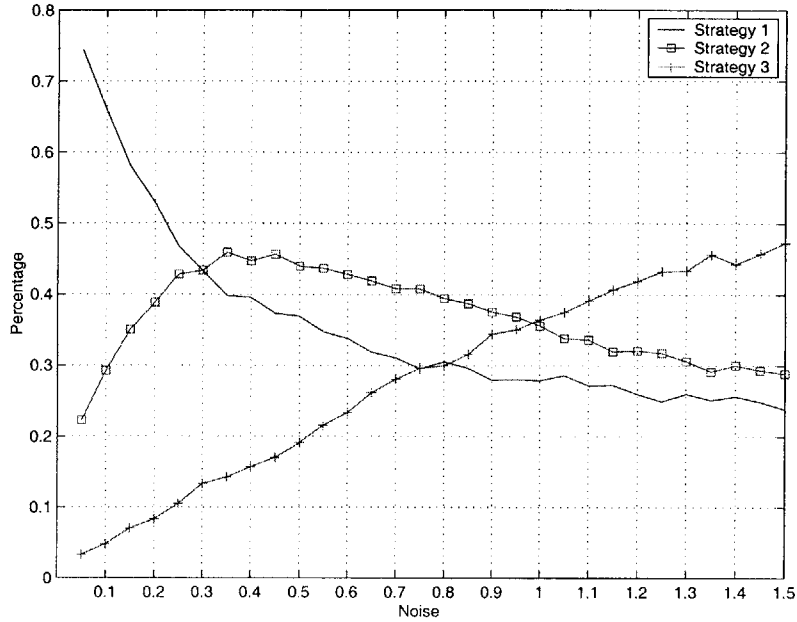


Figure 5-8: Percentages of the simulations of the Monte-Carlo method converges to Strategies 1, 2 and 3 in Experiment 2.

5.4 The Extended Model

The previous section demonstrates how reinforcement learning algorithms can be applied to market-making problems and successfully converge to optimum strategies under different circumstances. Although the basic model is useful because the experimental and theoretical results can be compared, one major limitation of the basic model is the equality of the bid and ask prices. Without the bid-ask spread the market-maker suffers a loss from the market due to the information disadvantage. A natural extension of the basic model is to let the market-maker quote bid and ask prices. This section studies a reinforcement learning strategy of the market-maker that balances the conflicting objectives of maximizing profit and market quality. Computer experiments demonstrate that the market-making agent successfully tracks the true price using the its bid and ask prices, and controls its average spread in a continuous scale.

To incorporate bid and ask prices to the model, the set of actions is augmented

to include the change of bid price and the change ask price:

$$(\Delta BID_t, \Delta ASK_t) \in \mathcal{A} \times \mathcal{A},$$

where $\mathcal{A} = \{-1, 0, 1\}$. Altogether there are nine possible actions. Now, to characterize a market scenario, the set of states should also include the spread, a measure of market quality. Specifically, the state vector becomes

$$\mathbf{s}_t = (IMB_t, SP_t),$$

where $IMB_t \in \{-1, 0, 1\}$ is the order imbalance and $SP_t = ASK_t - BID_t \in \{1, 2, 3, 4\}$ is the spread at time t . The spread also enters the reward function for the control of market quality maintained by a market-making algorithm. Recall that a market-maker may have multiple objectives. In the basic model, the market-maker only aims at maximizing profit. With spread added to the model, the market-maker would also need to consider the quality of market it provides. To balance between the two objectives, consider the following reward function that linearly combines the measures of profit and spread:

$$r_t = w_{pro}(\Delta PRO_t) + w_{qit}SP_t,$$

where the reward for profit now depends on the side of the order:

$$\Delta PRO_t = \begin{cases} ASK_t - p_t^* & \text{for a buy order} \\ p_t^* - BID_t & \text{for a sell order} \end{cases} \quad (5.13)$$

As for the reinforcement learning technique, an actor-critic method as described in Section 5.1.2 is used for the extended model. This algorithm allows the agent to expressly pick stochastic policies, which is important for two reasons. First, stochastic policies allow real-valued average spreads and profits. Essentially, this gives the agent more control over the fine-tuning of the trade-off between profit and market quality. For example, a policy which maintains a spread of 1 and 2 with equal probability of

1/2 would lead to an average spread of 1.5. Since the spread is intimately related to the profit (as shown as Section 5.3.2), the agent also indirectly controls the profit. Second, stochastic policies are particularly efficient in problems with partially observable states. This extended model pushes the partial observability of the environment much further.

The market-making agent should aim to set its bid and ask prices such that they enclose the observed true price: $BID_t \leq p_t^* \leq ASK_t$. Under this condition, the market-maker will gain from any trades (those of the uninformed traders) submitted to the market.

Three computer experiments are conducted for the extended model. In Experiments 3 and 3a, the market-maker simultaneously maximizes profit and market-quality. The weight w_{pro} is fixed but w_{qtl} is varied to demonstrate how spread can be fine-tuned. Experiment 3 applies the actor-critic method that yields stochastic policies; Experiment 3a considers the SARSA method that yields deterministic policies. It is interesting to compare the performance of the two approaches under partially observable environments.

Experiment 4 studies how one can directly control the profit by incorporating a target profit ΔPRO^* into the reward function:

$$r'_t = |\Delta PRO_t - \Delta PRO^*|.$$

The target profit ΔPRO^* is the desired average profit per unit time. Experiment 4 studies the particular case when $\Delta PRO^* = 0$. The resulting spread is the zero profit spread for the market-maker.

5.4.1 Simulation Results

As in the basic model, the performance of the market-maker is measured with variables including profit and average absolute price deviation. The end-of-episode profit

PRO_T measures how much the market-maker makes in an episode:

$$PRO_T = \sum_{t=1}^T \Delta PRO_t,$$

where ΔPRO_t is defined in Equation 5.13. The average absolute price deviation for an episode is calculated by considering both bid and ask prices:

$$\overline{\Delta_p} = \sum_{t=1}^T |BID_t - p_t^*| + |ASK_t - p_t^*|.$$

The episodic average spread for an episode is calculated as the average of the spread over time.

$$\overline{SP} = \frac{1}{T} \sum_{t=1}^T SP_t.$$

Figure 5-9 presents a typical run of Experiment 4. The accuracy in tracking the true price improves over the episodes. Figures 5-10a to 5-10d show the end-of-episode profit and inventory, average spread, and average absolute price deviation for a run of Experiment 3. The figures indicate that the algorithm converges approximately at episode 500.

To demonstrate the results of an actor-critic method, Figure 5-11 graphically depicts the details of a typical stochastic policy found in Experiment 3. The figure shows the probability distribution of actions in all twelve possible situations specified by the state vector (IMB, SP) . For each situation, the probabilities of the nine possible actions are shown as a grid of squares. The areas of the squares represent the probabilities of pairs of bid/ask actions under the policy. The bid/ask actions have been transformed into changes of the mid-quote, $(\Delta ASK + \Delta BID)/2$, and the changes of the spread, $\Delta ASK - \Delta BID$, for easier interpretation of the figure.

The policy adjusts the prices for two objectives: to react to the order imbalance and to control the spread. It behaves correctly by reducing, maintaining, and raising bid/ask prices under negative, zero, and positive imbalance respectively, for cases of $SP = 1, 2, 3$. For the case when $SP = 4$, order imbalance is ignored (*i.e.* the

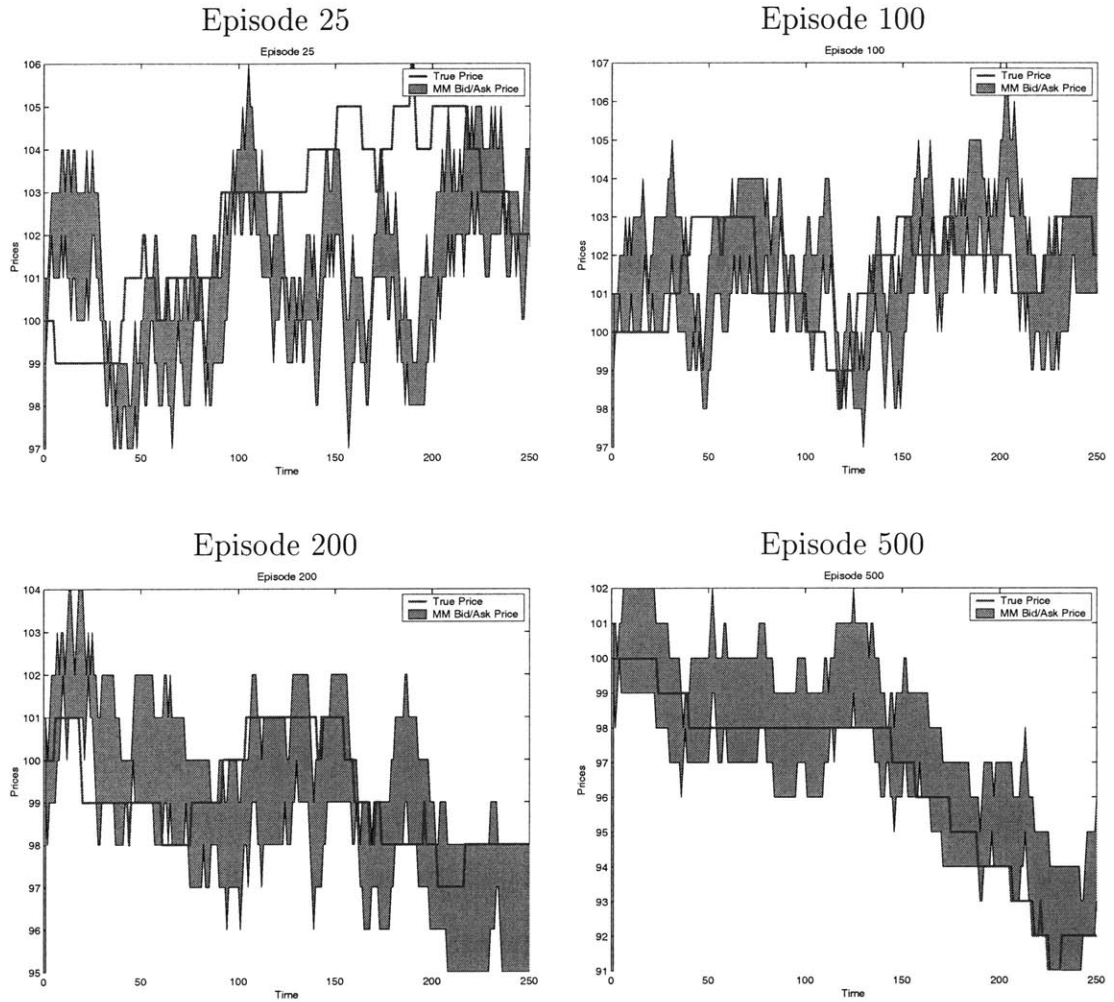


Figure 5-9: Episodes 25, 100, 200 and 500 in a typical realization of Experiment 3 with $w_{qt} = 0.1$. The bid and ask prices are shown in the shaded area, and the true price in the single solid line. The algorithm shows improvement in tracing the true price with the bid and ask prices over time.

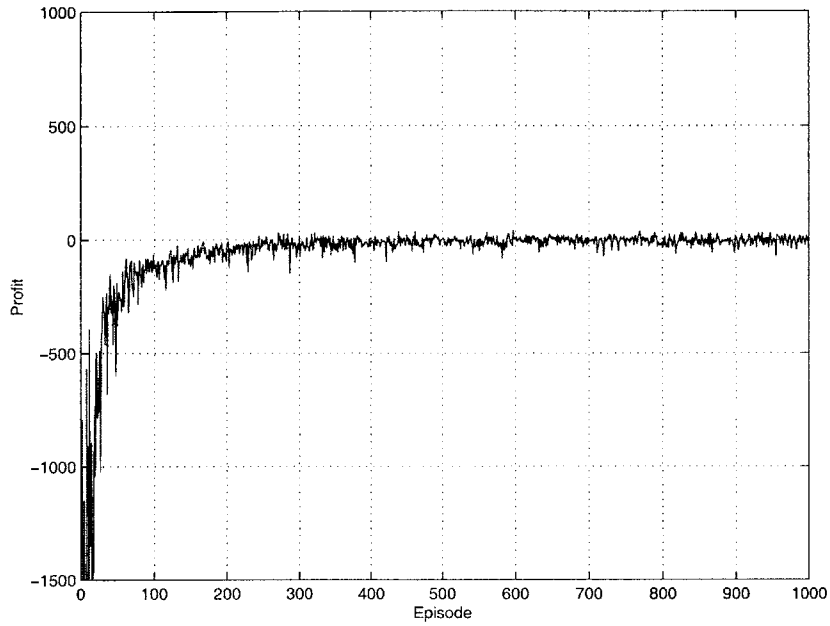


Figure 5-10a: End-of-episode profit, PRO_T , of a typical epoch of Experiment 3 with $w_{qll} = 0.1$.

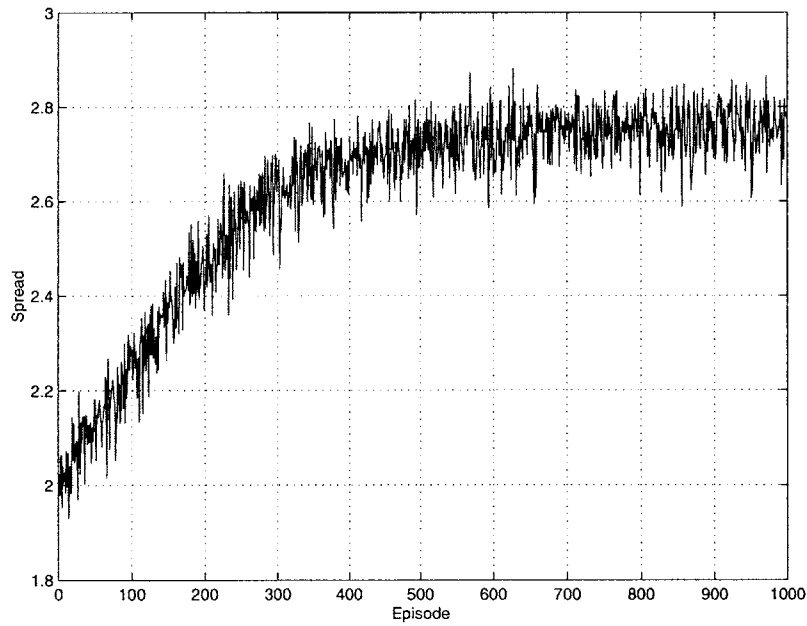


Figure 5-10b: Episodic average spread, \overline{SP} , of a typical epoch of Experiment 3 with $w_{qll} = 0.1$.

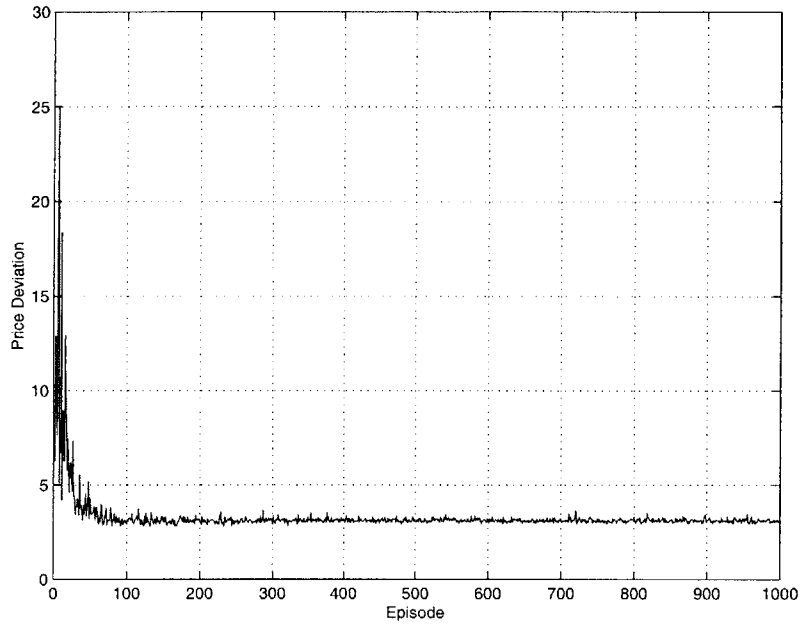


Figure 5-10c: Episodic average absolute price deviation, $\overline{\Delta_p}$, of a typical epoch of Experiment 3 with $w_{qit} = 0.1$.

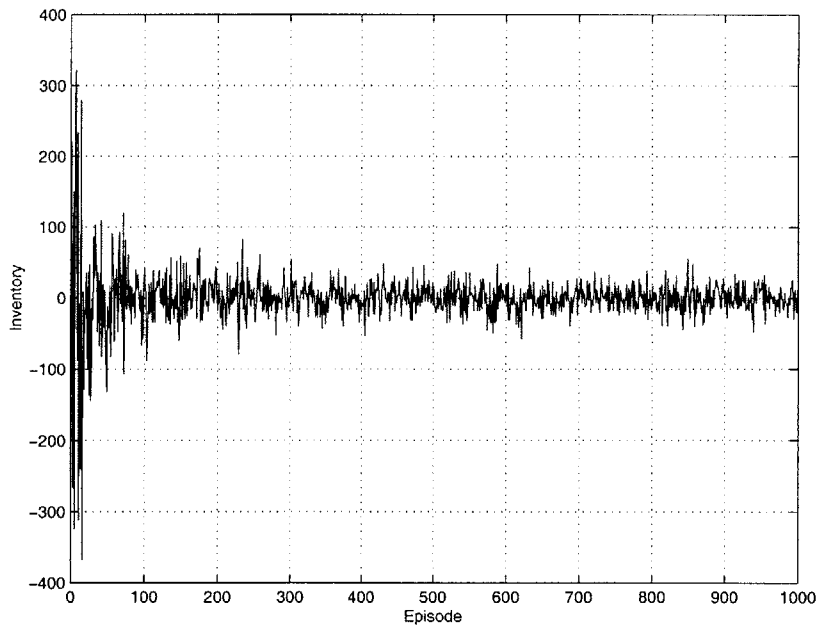


Figure 5-10d: End-of-episode inventory of a typical epoch of Experiment 3 with $w_{qit} = 0.1$.

adjustment of the mid-quote is not biased towards any direction). On the other hand, the policy tends to increase the spread for $SP = 1$, maintain or slightly increase the spread for $SP = 2$, and decrease the spread for $SP = 3, 4$. The mean and median spread resulting from this policy are both approximately 2.7.

By varying the spread parameter $w_{q,t}$, we can control the spread of the policy learned by either SARSA or actor-critic. The spread, as shown in Figure 5-12, decreases with an increasing value of $w_{q,t}$ in Experiments 3 and 3a. For each $w_{q,t}$, the mean, median and deciles of the average episodic spread are shown. The variance of the average spread is due to the stochastic nature of the algorithm, randomness in the order flow and true price process, and the imperfect state information. Comparing the results of Experiments 3 and 3a, we notice that stochastic policies yield a much lower variance for the resulting spread than deterministic policies do. As we expect, stochastic policies are better able to control partially observable environments.

Figure 5-13 presents the relationship between spread and profit in Experiment 3. Profit increases with spread as is expected. The results also indicate that to make a zero profit, the market-maker must maintain a spread approximately between 2.8 and 2.9.

In Experiment 4, the algorithm successfully enforces a zero profit in the market. The mean, median and standard error of profit are -0.48, 2.00 and 2.00 respectively, while the mean and median of spread are 2.83 and 2.84 respectively. This result agrees with the results from Experiment 3. The empirical distributions of profit and spread are shown in Figures 5-14a and 5-14b.

5.5 Conclusions

This paper presents an adaptive learning model for market-making in a reinforcement learning framework. We develop explicit market-making strategies, achieving multiple objectives under a simulated environment.

In the basic model, where the market-maker quotes a single price, we are able to determine the optimum strategies analytically and show that the reinforcement

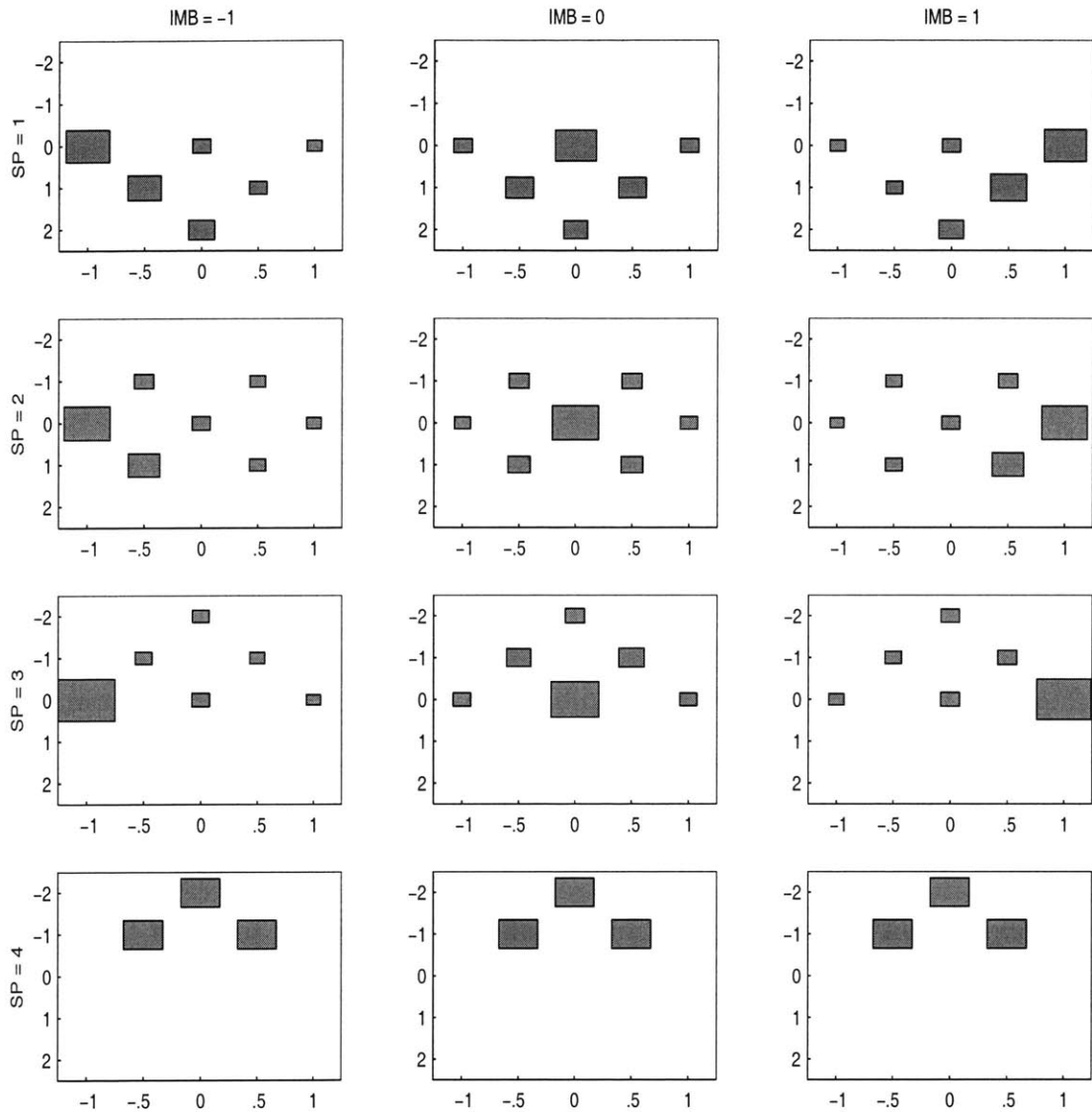
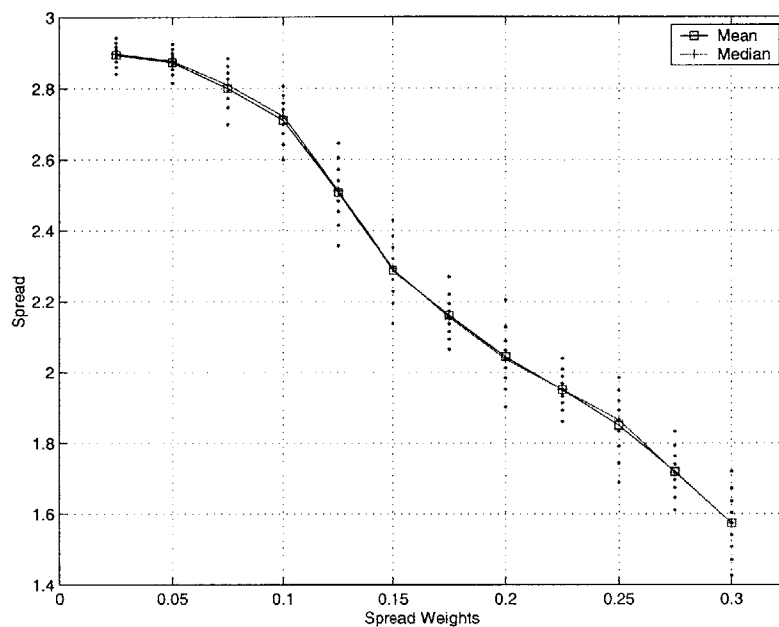


Figure 5-11: Conditional probability distribution of actions given imbalance and spread in a typical epoch of Experiment 3 with $w_{q|t} = 0.1$. Each probability distribution is depicted as a grid of squares whose areas represent the actual probability of pairs of bid/ask actions. In each panel, the change of the mid-quote and the change of the spread are shown on x-axis and y-axis respectively. For example, the panel at the third row and first column shows the conditional probability $\Pr(\mathbf{a} = \mathbf{a}' | IMB = -1, SP = 3)$. The action $\mathbf{a}' = (\Delta BID = -1, \Delta ASK = -1)$, which is equivalent to a change of mid-quote of -1 and a change of spread of 0, has the highest probability among all actions. In general, areas that appear in the upper (lower) portion of the panel represent a tendency to reduce (raise) the spread; areas that appear to the left (right) of the panel represent a tendency to decrease (increase) the mid-quote price.

Experiment 3: Stochastic Policies



Experiment 3a: Deterministic Policies

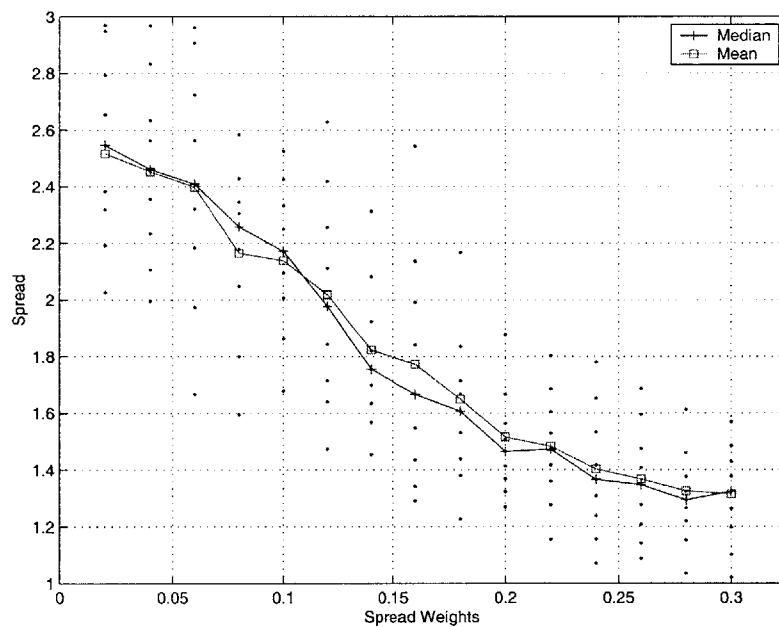


Figure 5-12: Spread weight versus episodic average spread in Experiment 3 and 3a. The deciles, median and mean of average episodic spread, \overline{SP} , of all the episodes over all epochs, are shown for different values of w_{qit} . For both experiments, the spread decreases with the weight parameter, but the variance of the spread is much lower for the actor-critic method that yields stochastic policies.

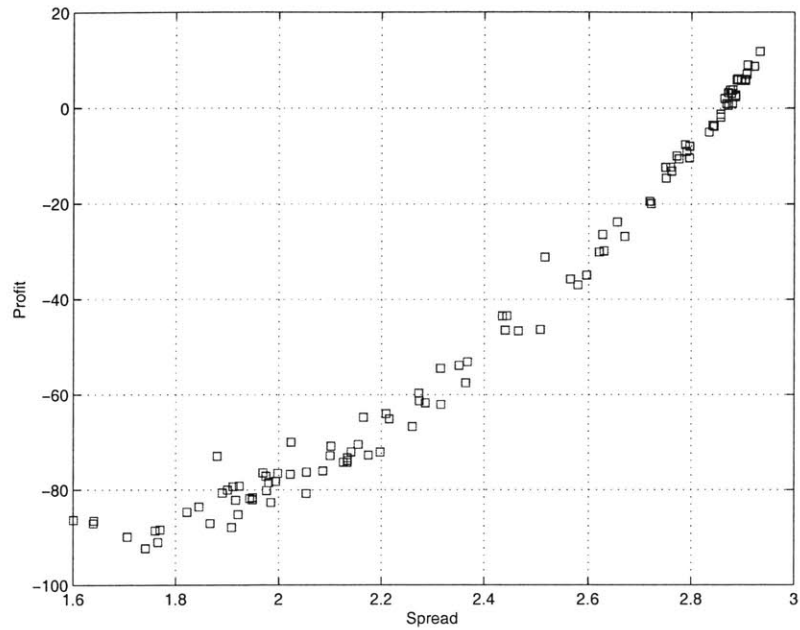


Figure 5-13: Episodic average spread versus end-of-period profit in Experiment 3. The figure presents the average \overline{SP} versus the average PRO_T over all episode of an epoch. The profit goes up with the spread.

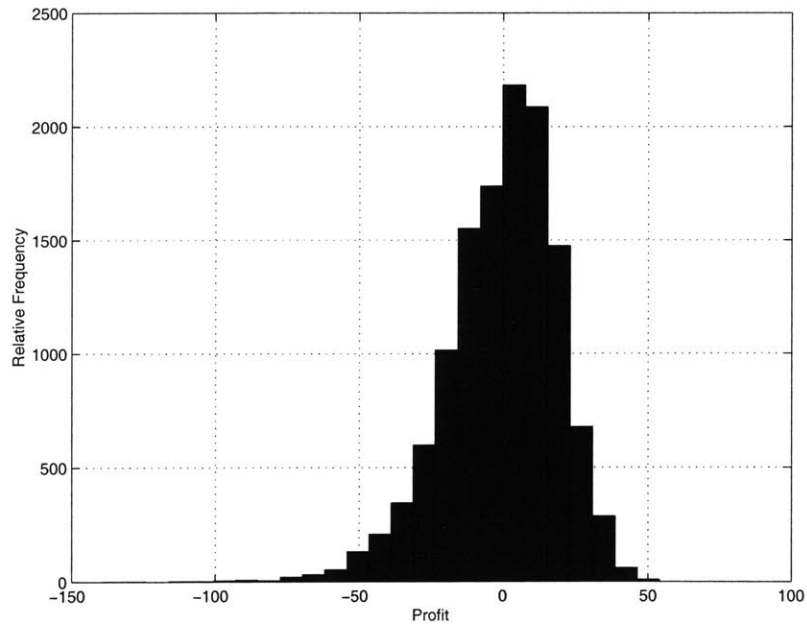


Figure 5-14a: Empirical distribution of end-of-episode profit, PRO_T , in Experiment 4. The mean, median and standard error of the profit is -0.48, 2.00 and 19.36 respectively.

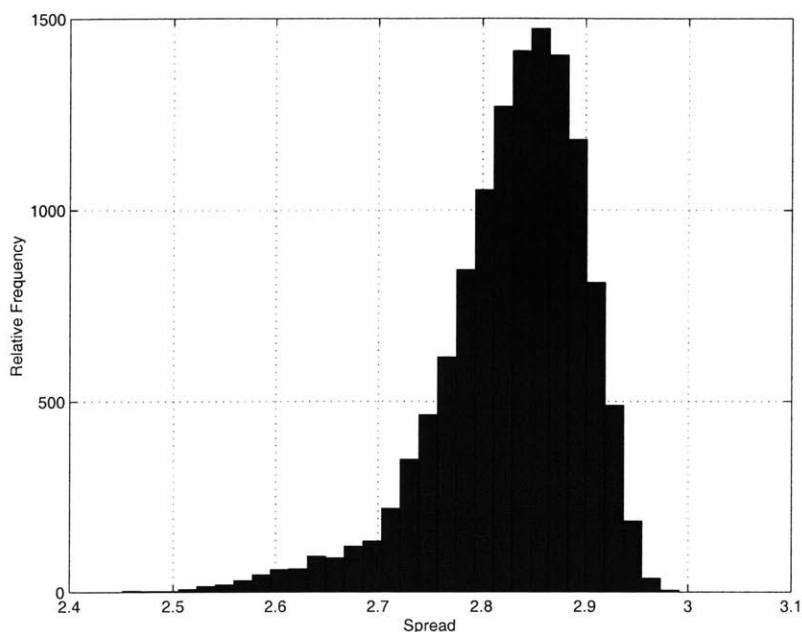


Figure 5-14b: Empirical distribution of average episodic spread, \overline{SP} in Experiment 4. The mean, median and standard error of the spread is 2.83, 2.84 and 0.07 respectively.

algorithms successfully converge to these strategies. In the SARSA experiment, for example, given the reward at each time step, a significant percentage of the epochs converges to the optimum strategies under moderate noise environments. It is also important to point out that the algorithm does not always converge to a single strategy, primarily due to the partial observability of the problem.

The basic model is then extended to allow the market-maker to quote bid and ask prices. While the market-maker only controls the direction of the price in the basic model, it has to consider both the direction of the price and the size of the bid-ask spread in the extended model. The actor-critic algorithm generates stochastic policies that correctly adjust bid/ask prices with respect to order imbalance and effectively control the trade-off between the profit and the spread. Furthermore, the stochastic policies are shown to out-perform deterministic policies in achieving a lower variance of the resulting spread.

Reinforcement learning assumes no knowledge of the underlying market environment. This means that it can be applied to market situations for which no explicit

model is available. We have shown initial success in bringing learning techniques to building market-making algorithms in a simple simulated market. We believe that it is ideal to use the agent-based approach to address some of the challenging problems in the study of market microstructure. Future extensions of this study may include the setup of more realistic and complex market environments, the introduction of additional objectives to the market-making model, and the refinement of the learning techniques to deal with issues such as continuous state variables.

Chapter 6

Conclusions

In this thesis, various simulated and experimental markets are constructed to study different aspects of markets, with emphasis on market structures, trading mechanisms, and the learning of the agents. Results from experimental and simulated markets are compared and contrasted. The market structure and economic environments of the simulations are carefully designed. The market setup of the simulations in Chapter 2, for instance, follows from that of the corresponding experimental markets. Some simulations are calibrated and parameterized using historical data. Specifically, the modeled order flow in Chapter 4 is derived from the NYSE TORQ data.

In the computational markets with empirical Bayesian traders, we are able to show that simple learning mechanisms enable software agents to aggregate and disseminate information through the trading process. The results are shown to be consistent with the findings from the corresponding experimental markets. The use of artificial agents also enables us to perform new experiments that are impossible in human-based markets—the cases of momentum and nearest-neighbor traders.

The role of markets as an efficient information aggregator is further examined in a series of market experiments with human subjects. Empirical evidence suggests that these markets successfully aggregate diverse preferences of the human traders on some virtual consumer products, and produce consistent forecast for the potential market share. An equilibrium model and simulations with artificial agents are proposed to offer some conjectures on the working of these markets.

Lastly the feasibility of automated market-making is studied through the use of heuristic rules and an adaptive learning model. The rule-based agent successfully supplies liquidity to the market and maintains high market quality. Simple rules regarding inventory control and speculative motive are also shown to be effective. On the other hand, the adaptive agent is able to learn in real-time and balance dual objectives under a simulated environment.

Agent-based models are also ideally suited to address some of the most challenging issues in market microstructure: What are the relative merits of a monopolistic market-maker versus multiple dealers? What are the likely effects of decimalization on the bid-ask spreads and volume? Do “circuit breakers” ameliorate or exacerbate market volatility? How do we define “liquidity”? And what are the social-welfare implications of the growing number of ECN’s and the corresponding fragmentation that they create? Although each of these issues has been subjected to theoretical and empirical scrutiny, the complexities of the interactions among market participants and institutional structure are so great that very few practical implications can be expected from such studies. Agent-based models provide a natural alternative, and we plan to explore these issues more fully in future research.

This thesis is a growing research program in which computer-simulated market interactions of AI-agents are yielding many insights into complex issues such as learning dynamics, the evolution of market structure, and the nature of human intelligence in an economic context. We hope to have provided a bridge between ad hoc learning models and market experiments with human subjects.

Future agent-based simulations need not be restricted to AI-agents. We believe that there are many interesting experiments to be performed with human and software agents combined, and these “mixed” experiments may provide new methods for exploring the nature of human cognition in economic settings. With the recent plethora of electronic day-trading companies and corresponding technologies, we may soon see AI-agents acting as broker/dealers for human clients, hence an agent-based modeling approach to financial markets may have practical implications as well.

Bibliography

- Amihud, Y., Ho, T. & Schwartz, R. (1985), *Market Making and the Changing Structure of the Securities Industry*, Lexington Books, Lexington, MA.
- Amihud, Y. & Mendelson, H. (1980), 'Dealership market: Market-making with inventory', *Journal of Financial Economics* **8**, 31–53.
- Arifovic, J. (1996), 'The behavior of the exchange rate in the genetic algorithm and experimental economies', *Journal of Political Economy* **104**, 510–541.
- Arthur, W. B., Holland, J., LeBaron, B., Palmer, R. & Tayler, P. (1997), Asset pricing under endogenous expectations in an artificial stock market, *in* W. B. Arthur, S. Durlauf & D. Lane, eds, 'The Economy as an Evolving Complex System II', Addison-Wesley, Reading, MA, pp. 15–44.
- Bertsekas, D. P. & Tsitsiklis, J. N. (1996), *Neural-Dynamic Programming*, Athena Scientific, Belmont, MA.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.
- Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1996), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.
- Cliff, D. & Bruten, J. (1997), 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.

- Cohen, K. J., Maier, S. F., Schwartz, R. A. & Whitcomb, D. K. (1986), *The Microstructure of Securities Markets*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Cohen, K. J., Maier, S. F., Schwartz, R. A. & Whitcomb, D. K. (1991), 'A simulation model of stock exchange trading', *Simulation* **4**, 181–191.
- Cohen, K., Maier, S., Schwartz, R. & Whitcomb, D. (1983), 'A simulation model of stock exchange trading', *Simulation* **41**, 181–191.
- Dahan, E. & Srinivansan, V. (2000), 'The predictive power of internet-based product concept testing using depiction and animation', *Journal of Product Innovation Management* **17**(2), 99–109.
- Davis, D. & Holt, C. (1993), *Experimental Economics*, Princeton University Press, Princeton, NJ.
- Evgeniou, T., Pontil, M. & Poggio, T. (2000), 'Regularization networks and support vector machines', *Advances in Computational Mathematics* **13**(1), 1–50.
- Forsythe, R. & Lundholm, R. (1990), 'Information aggregation in an experimental market', *Econometrica* **58**, 309–47.
- Forsythe, R., Nelson, F., Neumann, G. & Wright, J. (1993), 'The iowa presidential stock market: A field experiment', *Research in Experimental Economics* pp. 1–43.
- Forsythe, R., Palfrey, T. R. & Plot, C. R. (1982), 'Asset valuation in an experimental market', *Econometrica* **50**, 537–567.
- Garman, M. (1976), 'Market microstructure', *Journal of Financial Economics* **3**, 257–275.
- Glosten, L. R. & Milgrom, P. R. (1985), 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of Financial Economics* **14**, 71–100.

- Gode, D. K. & Sunder, S. (1993), 'Allocative efficiency of markets with zero intelligence traders: Markets as a partial substitute for individual rationality', *Journal of Political Economy* **101**, 119–37.
- Green, P. E. & Wind, Y. (1981), New way to measure consumers' judgement, *in* Y. Wind, V. Mahajan & R. N. Cardozo, eds, 'New-product forecasting', D.C. Heath and Company, Lexington, MA, chapter 3, pp. 89–108.
- Grossman, S. J. & Stiglitz, J. E. (1980), 'On the impossibility of informationally efficient markets', *American Economic Review* **70**(3), 393–408.
- Hakansson, N. H., Beja, A. & Kale, J. (1985), 'On the feasibility of automated market making by a programmed specialist', *Journal of Finance* **40**(1), 1–20.
- Hasbrouck, J. & Sofianos, G. (1993), 'The trades of market makers: An empirical analysis of nyse specialists', *Journal of Finance* **48**(5), 1565–1593.
- Hasbrouck, J., Sofianos, G. & Sosebee, D. (1993), New york stock exchange: Systems and trading procedures. NYSE Working Paper.
- Ho, T. & Stoll, H. R. (1981), 'Optimal dealer pricing under transactions and return uncertainty', *Journal of Financial Economics* **9**, 37–73.
- Ho, T. & Stoll, H. R. (1983), 'The dynamics of dealer markets under competition', *Journal of Finance* **38**, 1053–1074.
- Jaakkola, T., Jordan, M. I. & Singh, S. P. (1994), 'On the convergence of stochastic iterative dynamic programming algorithms', *Neural Computation* **6**, 1185–1201.
- Jaakkola, T., Singh, S. P. & Jordan, M. I. (1995), Reinforcement learning algorithm for partially observable markov decision problems, *in* G. Tesauro, D. S. Touretzky & T. K. Leen, eds, 'Advances in Neural Information Processing Systems', MIT Press, Cambridge, MA, pp. 345–352.

- Jamal, K. & Sunder, S. (1996), 'Bayesian equilibrium in double auctions populated by biased heuristic traders', *Journal of Economic Behavior and Organization* **31**, 273–291.
- Kaelbling, L. & Moore, A. (1996), 'Reinforcement learning: A survey', *Journal of Artificial Intelligence Research* pp. 237–285.
- Kagel, J. & Roth, A., eds (1995), *Handbook of Experimental Economics*, Princeton University Press.
- Keynes, J. M. (1958), *The General Theory of Employment, Interest and Money*, Harcourt Brace.
- Kyle, A. S. (1985), 'Continuous auctions and insider trading', *Econometrica* **53**, 1315–1335.
- LeBaron, B. (forthcoming 1999), 'Agent based computational finance: Suggested readings and early research', *Journal of Economic Dynamics and Control* .
- Lettau, M. (1997), 'Explaining the facts with adaptive agents: The case of mutual fund flows', *Journal of Economic Dynamics and Control* **21**, 1117–1148.
- Lutostanski, J. (1982), Liquidity and Market-making, PhD thesis, Massachusetts Institute of Technology.
- Madhavan, A. (1992), 'Trading mechanisms in securities markets', *Journal of Finance* **47**(2), 607–641.
- Madhavan, A. (2000), 'Market microstructure: A survey', *Journal of Financial Markets* **2**, 205–258.
- Madhavan, A. & Sofianos, G. (1998), 'An empirical analysis of nyse specialist trading', *Journal of Financial Economics* **48**, 189–210.
- Minsky, M. L. (1963), Steps towards artificial intelligence, in E. A. Feigenbaum & J. Feldman, eds, 'Computers and Thought', McGraw-Hill, pp. 8–30.

- O'Brien, J. & Srivastava, S. (1991), 'Dynamic stock markets with multiple assets', *Journal of Finance* **46**, 1811–38.
- O'Hara, M. (1995), *Market Microstructure Theory*, Blackwell Publishers, Cambridge, MA.
- O'Hara, M. & Oldfield, G. (1986), 'The microeconomics of market making', *Journal of Financial and Quantitative Analysis* **21**, 361–376.
- Pennock, D. M. & Wellman, M. P. (1997), Representing aggregate belief through the competitive equilibrium of a security market, *in* 'Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence', pp. 392–400.
- Plott, C. R. & Sunder, S. (1982), 'Efficiency of experimental security markets with insider information: An application of rational-expectations models.', *Journal of Political Economy* **90**, 663–98.
- Plott, C. R. & Sunder, S. (1988), 'Rational expectations and the aggregation of diverse information in laboratory security markets', *Econometrica* **56**, 1085–1118.
- Routledge, B. R. (1994), Artificial selection: Genetic algorithms and learning in a rational expectations model, Technical report, GSIA, Carnegie Mellon, Pittsburgh, PA.
- Rust, J., Miller, J. & Palmer, R. (1992), Behavior of trading automata in a computerized double auction market, *in* D. Friedman & J. Rust, eds, 'The double auction market: Institutions, Theories, and Evidence', Addison-Wesley, Reading, MA, pp. 155–198.
- Schwartz, R. (1993), *Reshaping the Equity Markets*, Business One Irwin, Homewood, IL.
- Sofianos, G. (1995), Specialist gross trading revenues at the new york stock exchange. NYSE Working Paper 95-01.

- Sunder, S. (1992), 'Market for information: Experimental evidence', *Econometrica* **60**, 667–95.
- Sutton, R. & Barto, A. (1998), *Reinforcement Learning – an Introduction*, The MIT Press, Cambridge, MA.
- Sutton, R. S. (1996), Generalization in reinforcement learning: Successful examples using sparse coarse coding, *in* D. S. Touretzky, M. C. Mozer & M. E. Hasselmo, eds, 'Advances in Neural Information Processing Systems: Proceedings of the 1995 Conference', MIT Press, Cambridge, MA, pp. 1038–1044.
- Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer.
- Vapnik, V. (1998), *Statistical Learning Theory*, Wiley.
- Watkins, C. J. C. H. (1989), Learning from Delayed Rewards, PhD thesis, Cambridge University.
- Watkins, C. J. C. H. & Dayan, P. (1992), 'Q-learning', *Machine Learning* **8**, 279–292.
- Youssefmir, M. & Huberman, B. A. (1997), 'Clustered volatility in multiagent dynamics', *Journal of Economic Behavior and Organization* **32**, 101–118.