

Testing Rationality of Financial Markets

Thesis by
Oleg P. Bondarenko

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California

1998
(Submitted January 5, 1998)

Dedication

To my family

Acknowledgements

My stay at Caltech has been truly a learning experience and I would like to thank all the people who helped me during these four years.

I am especially indebted to my advisor, Peter Bossaerts, for his encouragement, intellectual inspiration, and excellent guidance throughout the completion of this thesis. He introduced me to the exciting world of finance and had the most valuable impact on the direction of my research and career path. Perhaps most of all, I am grateful for his understanding and willingness to help.

I would like to express my gratitude to my committee members Jeff Strnad, Richard McKelvey, and Jonathan Katz for their help and involvement in this project. My sincere thanks also go to Pierre Hillion, Matt Jackson, Rod Kiewiet, Charles Plott, and Simon Wilkie for their kind support.

I would like to thank all my fellow students who supported me, especially Olga Shvetsova, Misha Filippov, Bill Hartnett, Hye Yeong Lee, Jin Yu, Robin Hanson, and Yan Chen. I am also glad to acknowledge the friendship of Ilia, Max, Stas, Kurt, Pete, Dean, Szilvia and Kelly, and many other people whom I met during my stay at Caltech.

I am grateful for the helpful staff in Baxter Hall. Special thanks to Laurel Auchampaugh for her assistance in countless matters, to Wes Boudville for installing SAS and helping with other software.

My warm thanks go to my parents and my sister for their love and encouragement over the years.

Finally, my deepest gratitude is reserved for my love and best friend, my beautiful wife Lilia. I thank her for the happiest moments of the past years, especially for our newest joint project of bringing up our adorable son Nikita who has always been an amazing source of joy, inspiration, and distraction.

Abstract

This thesis consists of three related chapters that investigate rationality of asset pricing. The point of view of this study differs from traditional analysis of rational expectations (market efficiency) in one important aspect: market participants are not required to have correct beliefs about all parameters of the economy. Instead, investors' subjective assessment of the future may be different from the objective probability distribution. Other than that, investors are assumed to be rational which means that they process new information to the market in a consistent and efficient manner.

This line of research has been inspired by Bossaerts (1996) who proposed a set of novel rationality restrictions on securities prices that are robust to investors' initial beliefs. These restrictions obtain for prices of special finite-lived securities whose payoff can be categorized into two states: "out-of-the-money" state where the payoff is zero, and "in-the-money" state where the payoff is random. The simplest example of such securities is an Arrow-Debreu (AD) type security, a binary security paying \$1 in some states of the world and zero in others. More general examples include option-like contracts and, most importantly, equity.

In Chapter 1, we discuss the model of Bayesian agents with potentially biased beliefs and derive several related theoretical results. The new rationality restrictions allow one to test the central part of rational expectations, namely investors' usage of Bayes' law to update their beliefs, or *rationality of learning*. The strongest predictions obtain for prices of Arrow-Debreu (AD) securities and we use them in two empirical applications.

In Chapter 2, we investigate rationality and biases in expectations in the Iowa Electronic Market (IEM). The IEM, an experimental market, runs, among others, a number of *winner-take-all* markets with contracts based on the future prices of US stock exchange traded securities and indices. The winner-take-all contracts are

classical examples of AD securities with “1-0” payoff structure and, therefore, their prices can be used in the new tests. We find that price histories of these contracts are in accordance with rational learning but not with rational expectations.

As the next application, in Chapter 3, we suggest how rationality of learning can be tested in real financial markets, where no AD securities are traded. The idea here is to estimate prices of particular AD securities, namely *digital options*. Each of these digitals pays \$1 whenever the S&P 500 Index closing price on the expiration date is above a predetermined cutoff level, or zero otherwise. Prices of digitals can be estimated from prices of corresponding traded S&P calls and puts. The estimation procedure that we use to recover prices of digitals is in the spirit of nonparametric techniques and has several interesting features.

First, and most importantly, the estimation procedure does not require *stationarity* of an underlying asset’s price dynamics. This means that the option formula is allowed to change over time. Second, the estimation procedure enforces rudimentary theoretical *shape restrictions* on options prices. Third, it proposes an alternative parametric option specification which is *parsimonious*, computationally *very simple*, and which significantly outperforms the Black-Scholes specification for the studied dataset.

The analysis of time series of the estimated digital options supports the hypothesis of rational learning. Several tests based on the martingale restriction on prices of digitals indicate that investors of the S&P 500 Index Options efficiently process new information to the market. Still, strong evidence against rational expectations surfaces, implying that investors *must* have started with biased priors. We characterize the investors’ expectational errors and, in particular, show that, over the five-year period from 1991 to 1995, investors repeatedly underestimated the S&P 500 Index expected return and overestimated its volatility.

Contents

Acknowledgements	iv
Abstract	v
Introduction	1
1 Selection-Bias Induced Martingale Restrictions on Securities Prices	9
1.1 Introduction	9
1.2 The Benchmark Model	11
1.2.1 Arrow-Debreu Securities	13
1.2.2 Option-like Securities	19
1.2.3 Nonzero Risk-free Rate	25
1.3 Extension to Rational Learning	26
1.4 Illustration	28
1.4.1 A Digital Option	29
1.4.2 A Call Option	31
1.5 Conclusion	32
1.6 Appendix	33
2 Expectations and Learning in Iowa	39
2.1 Introduction	39
2.2 The Iowa Experimental Market	41
2.3 Testing Rationality of Learning	43
2.4 Testing For Rational Expectations	46
2.5 Conclusion	48
3 Testing Rationality of S&P 500 Index Options	56
3.1 Introduction	56

3.2	Constructing and Pricing Digital Options	58
3.3	The Dataset and Test Specification	62
3.3.1	S&P 500 Index Options	62
3.3.2	Estimation Procedure	65
3.3.3	Digital Options and Their Prices	72
3.4	Empirical Results	73
3.4.1	Testing Rationality of Learning	74
3.4.2	Testing For Rational Expectations	78
3.4.3	Analysis Of Investors' Expectational Errors	80
3.5	Conclusion	83
3.6	Appendix	86
3.6.1	Constrained Convex Least Squares Regression	86
3.6.2	Properties of Quadratic Specification	88
3.6.3	One Day in the Life of the Dataset	91
	Conclusion	107
	Bibliography	108

Introduction

Throughout the last several decades, almost all empirical work in macroeconomics and finance has relied on the hypothesis of *rational expectations*. This hypothesis was first formally stated by Muth (1961) and it asserts that, when forecasting the future, individuals do not make systematic mistakes and that their expectations about parameters of the economy are correct on average. The hypothesis of rational expectations is based on the following two important assumptions. The first assumption states that, when making consumption, savings and investment decisions, agents weigh potential future outcomes using probabilities that reflect the frequencies with which these *factually* will occur. In other words, the agents' *subjective* probability measure coincide with the *objective* (or, true) probability measure in the system. The second assumption says that, when new information arrives to the market, agents correctly update their belief. They infer the necessary information from signals (which include prices in the market) applying the rules of conditional probability (Bayes' law).

In finance, the related idea of investors' rationality is called the *efficient markets hypothesis* (EMH). The origin of the EMH can be traced back at least as far as the pioneering theoretical contribution of Bachelier (1900). In the modern literature, this hypothesis is usually stated as follows:

The prices of securities observed at any time are based on *correct* evaluation of all information available at that time. In an efficient market, prices *fully reflect* available information" – Fama (1970).

In a perfectly efficient market, the return of any security must be unpredictable from past information and must only reflect risk. If a security earns higher average returns than another, it must be because it is more risky.¹

¹It is common to distinguish among three versions of the EMH: the weak, semistrong, and strong forms of the hypothesis. Formally, the market is said to be efficient with respect to some information set if security prices would be unaffected by revealing that information to all participants. Moreover,

The first empirical tests in 1960s and early seventies seemed to support the hypothesis of market efficiency. However, starting in the mid-1970s, more sophisticated techniques found considerable evidence of predictability of securities returns.

For example, DeBondt and Thaler (1985), Jegadeesh and Titman (1993), Chopra, Lakonishok, and Ritter (1992) have shown that past returns of a stock are useful in predicting future returns. Also, Fama and French (1992, 1993, 1996), Lakonishok, Shleifer, and Vishny (1994) have demonstrated that future returns can be predicted by book-to-market ratio, size, earning-to-price ratio, and sales growth. Some other “anomalies” that seem to contradict the theory of efficient markets include the equity premium puzzle (see Mehra, Prescott (1985, 1988) and Rietz (1988)), mean reversion in stock prices (see, for example, Lo, MacKinlay (1988) and Poterba, Summers (1988)), some of the event studies anomalies, like a strong association between the magnitude of an earnings announcement and the postannouncement performance of equity (see, for example, Foster (1977), Foster, Olsen, Shevlin (1984), and Ball (1992)).

There is considerable disagreement about the causes of these anomalies and different explanations have been suggested.

Some believe that they may result, at least in part, from various biases in data or methodology. For example, Kothari, Shanken, and Sloan (1995) and Brown, Goetzmann, and Ross (1995) have suggested that survivorship bias could account for many apparent violations of market efficiency. Black (1993) and Lo and MacKinlay (1990) have argued that the findings could be due to some form of “data mining” or data “snooping.”

Others accept the fundamental nature of the anomalies of market efficiency but interpret them differently. Since it is not possible to test for market efficiency directly

efficiency with respect to an information set implies that it is impossible to make economic profits by trading on the basis of that information set.

Strictly speaking, the original definition of the EMH does not specify theoretical assumptions required for a market to be efficient, however, standard finance theory and virtually the entire empirical literature have interpreted market efficiency as implying the same two assumptions as the rational expectations hypothesis does: (i) correct prior beliefs, and (2) correct updating of beliefs from news.

(one can only conduct *joint* tests of market efficiency *and* some model of equilibrium prices), they disagree on whether these findings reflect market inefficiency or time-varying expected risk premia.

The first group believes that the cross-sectional variation in stock returns is *expected* and *required* by investors and that this variation reflects risk premia (see, for example, Fama and French (1992, 1993, 1996); Ball, Kothari, and Shanken (1995)). They argue that while the nature of risk premia seem inconsistent with the predictions of the Capital Asset Pricing Model (CAPM), the premia could be consistent with market efficiency under other, multifactor models. These authors search for more complicated asset pricing models to replace the CAPM but their belief in market efficiency is not shaken.

The second group of authors does not believe that risk can explain all anomalies and attributes the differentials in stock returns to investors' irrationality (for example, DeBondt and Thaler (1985), Chopra, Lakonishok, and Ritter (1992), and Lakonishok, Shleifer, and Vishny (1994)). These authors reject the hypothesis of market efficiency and develop alternative behavioral models in which investors have non-rational expectations about future cash flows or have faulty risk perceptions. In these models, the differentials in stock returns result from cognitive errors committed by investors who *over-* or *under-*react to various events.

The present study takes a different point of view on the hypothesis of rational expectations. It accepts that rational expectations is a very restrictive model which is often inconsistent with the empirical data. As an alternative, we consider a more realistic version of rationality which allows for biased (or, subjective) priors. In this model, investors remain rational, in the sense that they update their beliefs in a consistent, efficient manner (using the rules of conditional probabilities). They, however, are not required to always have correct (objective) beliefs about all parameters of the economy. This means that of the two assumptions of rational expectations we relax the first one (of correct beliefs) but keep the second and central assumption (of rational updating).

The assumption of correct beliefs has been often criticized as too strong and un-

realistic. For example, Swamy, Barth, and Tinsley (1982) have focused on the critical role that objective probability distributions play in the theory of rational expectations, where individuals calculate the expected values of key variables on the basis of these objective distributions. However, in many realistic circumstances, it is difficult to imagine how individuals actually find or learn about these probability distributions in the first place. For example, it is typical for the economic system to undergo continual structural changes: institutions constantly adapt, government agencies become more complex and assume different functions, and new economic actors appear on the scene. Also, in new or nonrepetitive situations (like initial public offerings in new industries, company liquidations, political events, etc.), how can someone seriously contend that individuals' expectations coincide with *true* probabilities governing the system?

As for the assumptions of rational updating, it seems less controversial. It “only” requires investors to be good analysts and to understand how securities prices are set in the market.

Correctness of beliefs is not crucial for many theoretical insights, such as the Euler equation restrictions obtained in Lucas (1978), the existence of an equivalent martingale measure shown in Harrison and Kreps (1979), or the dealer market pricing equations of Kyle (1985). These continue to hold even if we allow for biases in investors' beliefs. However, expectations must then be taken with respect to agents' subjective beliefs instead of the “objective probability measure.” Even homogeneity of beliefs is unnecessary.²

Instead, the attraction of the assumption of correct beliefs is exclusively *empirical*. It facilitates estimation and testing of asset pricing models, because investors' beliefs can readily be obtained from the empirical distribution. Parsimonious model testing

²Beliefs may differ across agents as long as priors remain “consistent” in the sense of Harsanyi (1967), which means that any differences in priors should be as if they were the outcome from differences in information only. When beliefs are inconsistent, agents will agree to disagree (Aumann (1976)) and start to speculate openly, something that is absent in standard asset pricing models, where only covert speculation takes place (exploitation of superior information behind noise or liquidity trading). See Bossaerts and Biais (1997) for a comparison of these beliefs concepts and further references.

ensues, as exemplified by Hansen and Singleton (1982). The introduction of biases in beliefs would require one to be much more specific about the economy, and, in particular, about learning, with the consequent danger that the model is rejected because the evolution of beliefs was misspecified.

Even proponents of rational expectations admit that, in many circumstances, individual probability beliefs may differ from objective beliefs. However, they defend the assumption of correct beliefs by arguing that *no* economic reasoning will be of any value *if* agents have incorrect beliefs. Lucas (1977) has expressed this common view that “anything is possible when beliefs are biased,” in the following:

... without some way of inferring what an agent’s subjective view of the future is, this hypothesis [of Bayesian updating with subjective beliefs] is of no help in understanding his behavior. Even psychotic behavior can be (and today, is) understood as “rational” given a sufficiently abnormal view of relevant probabilities.

To employ economic theory in practice, Lucas has argued, one *must* know what probability distributions agents actually use.

This point of view has recently been challenged by Bossaerts (1996) who has proposed a new theory of Bayesian investors with biased beliefs.³ The novelty of his approach is to analyze time-series of special finite-lived securities for which survival is related to the payoff. The terminal payoff of such a security can be categorized into two states: “out-of-the-money” state, where the security pays zero, and “in-the-money” state, where its payoff is random. The simplest example of such securities is an Arrow-Debreu (AD) type security, a binary security paying \$1 in some states of the world and zero in others. More general examples include option-like contracts.

By using a modification of the traditional return measure and introducing a selection bias, Bossaerts has derived a set of novel empirically relevant martingale restrictions that are robust to investors’ beliefs and does not depend on specification of

³Bossaerts refers to it as Theory of Efficiently Learning Markets

likelihood functions. Investors may have biased beliefs about the eventual payoff of securities but they are otherwise rational.

The new theory of Bayesian investors with biased beliefs, thus, offers a more general and realistic alternative to the traditional hypothesis of rational expectations. More importantly, however, this relaxation of rational expectations has simple, well-specified predictions (which can easily be tested) meaning that it is potentially rejectable. And this is in contrast to both behavioral models and descriptive models of time-varying risk premia which tend to be complicated, with vague and often untestable predictions.

This thesis follows the approach by Bossaerts (1996) to investigate rationality of asset pricing. It consists of three related chapters.

In Chapter 1, we discuss the theory of Bayesian agents with potentially biased beliefs and derive several additional theoretical results for securities whose payoff have the survivorship feature. In particular, for the case of general option-like securities, we propose new sub- and super-martingale restrictions. These restrictions obtain under relatively weak assumptions on an underlying security price dynamics and can be used in many interesting applications. We also consider extensions of the results obtained in Bossaerts (1996) for the case of “out-of-the-money” securities and nonzero risk-free rates.

The new selection-bias induced rationality restrictions have two important empirical implications. First, under rational expectations, they provide the necessary information to measure the effect of the survivorship bias and allow one to disentangle the survivorship bias from other economic phenomena.

Second, the new rationality restrictions suggest simple tests to discriminate between traditional rational expectations (with correct beliefs) and more general rational learning (with potentially biased beliefs). The strongest predictions obtain for prices of Arrow-Debreu (AD) securities and, to gauge empirical relevance of the new theory, we use them in two applications in Chapters 2 and 3.

In Chapter 2, we investigate rationality and biases in expectations in the Iowa Electronic Market (IEM). The IEM is an experimental Internet-based market devel-

oped at the College of Business Administration at the University of Iowa for research purposes. The IEM runs, among others, a number of *winner-take-all* submarkets with contracts based on the future prices of US stock exchange traded securities and indices. The winner-take-all contracts are classical examples of AD securities with “1-0” payoff structure and, therefore, their prices can be used in the new tests. We find that price histories of these contracts are in accordance with rational learning but not with rational expectations.

As the next application, in Chapter 3, we suggest how rationality of learning can be tested in real financial markets, where no AD securities are traded. The idea here is to estimate prices of particular AD securities, namely *digital options*. A *digital call* with exercise price X is a contract that pays \$1 when the maturity price of an underlying asset S_T exceeds X . Otherwise, it pays nothing. Thus, digital options (or, simply *digitals*) amount to straight bets on the terminal value of the underlying asset S_T .

Prices of digital options are implicit in prices of traded puts and calls. In Chapter 3, we use prices of S&P 500 Index Options to recover prices of corresponding digitals. The estimation procedure that we use is in the spirit of nonparametric techniques and has several interesting features.

First, and most importantly, the estimation procedure does not require *stationarity* of an underlying asset’s price dynamics. This means that the option formula is allowed to change over time. Second, the estimation procedure enforces rudimentary theoretical *shape restrictions* on options prices. Third, it uses an alternative parametric option specification which is *parsimonious*, computationally *very simple*, and which significantly outperforms the Black-Scholes specification for the studied dataset.

The analysis of time series of the estimated digital options supports the hypothesis of rational learning. Several tests based on the martingale restriction on prices of digitals indicate that investors of the S&P 500 Index Options efficiently process new information to the market. Still, strong evidence against rational expectations surfaces, implying that investors *must* have started with biased priors. The dataset

allows one to characterize the investors' expectational errors. In particular, it is shown that, over the five-year period from 1991 to 1995, investors repeatedly underestimated the S&P 500 Index expected return and overestimated its volatility.

Chapter 1 Selection-Bias Induced Martingale Restrictions on Securities Prices

1.1 Introduction

In this chapter we examine the theoretical implications of survivorship bias on asset pricing. Survivorship (or, selection) bias has been a longstanding problem of the empirical finance most recently addressed in Kothari, Shanken, and Sloan (1995) and Brown, Goetzmann, and Ross (1995). Empirical analysis of time series of asset returns is implicitly conditioned upon the asset surviving (or being selected) into the sample. The survivorship bias, in general, induces spurious relationship between the asset's realized returns and the asset's riskiness. Brown, Goetzmann and Ross (1995) have argued that this spurious relationship could, at least in part, explain a number of "anomalies" of market efficiency such as, the equity premium puzzle (Mehra, Prescott (1985, 1988) and Rietz (1988)), mean reversion in stock prices (Lo, MacKinlay (1988) and Poterba, Summers (1988)), and some of the event studies anomalies, like a strong association between the magnitude of an earnings announcement and the postannouncement performance of equity (Foster (1977), Foster, Olsen, Shevlin (1984), and Ball (1992)).

To formally investigate the implications of the survivorship (selection) bias, Bossaerts (1996) has considered special finite-lived securities that do not pay dividends and whose terminal payoff can be categorized into two states: "out-of-the-money" state, where the payoff is zero, and "in-the-money" state, where the payoff is random. The simplest example of such securities is an Arrow-Debreu (AD) type security, a binary security paying \$1 in some states of the world and zero in others. Such a security can be used to model a wide range of economic and finance environments: it could be a company's bankruptcy status on a certain date; whether or not a firm will have

its technology accepted as unique nationwide standard by the government; it could represent an outcome of a sport betting, and so on.

More general examples of securities for which survival is related to the payoff include option-like contracts. These contracts could describe equity, an emerging market, a high-yield bond, and several derivative securities (including puts, calls, warrants, caps, floors, etc.).

By using a new performance measure, Bossaerts has derived a set of novel rationality restrictions on prices of “survived” securities. These martingale restrictions provide the necessary information to measure the effect of the survivorship bias and allow one to disentangle the survivorship bias from meaningful economic phenomena.

Even more importantly, the new rationality restrictions are robust to investors’ expectations. Investors may have biased belief about eventual payoff of securities but they are otherwise rational, i.e., they process new information to the market in an efficient manner: they are Bayesian. This suggests that the new restrictions may be used to distinguish between traditional rational expectations where investors’ beliefs are correct, and more general rational learning where investors may have potentially biased beliefs.

In this chapter we follow the approach of Bossaerts (1996) and derive several related results for securities whose payoff have this survivorship feature. In particular, for the case of general option-like securities, we propose new sub- and super-martingale restrictions. These results obtain under relatively weak assumptions on an underlying security’s price dynamics and can be used in many interesting applications. We also consider extensions of the results obtained in Bossaerts (1996) for the case of “out-of-the-money” securities and nonzero risk-free rates.

The remainder of the chapter is organized as follows. The next section discusses the model and states new results for Arrow-Debreu and option-like securities under rational expectations. Section 1.3 extends the results to the setting of rational learning. Section 1.4 illustrates the theoretical restrictions by means of simulation examples. We conclude in Section 1.5.

1.2 The Benchmark Model

Let $(\Omega, \mathcal{F}, \pi)$ be a probability space, where points $\omega \in \Omega$ represent possible states of the world, $\mathcal{F}$ is a σ -algebra of subsets of Ω (which are called events), and $\pi : \mathcal{F} \rightarrow R^+$ is a probability measure on $\mathcal{F}$. The probability measure π can be interpreted for now as a set of unanimously held objective probability assessments concerning the state of the world.

We study a finite horizon economy where risk-neutral investors trade at discrete points of time indexed $t = 0, 1, \dots, T$. For the moment, we assume that the risk-free rate is zero. Consider a security ν which does not pay dividends and whose eventual payoff at time $t = T$ is V , a random variable defined on Ω .

The payoff V depends on the true state of the world ω picked by Nature at the outset. Investors do not know ω ; however, they gather information as time passes by and update their beliefs using Bayes' rule. At each period $t = 1, \dots, T$ they receive a signal $s_t \in S_t$ about the state of the world. A signal s_t could be public information common among all investors, or it could be a vector of private signals to investors. In the latter case, investors learn about the true state of the world through information reflected in the equilibrium prices $p_1, \dots, p_t$ of the traded security. The last signal s_T reveals the realized value of V .

The history of signals $s_1, \dots, s_t$ determines (possibly through the observable equilibrium prices $p_1, \dots, p_t$) the information flow to the market at time t , which is represented by a filtration $\{\mathcal{F}_t\}_{t=1}^T$, i.e., a nondecreasing family of sub- σ -fields of $\mathcal{F}$, such that $\mathcal{F}_t \subseteq \mathcal{F}_{t+1} \subseteq \mathcal{F}$, for all $t = 1, \dots, T$.¹

Under REE, the equilibrium price p_t equals the conditional expectation of the payoff V given the current information $\mathcal{F}_t$:

$$p_t = E[V | \mathcal{F}_t]. \quad (1.1)$$

The representation (1.1) implies the following fundamental result of the theory of

¹ $\mathcal{F}_t = \sigma(s_1, \dots, s_t)$ – a σ -field generated by the random variables s_k , i.e., the smallest σ -field with respect to which s_k is measurable for every $k = 1, \dots, t$.

rational expectations. Let r_{t+1} denote the return over the period $[t, t + 1]$:

$$r_{t+1} = \frac{p_{t+1} - p_t}{p_t}. \quad (1.2)$$

Proposition 1 [*Samuelson*] *Under REE, for any security*

$$E[r_{t+1}|\mathcal{F}_t] = 0.$$

Proof: The result obtains immediately from the law of iterated expectations:

$$E[p_{t+1}|\mathcal{F}_t] = E[E[V|\mathcal{F}_{t+1}]|\mathcal{F}_t] = E[V|\mathcal{F}_t] = p_t. \quad \triangle$$

Proposition 1, originally formalized in Samuelson (1965), implies that, under REE, risk neutrality and zero discounting, realized security returns should form a martingale difference sequence with respect to the filtration $\{\mathcal{F}_t\}_{t=1}^T$, the property usually interpreted as unpredictability of security returns, also known as the *market efficiency* property (Fama (1970)). Historical averages of security returns should be zero and projections of return r_{t+1} onto past information should produce insignificant coefficients.

To model a selection bias, we now consider a random variable θ defined on Ω which can take on two values $\bar{\theta}$ and $\underline{\theta}$. In interpretation, θ is a payoff relevant variable that represents general process of censoring (or selecting). For example, it could determine the security surviving into the sample. For an equity, the event $\{\theta = \bar{\theta}\}$ denotes that a firm survived on a certain date. In the case of options, $\theta = \bar{\theta}$ for those options that eventually matured in-the-money and so on.

In many situations, only securities with $\theta = \bar{\theta}$ generate complete price records available to an empiricist and one of the objectives of this chapter is to derive the restrictions on prices of these *censored* securities. More specifically, we know that

returns r_{t+1} of *uncensored* securities form a martingale difference sequence (Proposition 1). Unfortunately, this result is of little relevance for an econometrician who observes time-series with only $\theta = \bar{\theta}$. In this case, the *actual* distribution of the signals s_t and, hence, the equilibrium prices p_t is determined by $\mathcal{F}_t \wedge \bar{\theta} = \sigma(s_1, \dots, s_t, \bar{\theta})$, the σ -algebra generated by the information that investors have *as well as* the event $\{\theta = \bar{\theta}\}$.

In general, $E[r_{t+1} | \mathcal{F}_t \wedge \bar{\theta}] \neq 0$ and one wonders if there are martingale-type restrictions on securities prices which obtain under the empirically relevant filtration $\{\mathcal{F}_t \wedge \bar{\theta}\}_{t=1}^T$.

1.2.1 Arrow-Debreu Securities

We first consider the simplest possible case of Arrow-Debreu type securities, i.e., binary securities paying \$1 if $\theta = \bar{\theta}$ and zero otherwise, i.e.:

$$V_T = \begin{cases} 1, & \text{if } \theta = \bar{\theta} \\ 0, & \text{if } \theta = \underline{\theta} \end{cases} \quad (1.3)$$

In this case,

$$p_t = E[V | \mathcal{F}_t] = \text{Prob}(\theta = \bar{\theta} | \mathcal{F}_t).$$

In what follows we will often assume that the information filtration satisfies the following technical *no-revelation* assumption:

A1. For all $t < T$, $(\theta = \underline{\theta}) \notin \mathcal{F}_t$ *a.e.*

This assumption means that for *almost all* $\omega \in \Omega$, investors do not know at time $t < T$ whether the event $(\theta = \underline{\theta})$ has occurred.

An example that does not satisfy A1 is an Arrow-Debreu security that represents a firm's bankruptcy status: at a certain period $t = T$ in the future, the firm has either survived (i.e., $\theta = \bar{\theta}$) or is bankrupt ($\theta = \underline{\theta}$). Since the firm can become bankrupt prior to time $t = T$, there is a positive probability that the event $\{\theta = \underline{\theta}\}$ is revealed at some $t < T$, and, hence, A1 does not hold. Note that A1 does not ban situations

when $\bar{\theta} \in \mathcal{F}_t$ with some positive probability, only that $\underline{\theta} \notin \mathcal{F}_t$ *a.e.* We shall discuss this condition later on.

We use notations $\pi(A)$ for the probability density of event A and $\pi_t(A)$ for the conditional density of event A at time t (given that signals $s_1, \dots, s_t$ have been observed). We assume that the objective probability measure π is “factored” into an “in-the-money” probability measure π^1 which governs signals $s_1, \dots, s_t$ when $\theta = \bar{\theta}$, and an “out-of-the-money” probability measure π^0 when $\theta = \underline{\theta}$. If the initial probability of the event $\{\theta = \bar{\theta}\}$ (i.e., priors) is denoted as q , then π can be represented as

$$\pi(\cdot) \equiv \pi^q(\cdot) = q\pi^1(\cdot) + (1 - q)\pi^0(\cdot). \quad (1.4)$$

In this notation, time- t and time- $(t + 1)$ prices of ν can be written as follows:

$$p_t = \frac{\pi^1(s_1, \dots, s_t)}{\pi(s_1, \dots, s_t)}, \quad (1.5)$$

and

$$p_{t+1} = p_{t+1}(s_{t+1}) = \frac{\pi_t^1(s_{t+1})}{\pi_t(s_{t+1})} \frac{\pi^1(s_1, \dots, s_t)}{\pi(s_1, \dots, s_t)} = \frac{\pi_t^1(s_{t+1})}{\pi_t(s_{t+1})} p_t \quad (1.6)$$

As mentioned before, returns r_{t+1} of Arrow-Debreu securities that eventually matured in-the-money ($\theta = \bar{\theta}$) do not form a martingale difference sequence (in fact, they are positive on average reflecting an upward drift in prices of “winners” – see below). To obtain martingale restrictions for this case, Bossaerts (1996) has introduced the *modified* return defined as follows:

$$x_{t+1} = \frac{p_{t+1} - p_t}{p_{t+1}}. \quad (1.7)$$

The only difference from the traditional return r_{t+1} (1.2) is that the future price p_{t+1} instead of the past price p_t is used as a basis.

It turns out that the modified return x_{t+1} of Arrow-Debreu securities that eventually matured in-the-money plays a role similar to that of the traditional return x_{t+1} of uncensored securities. More specifically,

Proposition 2 [Bossaerts] Under REE and A1,

$$E[x_{t+1}|\mathcal{F}_t \wedge \bar{\theta}] = 0,$$

where the return x_{t+1} is computed as in (1.7) using prices of an Arrow-Debreu security with the payoff structure (1.3).

Proof: From (1.6)

$$E\left[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] = \int_{s_{t+1}} \frac{p_t}{p_{t+1}(s_{t+1})} \pi_t^1(s_{t+1}) ds_{t+1} = \int_{s_{t+1}} \pi_t(s_{t+1}) ds_{t+1} = 1,$$

and the result follows immediately from the definition of x_{t+1} . $\triangle$

Consequently, the histories of the modified return of Arrow-Debreu securities that matured in-the-money satisfy a martingale difference sequence and must be unpredictable from investors' past information. Propositions 1 and 2 together solve the problem of the survivorship bias for the case of Arrow-Debreu securities. Indeed, under REE, the difference between the traditional and modified returns of Arrow-Debreu securities that matured in-the-money “measures” the effect of the selection bias which now can be disentangled from other economic phenomena.

Proposition 2 can be easily extended to the case of Arrow-Debreu securities that matured *out-of-the-money*, i.e., when $\theta = \underline{\theta}$. Define a new Arrow-Debreu security $\hat{\nu}$ as complementary to ν . This security pays \$1 if $\theta = \underline{\theta}$ and zero if $\theta = \bar{\theta}$. Let $\hat{p}_t$ denote the equilibrium price of $\hat{\nu}$. The corresponding traditional and modified returns $\hat{r}_{t+1}$ and $\hat{x}_{t+1}$ are defined similar to (1.2) and (1.7). From the no-arbitrage condition $p_t + \hat{p}_t = 1$, returns $\hat{r}_{t+1}$ and $\hat{x}_{t+1}$ can be rewritten in terms of prices of ν as follows:

$$\hat{r}_{t+1} \equiv \frac{\hat{p}_{t+1} - \hat{p}_t}{\hat{p}_t} = -\frac{p_{t+1} - p_t}{1 - p_t}, \quad (1.8)$$

$$\hat{x}_{t+1} \equiv \frac{\hat{p}_{t+1} - \hat{p}_t}{\hat{p}_{t+1}} = -\frac{p_{t+1} - p_t}{1 - p_{t+1}}. \quad (1.9)$$

When Arrow-Debreu security ν matures out-of-the-money, the complementary

Arrow-Debreu security $\hat{\nu}$ matures in-the-money and we can apply Proposition 2 to the modified return $\hat{x}_{t+1}$ of ν . The assumption which is “dual” to A1 is:

A2. For all $t < T$, $(\theta = \bar{\theta}) \notin \mathcal{F}_t$ *a.e.*

Note that assumption A2 is not the same as A1. The condition A2 means that, for *almost all* $\omega \in \Omega$, the event $\{\theta = \bar{\theta}\}$ cannot possibly be revealed at any date prior to $t = T$. It assumes nothing about possibility for event $(\{\theta = \underline{\theta}\})$ to be revealed before $t = T$. An Arrow-Debreu security that represents a firm bankruptcy status is an example that does not satisfy A1 but satisfies A2 – the fact that a firm has survived $(\theta = \bar{\theta})$ is only known *after* $t = T$.²

The following corollary of Proposition 2 holds for Arrow-Debreu securities that matured out-of-the-money.

Corollary 1 *Under REE and A2,*

$$E[\hat{x}_{t+1} | \mathcal{F}_t \wedge \underline{\theta}] = E\left[-\frac{p_{t+1} - p_t}{1 - p_{t+1}} | \mathcal{F}_t \wedge \underline{\theta}\right] = 0,$$

where the return $\hat{x}_{t+1}$ is computed as in (1.9) using prices of an Arrow-Debreu security with the payoff structure (1.3).

While being a trivial extension to Proposition 2, Corollary 1 is important because it allows one to analyze time series of Arrow-Debreu securities that fail to satisfy assumption A1 (but still satisfy A2) – potentially a wide class of interesting applications which includes such examples as survival of a company, emerging markets, peso problems, default of a low-quality bond, and others.³

²Another example would be betting on sport events. Consider an Arrow-Debreu security which pays \$1 if a certain team wins the NBA championship in 1996 (i.e., $\theta = \bar{\theta}$). This security does not satisfy A1 since $(\theta = \underline{\theta})$ can be revealed at some intermediate time (for example, when the team is eliminated in a first round playoff). Nevertheless, it satisfies A2 – bettors cannot possibly know that the event $\{\theta = \bar{\theta}\}$ has occurred before the playoff ends.

³Bossaerts’ Theorem 4 states another result for Arrow-Debreu securities that matured out-of-the-money. That theorem is derived for return series in *reverse time* and it requires signals s_t to be conditionally independent in reverse time. Unfortunately, this assumption is very restrictive and is not satisfied in most realistic environments. The problem here is that the reversed sequence of prices $p_T, \dots, p_1$ “converges” to the starting price $p_0 \equiv q$ (i.e., investors’ priors about θ). This means that

The proof of Proposition 2 reveals the role of assumption A1. The derivation depends on the cancelation of term $\pi_t^1(s_{t+1})$ which has been assumed positive *almost everywhere* in view of A1. The following example is intended to illustrate what happens when an Arrow-Debreu security does not satisfy A1.

Example 1. There are 16 teams in the NBA playoff which consists of four rounds denoted $t = 1, 2, 3, 4$. Assume that all teams have the equal chance to win in any round. Consider an Arrow-Debreu security ν that at time $t = 5$ pays one dollar if a certain team wins the championship (i.e., $\theta = \bar{\theta}$). Before round t , price of the Arrow-Debreu security p_t is 2^{t-5} or zero if the team has been already eliminated.

There are five possible price series for ν :

$\{ \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1 \}$ – with probability $\frac{1}{16}$;

$\{ \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 0 \}$ – with probability $\frac{1}{16}$;

$\{ \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 0, 0 \}$ – with probability $\frac{1}{8}$;

$\{ \frac{1}{16}, \frac{1}{8}, 0, 0, 0 \}$ – with probability $\frac{1}{4}$;

$\{ \frac{1}{16}, 0, 0, 0, 0 \}$ – with probability $\frac{1}{2}$.

We first observe that the Arrow-Debreu security ν does not satisfy assumption A1 (for example, with the probability $\frac{1}{2}$ the event $\{\theta = \underline{\theta}\}$ is announced after round 1). If an empiricist now computes the time-(t+1) average modified return $\bar{x}_{t+1}$ for the securities that eventually matured in-the-money, she would find that

$$\bar{x}_{t+1} = \frac{1}{2} \neq 0, \quad t = 1, 2, 3, 4.^4$$

signals s_t and prices p_t , in general, will not be independent in *reverse time* (even despite the fact that they might be conditionally independent in *forward time*). For an illustration of this point, see Section 1.4.

In contrast, Corollary 1 applies to general informational flows.

⁴Bossaerts (1996) has used a simple limiting argument to provide theoretical justification of the rationality restriction in Proposition 2 when A1 is not satisfied. However, for an *empirical* analysis of prices of surviving securities, this limiting argument will be often of little help.

However, ν satisfies assumption A2 and, as Corollary 1 predicts,

$$E[\hat{x}_2|\mathcal{F}_1 \wedge \underline{\theta}] = \frac{7}{15} \left(-\frac{\frac{1}{8} - \frac{1}{16}}{1 - \frac{1}{8}} \right) + \frac{8}{15} \left(-\frac{0 - \frac{1}{16}}{1 - 0} \right) = 0.$$

Similarly,

$$E[\hat{x}_3|\mathcal{F}_2 \wedge \underline{\theta}] = \frac{3}{7} \left(-\frac{\frac{1}{4} - \frac{1}{8}}{1 - \frac{1}{4}} \right) + \frac{4}{7} \left(-\frac{0 - \frac{1}{8}}{1 - 0} \right) = 0,$$

$$E[\hat{x}_4|\mathcal{F}_3 \wedge \underline{\theta}] = \frac{1}{3} \left(-\frac{\frac{1}{2} - \frac{1}{4}}{1 - \frac{1}{2}} \right) + \frac{2}{3} \left(-\frac{0 - \frac{1}{4}}{1 - 0} \right) = 0.$$

From the proof of Proposition 2 it is clear that when an Arrow-Debreu security does not satisfy A1 (A2),⁵ returns x_{t+1} ($\hat{x}_{t+1}$) are nonnegative on average and we state that as

Corollary 2 *Under REE,*

$$E[x_{t+1}|\mathcal{F}_t \wedge \bar{\theta}] = \text{Prob}(p_{t+1} = 0|\mathcal{F}_t) \geq 0,$$

$$E[\hat{x}_{t+1}|\mathcal{F}_t \wedge \underline{\theta}] = \text{Prob}(p_{t+1} = 1|\mathcal{F}_t) \geq 0,$$

where the returns x_{t+1} and $\hat{x}_{t+1}$ are computed as in (1.7) and (1.9) using prices of an Arrow-Debreu security with the payoff structure (1.3).

By using Proposition 1, Corollaries 1-2 and applying inequality

$$r_{t+1} - x_{t+1} \equiv \frac{(p_{t+1} - p_t)^2}{p_{t+1}p_t} > 0, \tag{1.10}$$

we immediately obtain the following restrictions given here for completeness.

Corollary 3 *Under REE,*

$$E[x_{t+1}|\mathcal{F}_t] \leq 0,$$

⁵Note that at time T neither A1 nor A2 are satisfied ($\theta \in \mathcal{F}_T$ by definition). This means that Proposition 2 and Corollary 1 do not apply to the last period returns x_T and $\hat{x}_T$.

$$E[r_{t+1}|\mathcal{F}_t \wedge \bar{\theta}] \geq 0,$$

$$E[x_{t+1}|\mathcal{F}_t \wedge \underline{\theta}] \leq E[r_{t+1}|\mathcal{F}_t \wedge \underline{\theta}] \leq 0,$$

where the returns r_{t+1} and x_{t+1} are computed as in (1.2) and (1.7) using prices of an Arrow-Debreu security with the payoff structure (1.3).

Remarks. Similar restrictions obtain for returns $\hat{r}_{t+1}$ and $\hat{x}_{t+1}$. Note that Corollary 3 does not depend on assumption A1 (or A2). All inequalities in Corollary 3 are strict unless $p_{t+1}(s_{t+1})$ is equal to p_t almost everywhere.⁶

1.2.2 Option-like Securities

We now turn to the case of more general securities which are subject to censoring by θ . Specifically, let us consider a security ν which pays some random amount $V(\omega) > 0$ when $\theta = \bar{\theta}$ and zero when $\theta = \underline{\theta}$:

$$V_T = \begin{cases} V(\omega), & \text{if } \theta = \bar{\theta} \\ 0, & \text{if } \theta = \underline{\theta} \end{cases} \quad (1.11)$$

We refer to ν as an *option-like* security because its payoff structure is similar to that of a call option

$$C_T(E) = \begin{cases} X_T - E, & \text{if } X_T > E \\ 0, & \text{if } X_T \leq E \end{cases}$$

where E is exercise price and X_T is price of an underlying security at the maturity date $t = T$.

An equity, index of an emerging capital market and some of contingent claims are typical examples that can be modeled by an option-like security.

The equilibrium price of ν is p_t and returns r_{t+1} and x_{t+1} are defined as before

⁶I.e., when s_{t+1} is independent of θ (for *almost all* $s_{t+1} \in S_{t+1}$). In this case, the signal s_{t+1} reveals no additional information about whether or not the event $\{\theta = \bar{\theta}\}$ has occurred and $p_{t+1} = p_t$ a.e. (see (1.6)).

in (1.2) and (1.7). We observe that the price $p_t = E[V_T|\mathcal{F}_t]$ can be written as:

$$p_t = q_t V_t,$$

where

$$q_t \equiv \text{Prob}(\theta = \bar{\theta}|\mathcal{F}_t \wedge \bar{\theta}),$$

$$V_t \equiv E[V_T|\mathcal{F}_t \wedge \bar{\theta}].$$

In this representation, q_t denotes the probability of the event $\{\theta = \bar{\theta}\}$ and, thus, the results for Arrow-Debreu securities from the previous section apply. In particular (provided that A1 is satisfied):

$$E\left[\frac{1}{q_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] = \frac{1}{q_t}. \quad (1.12)$$

On the other hand, V_t denotes the expectation of the terminal payoff V_T conditional on $\theta = \bar{\theta}$. The law of iterated expectations with respect to the enhanced filtration $\{\mathcal{F}_t \wedge \bar{\theta}\}_{t=1}^T$ implies that

$$E[V_{t+1}|\mathcal{F}_t \wedge \bar{\theta}] = V_t. \quad (1.13)$$

We also know that Proposition 1 holds for the return r_{t+1} of uncensored time series of option-like securities :

$$E[p_{t+1}|\mathcal{F}_t] = p_t. \quad (1.14)$$

As for restrictions derived for the information filtration $\{\mathcal{F}_t \wedge \bar{\theta}\}_{t=1}^T$, the analogue of Proposition 2 does not obtain. Modified returns x_{t+1} of an option-like security that eventually matured in-the-money do *not* generally form a martingale difference sequence. In this case, however, the following modification of Proposition 2 can be proved.

Proposition 3 [Bossaerts] Under REE and A1,

$$E[x_{t+1} V_T | \mathcal{F}_t \wedge \bar{\theta}] = 0,$$

where the return x_{t+1} is computed as in (1.7) using prices of an option-like security with the payoff structure (1.11).

Proof: In view of (1.12) and (1.13)

$$E[p_t \frac{V_T}{p_{t+1}} | \mathcal{F}_t \wedge \bar{\theta}] = E[p_t \frac{V_{t+1}}{p_{t+1}} | \mathcal{F}_t \wedge \bar{\theta}] = V_t$$

and

$$E[V_T | \mathcal{F}_t \wedge \bar{\theta}] = V_t.$$

This implies that $E[x_{t+1} V_T | \mathcal{F}_t \wedge \bar{\theta}] = 0$. $\triangle$

In Proposition 3, V_T can be replaced by any V_τ ($t+1 \leq \tau \leq T$). In practice, however, conditional expectations of the payoff V_t 's are generally not known except for the terminal value V_T .⁷ Notice that when V_T is independent of signals $s_1, \dots, s_T$, Proposition 3 coincides with Proposition 2 for the case of Arrow-Debreu securities. In most interesting applications, however, V_T is correlated with signals $s_1, \dots, s_T$. In this case, Proposition 3 is weaker than Proposition 2. Since V_T remains fixed in a single time series, the empirical scope of Proposition 5 is limited to cross-sectional analysis when V_T varies across different securities.

To derive restrictions on prices of option-like securities that can be used for a time-series analysis, we state two *monotonicity* assumptions on an underlying security price dynamics.

A3. For all time- t signals s'_t, s''_t ($t < T$),

⁷Also, by applying the out-of-the-money restrictions to the Arrow-Debreu price $q_t = p_t/V_t$ we obtain that

$$E[\frac{p_{t+1}V_t - p_tV_{t+1}}{V_{t+1} - p_{t+1}} | \mathcal{F}_t \wedge \underline{\theta}] = 0,$$

however, this restriction is not interesting because it involves unobservable values V_t and V_{t+1} which cannot be conditioned out.

$$\begin{aligned} \text{if} \quad & p_t(s'_t) > p_t(s''_t) \\ \text{then} \quad & q_t(s'_t) \geq q_t(s''_t) \quad a.e.. \end{aligned}$$

A4. For all time- t signals s'_t, s''_t ($t < T$),

$$\begin{aligned} \text{if} \quad & p_t(s'_t) > p_t(s''_t) \\ \text{then} \quad & V_t(s'_t) \geq V_t(s''_t) \quad a.e.. \end{aligned}$$

Assumption A3 simply says that, at any time t , a higher price of ν implies a higher probability for ν to mature in-the-money. Similarly, assumption A4 says that a higher price of ν implies a higher expected payoff V_T (conditional on $\theta = \bar{\theta}$). Consequently, conditions A3 and A4 are intuitive and natural in many applications. We shall discuss them after we derive new results for option-like securities.

Proposition 4 *Under REE and A3,*

$$E[r_{t+1} | \mathcal{F}_t \wedge \bar{\theta}] \geq 0,$$

where the return r_{t+1} is computed as in (1.2) using prices of an option-like security with the payoff structure (1.11).

Proof: Proof in the Appendix.

Proposition 5 *Under REE, A1, and A4,*

$$E[x_{t+1} | \mathcal{F}_t \wedge \bar{\theta}] \leq 0,$$

where the return x_{t+1} is computed as in (1.7) using prices of an option-like security with the payoff structure (1.11).

Proof: Proof in the Appendix.⁸

⁸See also Theorem 6 in Bossaerts (1996) where it has been shown that $E[x_{t+1} | \mathcal{F}_t \wedge \bar{\theta}] \leq 0$ under a different restriction on the underlying security's price dynamics.

Propositions 4 and 5 suggest a simple test of the rational expectations hypothesis when only price records for surviving securities ($\theta = \bar{\theta}$) are available to an empiricist: one needs to verify that the traditional return r_{t+1} of such a security forms a *submartingale* difference sequence (i.e., it is positive on average) while the modified return x_{t+1} forms a *supermartingale* difference sequence (it is negative on average). In other words, the interval $[x_{t+1}, r_{t+1}]$ must include zero on average.

This test is potentially powerful because the difference between the traditional and modified returns is usually small (it is of second order of magnitude since $r_{t+1} - x_{t+1} = r_{t+1} x_{t+1}$). This means that Propositions 4 and 5 specify rather tight bounds on securities prices under REE. Violation of these restrictions would reject rational expectations.

One can also interpret the difference between the traditional and modified returns of the survived securities as an upper bound on the survivorship bias because under REE

$$E[x_{t+1}|\mathcal{F}_t \wedge \bar{\theta}] \leq E[r_{t+1}|\mathcal{F}_t] = 0 \leq E[r_{t+1}|\mathcal{F}_t \wedge \bar{\theta}].$$

This, in particular, means that if some anomaly of market efficiency do not “disappear” when the modified return x_{t+1} of the survived securities is used (instead of the traditional return r_{t+1}), then the survivorship bias *alone* cannot explain the anomaly.

To use Propositions 4 and 5, however, we need to assume that the underlying security price process satisfies both A3 and A4 which means the following. If signal s'_t is “better” than signal s''_t (i.e., $p_t(s'_t) > p_t(s''_t)$), then *both* the probability of maturing in-the-money q_t *and* the expected payoff V_t conditional on expiration in-the-money for signal s'_t are no less than those for signal s''_t :

$$q_t(s'_t) \geq q_t(s''_t),$$

$$V_t(s'_t) \geq V_t(s''_t).$$

Usually, one can directly verify assumptions A3 and A4 when the underlying security price dynamics is known. For example, one can check that assumptions A3

and A4 are satisfied for call and put options in the Black-Scholes model.⁹ One can also show that when processes q_t and V_t are independent, the restrictions of Propositions 4 and 5 obtain (even though assumptions A3 and A4 may not be satisfied).

Assumption A3 (A4) is sufficient but not necessary for Proposition 4 (Proposition 5) to obtain. However, when A1 (A2) is not satisfied, the prediction of Proposition 4 (Proposition 5) may not hold. This is illustrated in the following two examples.

Example 2. Consider a simple finite-state world where the first signal s_1 can take on only two values, s' and s'' . We assume that the signal probabilities are as in the following matrix:

	$s_1 = s'$	$s_1 = s''$
$\theta = \bar{\theta}$	0.1	0.1
$\theta = \underline{\theta}$	0.1	0.7

We also assume that $V_1(s') = 1$ and $V_1(s'') = 9$. From these data we can easily compute that

$$\begin{aligned} q_1(s') &= \frac{1}{2}, & V_1(s') &= 1, & p_1(s') &= \frac{1}{2}, \\ q_1(s'') &= \frac{1}{8}, & V_1(s'') &= 9, & p_1(s'') &= \frac{9}{8}, \\ q_0 &= \frac{1}{5}, & V_0 &= 5, & p_0 &= 1. \end{aligned}$$

Clearly, this example does not satisfy the monotonicity condition A3 ($q_1(s') > q_1(s'')$ while $p_1(s') < p_1(s'')$), and we find that

$$E\left[\frac{p_1 - p_0}{p_0} | \mathcal{F}_0 \wedge \bar{\theta}\right] = \frac{1}{2} \left(\frac{1}{2} + \frac{9}{8} \right) - 1 = -\frac{3}{16} < 0.$$

Example 3. Consider a slight modification of the previous example. The signal probabilities are the same:

⁹See Section 1.4 for a numerical illustration.

$$s_1 = s' \quad s_1 = s''$$

$\theta = \bar{\theta}$	0.1	0.1
$\theta = \underline{\theta}$	0.1	0.7

However, now $V_1(s') = 2$ and $V_1(s'') = 4$. For the price process of this example,

$$\begin{aligned} q_1(s') &= \frac{1}{2}, & V_1(s') &= 2, & p_1(s') &= 1, \\ q_1(s'') &= \frac{1}{8}, & V_1(s'') &= 4, & p_1(s'') &= \frac{1}{2}, \\ q_0 &= \frac{1}{5}, & V_0 &= 3, & p_0 &= \frac{3}{5}. \end{aligned}$$

In this example, the monotonicity condition A4 is not satisfied ($V_1(s') < V_1(s'')$ while $p_1(s') > p_1(s'')$), and we find that

$$E\left[\frac{p_1 - p_0}{p_1} | \mathcal{F}_0 \wedge \bar{\theta}\right] = 1 - \frac{1}{2} \left(\frac{3}{5} + \frac{6}{5} \right) = \frac{1}{10} > 0.$$

1.2.3 Nonzero Risk-free Rate

To complete the set of restrictions on censored time series, we now consider relaxing the assumption of a zero risk-free rate. We continue to study a general option-like security ν with the payoff structure as in (1.11).

Let $\mu_t \geq 0$ denote the risk-free rate over the period $[t, t+1]$. We can define

$$B_{t+1} = \prod_{\tau=0}^t (1 + \mu_\tau), \tag{1.15}$$

the time- $(t+1)$ value of one dollar invested at time 0 at the risk-free money account. We observe that all previous results continue to hold when the risk-free asset B_t is used as *numeraire*, i.e., securities prices are expressed in terms of B_t rather than dollars. For example, instead of Proposition 1 we obtain the following martingale property:

$$E\left[\frac{p_{t+1}}{B_{t+1}} | \mathcal{F}_t\right] = \frac{p_t}{B_t}.$$

Also, under REE and A1 (Proposition 2):¹⁰

$$E\left[\frac{p_{t+1}B_t - p_tB_{t+1}}{p_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] = 0. \quad (1.16)$$

In other words, if we define a new option-like security $\tilde{\nu}$ with the payoff $\tilde{V}_T = V_T/B_T$ when $\theta = \bar{\theta}$ and zero when $\theta = \underline{\theta}$, then all results of Section 1.2.2 apply to prices $\tilde{p}_t$ of $\tilde{\nu}$.

However, in the special case when risk-free rates μ_t can be viewed as *independent* of signals $s_1, \dots, s_T$ (which also includes constant and deterministic rates), these extensions to the case of nonzero risk-free rate become particularly simple. In this case,

$$\begin{aligned} E[p_{t+1}|\mathcal{F}_t] &= p_t(1 + \mu_t), \\ E\left[\frac{1}{p_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] &= \frac{1}{p_t(1 + \mu_t)}, \end{aligned}$$

and we obtain the following

Corollary 4 *When risk-free rates μ_t are independent of signals $s_1, \dots, s_T$, Propositions 1-5 and Corollaries 1-3 continue to hold if returns of ν are defined in excess of the risk-free rate, i.e.,*

$$\begin{aligned} r_{t+1} &= \frac{p_{t+1} - p_t}{p_t} - \mu_t, \\ x_{t+1} &= \frac{p_{t+1} - p_t}{p_{t+1}} - \frac{\mu_t}{1 + \mu_t}, \end{aligned}$$

and similarly for $\hat{r}_{t+1}$, $\hat{x}_{t+1}$.

1.3 Extension to Rational Learning

In the previous section, the results for censored securities (Propositions 2-5) have been all derived under the assumption that securities prices are set in REE. These results,

¹⁰See also Bossaerts' Theorem 5 for the pricing kernel $M_t = 1/B_t$.

however, continue to be valid in more general setting where investors' expectations may be biased. In this setting investors are Bayesian and rationally process information but they update their beliefs using *subjective* probability measure π^* instead of *objective* probability measure $\pi \equiv \pi^q = q\pi^1 + (1 - q)\pi^0$ that they use in REE. We assume that π^* can also be factored into π^1 and π^0 , i.e., that there exists $q^* \geq 0$ (representing investors' biased priors about the event $\{\theta = \bar{\theta}\}$) such that

$$\pi^* = \pi^{q^*} = q^*\pi^1 + (1 - q^*)\pi^0. \quad (1.17)$$

The interpretation behind the representation in (1.17) is that investors priors about θ may deviate from *true* priors q , yet, they make *correct* inferences about θ when new information (signals s_t) arrives to the market. In this case, we say that investors have *correct conditional beliefs* and refer to the corresponding equilibrium concept as Rational Learning Equilibrium (RLE).

Consider again an option-like security ν with the payoff structure (1.11). Its equilibrium price is p_t under rational expectations (i.e., when investors use the objective probability measure π) and p_t^* under more general rational learning (i.e., when investors use the subjective probability measure π^*):

$$p_t^* = E^*[V_T | \mathcal{F}_t],$$

where $E^*[\cdot]$ denotes integration with respect to π^* .

When investors' conditional beliefs are correct,

$$E[Z | \mathcal{F}_t \wedge \bar{\theta}] = E^*[Z | \mathcal{F}_t \wedge \bar{\theta}],$$

for *any* measurable with respect to $\mathcal{F}$ random variable Z (in both cases integration is carried out with respect to π^1). For example, under A1

$$E\left[\frac{p_{t+1}^* - p_t^*}{p_{t+1}^*} | \mathcal{F}_t \wedge \bar{\theta}\right] = E^*\left[\frac{p_{t+1}^* - p_t^*}{p_{t+1}^*} | \mathcal{F}_t \wedge \bar{\theta}\right] = 0.$$

Now it becomes clear why *all* previous results for censored securities are robust to investors' priors. Those results (Propositions 2-5 and Corollaries 1-4) are stated in terms of expectations with respect to filtration $\{\mathcal{F}_t \wedge \bar{\theta}\}_{t=1}^T$ (as opposed to Proposition 1 which is stated with respect to filtration $\{\mathcal{F}\}_{t=1}^T$). Since expectations $E[\cdot|\mathcal{F}_t \wedge \bar{\theta}]$ can everywhere be replaced with $E^*[\cdot|\mathcal{F}_t \wedge \bar{\theta}]$ we arrive to the same conclusions under the subjective probability measure π^* . For example, under A1

$$E\left[\frac{p_{t+1}^* - p_t^*}{p_{t+1}^*} \middle| \mathcal{F}_t \wedge \bar{\theta}\right] = E^*\left[\frac{p_{t+1}^* - p_t^*}{p_{t+1}^*} \middle| \mathcal{F}_t \wedge \bar{\theta}\right] = 0.$$

Hence, the following result holds.

Proposition 6 *[Bossaerts] Propositions 2-5 and Corollaries 1-4 continue to hold under RLE.*

Proposition 6 suggests that the restrictions on prices of censored securities can be used to distinguish between traditional rational expectations (with correct beliefs) and more general rational learning (with potentially biased beliefs). In particular, when Proposition 1 is rejected, rational expectations must be rejected as well. That would not necessarily reject rationality of learning, however. It could be the case that REE fails to hold *only* because investors have started with biased beliefs, but they are otherwise rational (i.e., they correctly apply the rules of conditional probability). That is the core of the more general class of RLE for which only Propositions 2-5 hold.

1.4 Illustration

For our simulation exercises we consider a financial market in which the price of an underlying asset is modeled by a geometric Brownian motion (GBM) with constant drift α and volatility σ :

$$dX_t = \alpha X_t dt + \sigma X_t dW_t, \tag{1.18}$$

where dW_t denotes a standard Wiener process. The risk-free interest rate is μ and the initial price of the asset is X_0 . For simplicity, we assume that the asset does not pay dividends.

Taking the number of days per year to be $252 = 12\text{mos} \times 21\text{days}$ we draw signals s_t from the normal distribution $N(\alpha/365, \sigma^2/252)$ to approximate the sample paths of the geometrical Brownian motion.¹¹ The price of the underlying asset at time t , $t = 1, \dots, T$ is given by

$$X_t = X_0 \exp \left(\sum_{i=1}^t s_i \right).$$

Let $C_t(X_t, E)$ denote the time- t value of European call with maturity date T and exercise price E when the current price of the underlying asset is X_t . Time to maturity $\tau = T - t$. From the Black-Scholes formula

$$C_t(X_t, E) = X_t \Phi(d_1) - E e^{-\mu\tau} \Phi(d_2), \quad (1.19)$$

where

$$d_1 = \frac{\ln(X_t/E) + (\mu + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}},$$

$$d_2 = \frac{\ln(X_t/E) + (\mu - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}},$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

1.4.1 A Digital Option

For the first example we consider a particular Arrow-Debreu security ν which pays \$1 when $X_T > E$ and zero dollars otherwise. This security is often referred to as a *digital (or binary) call*. The REE price of ν , denoted p_t , is equal to the probability that the call option $C_t(X_t, E)$ eventually matures in-the-money ($\theta = \bar{\theta}$):

$$p_t = \text{Prob}(\theta = \bar{\theta} \mid \mathcal{F}_t) = \text{Prob}(X_T \geq E \mid \mathcal{F}_t).$$

¹¹Following the common convention in the finance literature, when we calculate option prices, annual values are converted to a daily frequency by dividing by 365 for interest rates and returns, and by $\sqrt{252}$ for standard deviations.

Using the familiar result in Breeden and Litzenberg (1978), the price p_t can be computed from the call function as follows:

$$p_t = -e^{r(T-t)} \frac{\partial C(X_t, E, T-t)}{\partial E}. \quad (1.20)$$

From this expression¹² we obtain that

$$p_t = \Phi(d_2). \quad (1.21)$$

We take annual volatility $\sigma = 30\%$, $\mu = \alpha = 0$, and a time-to-maturity of two months ($T = 42$ days). The initial price of the underlying security X_0 is normalized to 1.

For the "at-the-money" strike price $E = 1.0$ we generate a sample path of the underlying asset $X_0, X_1, \dots, X_T$ and calculate a price path $p_0, p_1, \dots, p_T$ of the corresponding digital option from (1.21).

We generate $k = 1, \dots, 10,000$ sample paths of the underlying asset, and compute the time-series average returns r^k , x^k , $\hat{r}^k$, and $\hat{x}^k$ for each price path k of the corresponding digital option:¹³

$$r^k = \frac{1}{37} \sum_{t=1}^{37} r_{t+1}^k,$$

and similarly for returns x^k , $\hat{r}^k$, and $\hat{x}^k$.

After that we average returns r^k , x^k , $\hat{r}^k$ and $\hat{x}^k$ across sample paths that matured in-the-money, ($\theta = \bar{\theta}$), across sample paths that matured out-of-the-money ($\theta = \underline{\theta}$), and across all simulations. These (cross-sectional) average returns r , x , $\hat{r}$, and $\hat{x}$ are reported in Table 1.1.

The reported results supports the predictions of Propositions 1-2. In particular, the average modified return x for digital options that matured in-the-money insignif-

¹²See Chapter 3 for another derivation of the result in (1.20).

¹³We exclude the last five prices in each series because near-maturity prices usually converge very quickly to 1 (for digital options that eventually matured in-the-money) or to 0 (for digital options that eventually matured out-of-the-money).

icantly differs from zero (Proposition 2) while the average traditional return r is positive and significant (as Corollary 3 predicts).

A similar situation appears for the digital options that matured out-of-the-money. The average return $\hat{x}$ is insignificantly different from zero (Corollary 1), while $\hat{r}$ is positive and significant.

Table 1.1 also illustrates the observation about the martingale restriction in reverse time (Theorem 4 in Bossaerts (1996)). According to this theorem, the return r_{t+1} must be zero on average if signals s_t are (conditionally) independent in reverse time. In this example, signals s_t are independent in forward time but *not* in reverse time since a reversed sequence of underlying security prices $X_T, \dots, X_1$ always converges to $X_0 = 1.0$. The significantly negative average return r for digital options that matured out-of-the-money reflects this fact.

Table 1.2 and Table 1.3 report two independent sets of the same simulations for in-the-money digital options (strike price of $E = 0.9$) and for out-of-the-money digital options (strike price of $E = 1.1$) respectively.

1.4.2 A Call Option

The second example illustrates Propositions 3, 4 and 5 for option-like securities. We keep all the parameters the same as in the previous example and consider time series of call option prices. In this case, the equilibrium price p_t is given by:

$$p_t = C_t(X_t, E).$$

For each of three strike prices, in-the-money $E = 0.9$, at-the-money $E = 1.0$, and out-of-the-money $E = 1.1$, we generate 10,000 price series and compute average returns r and x across sample paths that matured in-the-money. Also, to illustrate Proposition 3, for each sample path k of that matured in-the-money we compute the

modified “weighted” return over the period $[0, T/2]$, i.e.:

$$v^k = C_T^k(X_T^k, E) \frac{p_{T/2}^k - p_0^k}{p_{T/2}^k} = (X_T^k - E) \frac{p_{T/2}^k - p_0^k}{p_{T/2}^k}.$$

Let v denote the average of v^k across all matured in-the-money calls. The results are reported in Table 1.4 and they confirm the predictions of Propositions 3-5 for option-like securities.

For all three strike prices, average returns r are significantly positive (Proposition 4), average returns x are significantly negative (Proposition 5) and average returns v are insignificantly different from zero (Proposition 3).

1.5 Conclusion

In this chapter we examine the implications of the survivorship (selection) bias on asset pricing. Following the model developed in Bossaerts (1996) we derive several related empirically relevant restrictions on prices of securities whose payoff have a survivorship feature.

These selection-bias induced rationality restrictions have two important empirical implications. First, under rational expectations, they provide the necessary information to measure the extent of the survivorship bias problem.

Second, they can be used to distinguish between the traditional rational expectations (when beliefs are required to be correct) and the more general rational learning (when beliefs may be biased). The strongest restrictions obtain for prices of Arrow-Debreu securities and we use them in two applications in Chapters 2 and 3.

1.6 Appendix

Proof of Proposition 4

The idea of the proof is to show that under A3 the following inequality must be true:

$$E\left[\frac{p_{t+1}}{p_t}|\mathcal{F}_t \wedge \bar{\theta}\right] E[1|\mathcal{F}_t] \geq E\left[\frac{V_{t+1}}{V_t}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] \quad (1.22)$$

Indeed,

$$\begin{aligned} I_1 &\equiv E\left[\frac{p_{t+1}}{p_t}|\mathcal{F}_t \wedge \bar{\theta}\right] E[1|\mathcal{F}_t] = \\ &\int \frac{p_{t+1}(s'_{t+1})}{p_t} \pi_t^1(s'_{t+1}) ds'_{t+1} \int 1 \pi_t(s''_{t+1}) ds''_{t+1} = \\ &\frac{1}{2} \int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{p_t q_t} \left(p_{t+1}(s'_{t+1}) q_{t+1}(s'_{t+1}) + p_{t+1}(s''_{t+1}) q_{t+1}(s''_{t+1}) \right) ds'_{t+1} ds''_{t+1}. \end{aligned}$$

Similarly,

$$\begin{aligned} I_2 &\equiv E\left[\frac{V_{t+1}}{V_t}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] = \\ &\int \frac{V_{t+1}(s'_{t+1})}{V_t} \pi_t^1(s'_{t+1}) ds'_{t+1} \int \frac{q_{t+1}(s'_{t+1})}{q_t} \pi_t(s''_{t+1}) ds''_{t+1} = \\ &\frac{1}{2} \int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{p_t q_t} \left(p_{t+1}(s'_{t+1}) q_{t+1}(s''_{t+1}) + p_{t+1}(s''_{t+1}) q_{t+1}(s'_{t+1}) \right) ds'_{t+1} ds''_{t+1}. \end{aligned}$$

Finally,

$$\begin{aligned} 2(I_1 - I_2) &= \\ &\int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{p_t q_t} \left(p_{t+1}(s'_{t+1}) - p_{t+1}(s''_{t+1}) \right) \left(q_{t+1}(s'_{t+1}) - q_{t+1}(s''_{t+1}) \right) ds'_{t+1} ds''_{t+1}. \end{aligned}$$

The last representation implies that $I_1 - I_2 \geq 0$ (under assumptions A3). Now, since we also know that

$$E[1|\mathcal{F}_t] = E\left[\frac{V_{t+1}}{V_t}|\mathcal{F}_t \wedge \bar{\theta}\right] = E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] = 1,$$

the result follows immediately from (1.22). $\triangle$

Proof of Proposition 5

The proof is similar to that of Proposition 4 except now we use a different inequality. We want to show that, under A4:

$$E\left[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{p_{t+1}}{p_t}|\mathcal{F}_t\right] \geq E\left[\frac{q_t}{q_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] \quad (1.23)$$

In this case, a similar trick yields that

$$\begin{aligned} I_1 &\equiv E\left[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{p_{t+1}}{p_t}|\mathcal{F}_t\right] = \\ &\int \frac{p_t}{p_{t+1}(s'_{t+1})} \pi_t^1(s'_{t+1}) ds'_{t+1} \int \frac{p_{t+1}(s''_{t+1})}{p_t} \pi_t(s''_{t+1}) ds''_{t+1} = \\ &\frac{1}{2} \int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{q_t} \left(\frac{p_{t+1}(s'_{t+1})}{V_{t+1}(s''_{t+1})} + \frac{p_{t+1}(s''_{t+1})}{V_{t+1}(s'_{t+1})} \right) ds'_{t+1} ds''_{t+1}, \end{aligned}$$

and

$$\begin{aligned} I_2 &\equiv E\left[\frac{q_t}{q_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] = \\ &\int \frac{q_t}{q_{t+1}(s'_{t+1})} \pi_t^1(s'_{t+1}) ds'_{t+1} \int \frac{q_{t+1}(s''_{t+1})}{q_t} \pi_t(s''_{t+1}) ds''_{t+1} = \\ &\frac{1}{2} \int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{q_t} \left(\frac{p_{t+1}(s'_{t+1})}{V_{t+1}(s'_{t+1})} + \frac{p_{t+1}(s''_{t+1})}{V_{t+1}(s''_{t+1})} \right) ds'_{t+1} ds''_{t+1}. \end{aligned}$$

Using these representations for I_1 and I_2 we obtain that

$$\begin{aligned} 2(I_1 - I_2) &= \\ &\int \int \frac{\pi_t(s'_{t+1}) \pi_t(s''_{t+1})}{q_t} (p_{t+1}(s'_{t+1}) - p_{t+1}(s''_{t+1})) \left(\frac{1}{V_{t+1}(s''_{t+1})} - \frac{1}{V_{t+1}(s'_{t+1})} \right) ds'_{t+1} ds''_{t+1}. \end{aligned}$$

This proves that, under assumption A4, $I_1 - I_2 \geq 0$. Finally, we recall that

$$E\left[\frac{p_{t+1}}{p_t}|\mathcal{F}_t\right] = E\left[\frac{q_{t+1}}{q_t}|\mathcal{F}_t\right] = 1,$$

and that, under A1,

$$E\left[\frac{q_t}{q_{t+1}}|\mathcal{F}_t \wedge \bar{\theta}\right] = 1,$$

and the result follows from (1.23). $\triangle$

Table 1.1: Average returns r , x , $\hat{r}$, and $\hat{x}$ of digital options computed across sample paths that matured in-the-money, ($\theta = \bar{\theta}$), across sample paths that matured out-of-the-money ($\theta = \underline{\theta}$) and across all simulations. The strike price $E = 1.0$. Results are based on 10,000 independent time series (t -statistics in parentheses).

	N	r	x	$\hat{r}$	$\hat{x}$
$\theta = \bar{\theta}$	4920	0.0265 (121.82)	-0.0002 (-0.50)	-0.0250 (-15.76)	-8.7684 (-1.66)
$\theta = \underline{\theta}$	5080	-0.0235 (-18.87)	-7.6318 (-1.84)	0.0225 (118.80)	-0.0003 (-0.73)
All	10000	0.0011 (1.39)	-3.8771 (-1.85)	-0.0009 (-1.07)	-4.3142 (-1.66)

Table 1.2: Average returns r , x , $\hat{r}$, and $\hat{x}$ of digital options computed across sample paths that matured in-the-money, ($\theta = \bar{\theta}$), across sample paths that matured out-of-the-money ($\theta = \underline{\theta}$) and across all simulations. The strike price $E = 0.9$. Results are based on 10,000 independent time series (t -statistics in parentheses).

	N	r	x	$\hat{r}$	$\hat{x}$
$\theta = \bar{\theta}$	8006	0.0073 (92.03)	0.0000 (0.39)	-0.0135 (-8.83)	-5.7651 (-6.15)
$\theta = \underline{\theta}$	1994	-0.0274 (-23.23)	-0.3533 (-7.77)	0.0626 (95.19)	-0.0015 (-1.42)
All	10000	0.0004 (1.29)	-0.0704 (-7.67)	0.0017 (1.34)	-4.6159 (-6.15)

Table 1.3: Average returns r , x , $\hat{r}$, and $\hat{x}$ of digital options computed across sample paths that matured in-the-money, ($\theta = \bar{\theta}$), across sample paths that matured out-of-the-money ($\theta = \underline{\theta}$) and across all simulations. The strike price $E = 1.1$. Results are based on 10,000 independent time series (t -statistics in parentheses).

	N	r	x	$\hat{r}$	$\hat{x}$
$\theta = \bar{\theta}$	2085	0.0570 (107.88)	0.0012 (1.22)	-0.0218 (-30.88)	-0.0613 (-30.58)
$\theta = \underline{\theta}$	7915	-0.0115 (-13.13)	-0.2600 (-57.08)	0.0052 (60.56)	-0.0002 (-1.34)
All	10000	0.0028 (1.22)	-0.2055 (-18.24)	-0.0004 (-0.73)	-0.0130 (-8.60)

Table 1.4: Average returns r , x , and v for call options that matured in-the-money. Results are based on 10,000 independent time series for strike prices $E = 0.9$, $E = 1.0$, and $E = 1.1$ (t -statistics in parentheses).

	N	r	x	v
$E = 0.9$	7780	0.0105 (50.9)	-0.0250 (-56.5)	-0.0007 (-0.38)
$E = 1.0$	4977	0.0402 (106.2)	-0.0289 (-37.2)	-0.0011 (-0.82)
$E = 1.1$	2452	0.0854 (105.7)	-0.0346 (-23.9)	-0.0004 (-0.65)

Chapter 2 Expectations and Learning in Iowa

2.1 Introduction

The purpose of this chapter is to illustrate how the theoretical results discussed in Chapter 1 can be implemented to test for the Rational Learning Equilibrium (RLE) against the more restrictive traditional Rational Expectations Equilibrium (REE).

We investigate the pricing of “winner take all” contracts traded in the Iowa Electronic Markets (IEM), computerized experimental markets developed at the University of Iowa for academic purposes. Winner-take-all contracts are classical examples of Arrow-Debreu securities: they pay \$1 in one state of the world; \$0 in all others. We study the contracts which derive their payoff from price changes of stocks in the computer industry and the S&P 500 Index. These winner-take-all contracts allow us to use the strongest restrictions from Chapter 1 derived for prices of Arrow-Debreu securities.

Specifically, let r_{t+1} and x_{t+1} denote the traditional and modified returns of an Arrow-Debreu security with time- t price p_t :

$$r_{t+1} = \frac{p_{t+1} - p_t}{p_t}, \quad (2.1)$$

$$x_{t+1} = \frac{p_{t+1} - p_t}{p_{t+1}}. \quad (2.2)$$

Under rational expectations, the traditional return r_{t+1} is unpredictable from investors' past information:

$$E[r_{t+1} | \mathcal{F}_t] = 0. \quad (2.3)$$

Under more general rationality of learning, the traditional return r_{t+1} and modified return x_{t+1} of Arrow-Debreu securities that matured in-the-money must satisfy the

following restrictions:

$$E[r_{t+1}|\mathcal{F}_t, \bar{\theta}] > 0, \quad (2.4)$$

$$E[x_{t+1}|\mathcal{F}_t, \bar{\theta}] = 0. \quad (2.5)$$

Under REE, restriction (2.3) can be verified in a cross section of random histories of prices of Arrow-Debreu securities. If this property is rejected, rational expectations must be rejected as well. That would not reject rationality, however. Agents could have started with biased beliefs. For rationality, they need only apply correctly the rules of conditional probability. That is the core of the more general class of RLE. In those, restrictions (2.4)-(2.5) for returns of securities that matured in-the-money hold. Thus, a cross section of price histories of “winning” Arrow-Debreu securities allows us to test the proposition that investors’ learning is rational.

The analysis of time series of winner-take-all contracts suggests that prices in the Iowa markets are in accordance with rational learning. Evidence against rational expectations surfaces, however, indicating that investors may have started with biased priors. Among other things, there is a pronounced negative correlation between the price level and the subsequent price change. This finding is not unlike the evidence from stock markets (e.g., Keim and Stambaugh (1986)).

It deserves emphasis that the latter does not necessarily imply that there is money to be made. The priors of participants in the Iowa markets are as good as any other. An outsider may have different priors, but whether this would lead her to “make money” depends on the circumstances. She would if her priors *happened to be better* (this is a chance event). Her trading would amount to pure speculation, because our results indicate that participants in the Iowa market certainly do not leave “money on the table” because they do not know how to correctly update their beliefs (they could be over-reacting or under-reacting to new information).

The remainder of this chapter is organized as follows. The next section describes the Iowa experimental markets that are the focus of the study. Section 2.3 presents the results of the tests of rationality of learning. Section 2.4 investigates rational

expectations. Section 2.5 concludes.

2.2 The Iowa Experimental Market

The College of Business Administration at the University of Iowa runs a number of markets in winner-take-all contracts based on the returns or prices of U.S. stock exchange listed securities and indices. Each month, a new set of contracts is offered. Contract liquidation values are determined by rates of return or changes in closing prices of the underlying stock measured from the third Friday of one month to the third Friday of the next month. Trade starts on Monday following the third Friday of the month. Trading halts when the liquidation value is determined. Rates of return on the underlying stock securities are adjusted for dividend payments. Rates of return on indices are not dividend-adjusted.

Iowa manages another set of markets, where the payoffs on the traded securities depend on political events, such as the outcome of presidential or parliamentary elections. These markets pre-date the ones that we are studying here. For an in-depth analysis of the political markets, see Forsythe, Nelson, Neumann and Wright (1992).

We investigate two sets of markets: the computer industry returns markets and the Microsoft price level markets. Four winner-take-all contracts are traded in the former. They pay one dollar depending on which underlying security had the highest return: Apple, IBM, Microsoft or the S&P 500, respectively. Contracts written on losers expire worthless. Hence, these contracts are Arrow-Debreu securities written on a state space determined by the returns of the aforementioned stocks and the S&P 500 Index. In the Microsoft markets, two contracts are traded. One, the “High” contract, pays a dollar when Microsoft’s next month closing price is above a predetermined cut-off level. The other one, the “Low,” pays one dollar in the complementary state. The cut-off level is determined by the exercise price of the closest-at-the-money, exchange-traded option (CBOE) on Microsoft.

The trading mechanism of the Iowa Electronic Market is the computerized double

auction. Participation extends beyond the University of Iowa (the market is Internet-based), but is restricted to investors with academic affiliations. In addition to trading individual contracts, a “market” portfolio can be purchased and sold. This is a unit portfolio which bundles all contracts in a set of markets, and, hence, guarantees a unit payout. It can be purchased from or sold to the system for a unit price, or bought/sold in the market (the latter requires, however, that there be quotes in all contracts).

We limit our attention to daily closing prices, which are defined to be the last transaction price before midnight, or, if no transactions took place, the previous closing price. Notice that closing prices on complementary contracts are not necessarily synchronous. This fact along with the bid-ask spread imply, among other things, that the closing prices do not necessarily add up to one, even if a bundle of complementary contracts guarantees a unit payoff, and that returns are not collinear. Let us look at the data to gauge the importance of nonsynchrony.¹

Table 2.1 displays some descriptive price and volume statistics. Figure 2.1 plots the price paths for the two Microsoft securities over four sample months. Table 2.1 lists the average and standard deviation of the absolute deviation of the sum of the prices of complementary securities from one. The numbers are sizeable, indicating that nonsynchrony is a serious problem. It also means that the returns on complementary securities are not necessarily collinear, a fact that we will exploit in our tests. The lack of collinearity is also apparent from Figure 2.1: if synchronous, the price paths of the two Microsoft securities should be perfectly symmetric with respect to a line through \$0.50. They are not.

Table 2.1 also documents substantial positive skewness and extreme kurtosis in the distribution of returns. Readers familiar with derivatives markets (such as the winner-take-all markets we are studying here) will not be surprised. The kurtosis is also apparent in Figure 2.1: the history for the 10/95 Microsoft High contract (which

¹Note that one can earn risk-free profit *if* at any time sum of the bid prices of complimentary contracts is greater than one or sum of the ask price is lower than one. However, when transactions are nonsynchronous and could take place either at the bid or at the ask, the fact that the closing prices on complimentary contracts do not always add up to one does not imply that arbitrage opportunities are available in the Iowa markets.

pays \$1 if Microsoft's price increased from September 1995) includes a day (10/17/95) when the price jumped from \$0.024 to \$0.455, a return of 1,796%! The skewness is reflected in the surprisingly high average daily return. See Table 2.1. The averages are less than two standard errors from zero, however, except in the pooled sample.

The next two sections present the tests for rationality of learning and rational expectations based on the martingale results discussed in Chapter 1. These results can be tested in many different ways, however, due to data problems, in analyzing returns of the Iowa winner-take-all contracts, we concentrate on the simplest predictions only. We provide more complete analysis of the theoretical results in the context of the higher quality options data in Chapter 3.

2.3 Testing Rationality of Learning

Restriction (2.5) demonstrates that one can use the modified return (where the end-of-period price is used to deflate payoffs) of winning securities to test for the rationality of learning. This result solely depends on agents' correctly applying the rules of conditional probability (Bayes' law), and not on the initial beliefs, which can be arbitrarily biased.

Let us look at the returns and modified returns *on winning securities* in the Iowa market data described in the previous section. Because of restriction (2.4), we expect the traditional return to be positive on average. This contrasts with the modified return, which should be zero on average.

We computed average daily returns and modified returns for each month-long price history and then tested whether the average (of this average) was zero across the monthly histories. Since the dataset covers 16 months, we will have a cross section of sixteen averages for each individual set of markets (the Microsoft markets and the computer industry markets separately) and 32 averages when we pool both sets of markets. The outcomes in the cross section can be safely assumed to be independent.

We can also look at the securities with payoffs complementary to those of the losing securities. Of course, the complementary securities are winners as well. Com-

plementary securities can be created by purchasing the market portfolio from the system for \$1 and selling one traded security. We take the price of the complementary security to be the value of this portfolio. In the Microsoft case, each security is the complement of the other, so that results are expected to be similar to those obtained from actually traded winning securities. Discrepancies will arise, however, because securities trade nonsynchronously. In the computer industry markets, the complementary security to, say, Apple, is the one that pays \$1 if any other security than Apple wins.

In the computer industry markets, there will be correlation across monthly histories of complementary securities that traded simultaneously, i.e., during the same month. The results that we are about to report do not correct for this. When we attempted to address the problem (mainly by running Seemingly Unrelated Regressions), the results were not altered qualitatively.

Table 2.2 displays the average of the time series daily return average, as well as its standard error. In all cases is this average more than two standard errors above zero, corroborating restriction (2.4). Table 2.2 also lists results for the average of the time series daily *modified* return average. It is always within two standard errors from zero, in accordance with restriction (2.5).²

In Table 2.1, we documented extreme skewness and kurtosis in the distribution of returns. One wonders to what extent these affect the inference in Table 2.2. Therefore, Table 2.2 also displays the skewness and excess kurtosis of the month-long averages. The values are far more acceptable, rendering credibility to our results. This is certainly so for the modified returns.

We can conclude provisionally that learning in the Iowa markets is rational: there are no biases in the modified returns of winning securities.

In fact, restriction (2.5) is stronger than this simple test of the average modified return may indicate. For learning to be rational, the average modified return of

²We only report standard errors and whether the statistic is 1.8 or 2 standard errors from zero. In order to avoid unnecessary controversy, we refrain from attaching a specific p -level to the estimates. We allow the reader to use his or her favorite distributional theory to determine what the corresponding p -level is.

winning securities ought to be unpredictable from *any* past information. We can test this in two ways.

First, we can correlate the modified return with past information, by computing the average of the modified return times, say, the lagged price, or the lagged price squared. Second, we can simply regress the modified return onto lagged information, e.g., the lagged price.

Let us discuss results from the first strategy. We compute monthly averages of $x_{t+1}p_t$ and $x_{t+1}(p_t)^2$, and test whether they are zero across the months. Of course, we continue to limit our attention to winners. For comparison, we also report results for monthly averages based on the traditional return, i.e., from $r_{t+1}p_t$ and $r_{t+1}(p_t)^2$.³

Table 2.3 reports the results. In no case is the monthly average of $x_{t+1}p_t$ or $x_{t+1}(p_t)^2$ significantly different from zero, supporting rationality of learning. This contrasts with the averages of $r_{t+1}p_t$ and $r_{t+1}(p_t)^2$, which are all significantly positive. Taking the results of Tables 2 and 3 together, the modified return of winning contracts is not predictable from past information embedded in the price level.

The second strategy, whereby we project the (modified) return onto the beginning-of-period price using OLS, deserves more discussion. Many researchers, starting with Keim and Stambaugh (1986), have found that stock returns are negatively related to the beginning-of-period price. For winners, however, the projection of the *modified* return onto the lagged price should produce insignificant results, at least if prices are set in a RLE. We verify this for the Iowa markets.

When plainly projecting the modified return on the lagged price, however, the error exhibits substantial heteroscedasticity: its variance is far higher for low price levels. This obviously affects the inference. It turns out that there is a simple way to correct for the heteroscedasticity in the error term, namely, by regressing the modified

³Later in Chapter 3, we analyze the traditional and modified returns r_{t+1} and x_{t+1} after multiplying them by p_t^γ for some real γ . We also argue that the transformed returns, $r_{t+1}p_t^\gamma$ and $x_{t+1}p_t^\gamma$, may be interpreted as returns of particular dynamic trading strategies. Specifically, the case $\gamma > 1$ corresponds to the “follow market” strategy: sell securities when their price goes down and buy them when price goes up. The case $\gamma < 1$ corresponds to the “contrarian” strategy: sell securities when their price goes up and buy them when price goes down. Finally, when $\gamma = 1$ (the case we investigate for the Iowa contracts) the trading strategy is “buy and hold” where composition of the portfolio remains constant over time.

return times the lagged price ($x_{t+1}p_t$) onto the lagged price (p_t).

Table 2.4 reports results of this alternative projection. For comparison purposes, we also display results for OLS projections of the traditional return multiplied by the lagged price (i.e., the price change) onto the lagged price. The contrast between the highly significant coefficients for the traditional return against the small and insignificant ones for the modified return is striking. As before, rational learning is supported.⁴

2.4 Testing For Rational Expectations

As is well-known, REE can be tested by verifying (in-sample) predictability of returns using past information. We must not, however, condition on future information, such as, e.g., the eventual payoff on a security. In the previous section, we did so, because the tests were inspired by restriction (2.5). In contrast, tests of rational expectations are based on restriction (2.3), where one conditions out (i.e., averages across) eventual outcomes.

We initially tested REE by projecting returns onto lagged prices. As with the projections leading to Table 2.4, however, we discovered pronounced heteroscedasticity in the error term: its variance decreased with the price level. Again, a simple transformation purged the error of heteroscedasticity, namely, multiplication of the return by the lagged price. Of course, this means that the regressand became simply the change in the price over the subsequent day.⁵

Figure 2.2 provides a scatter plot of daily price changes against lagged prices for the Microsoft contracts. The negative trend stands out. Also, there is little evidence of heteroscedasticity as a function of the price level.

Table 2.5 reports the results of the OLS projections of price changes onto lagged prices. The negative trend apparent in Figure 2.2 translates into significantly positive intercepts and negative slope coefficients. It is interesting to note that this is precisely

⁴The results for the securities that are complementary to the losers are similar to those for the winning contracts. They are not reported to conserve spaces.

⁵See also footnote 3.

what has been observed for U.S. stock returns as well (see, e.g., Keim and Stambaugh (1986)).

The evidence in Table 2.5 leads us to refute REE at high confidence levels. Because we concluded in the previous section that closing prices in the Iowa markets reflect rational learning, however, we cannot attribute this rejection to over-reaction (the explanation usually advanced for this phenomenon in the stock markets). It must be agents' beliefs. They are biased, in violation of one of the assumptions behind rational expectations.

In what sense are beliefs biased? A closer investigation of the data revealed the following. Take the Microsoft markets, for instance. The Microsoft High contract matured in the money 12 times out of 16 (months). This seems to indicate that chances for Microsoft High to pay off \$1 are 75%, which would lead us to assign a value of \$0.75 at the beginning of trading. We noticed, however, that the first closing price for Microsoft High was often substantially different from \$0.75. The deviation for the 9/95 contract, for instance, was more than \$0.30 (see Figure 2.1).

By themselves, these deviations from what we consider to be a fair value for the Microsoft High contract may merely be evidence of the superior information that markets possess over us, or they may reflect that the Microsoft High contract does not always start trading at the money. However, together with the negative relationship between price levels and subsequent returns, the deviations in the initial Microsoft High prices from fair value seem to indicate that Iowans started out with more extreme priors than can be justified *ex post*. Beliefs subsequently reverted back to some more acceptable level.

Unfortunately, we observe only 16 predictions. The forecasts reflected in the initial closing prices did not deviate sufficiently to be able to make formal statements about the extent to which agents' beliefs were biased.

In this respect, it is worth emphasizing that a comparison between the negative correlation between price levels and subsequent price changes (documented in Table 2.5), on the one hand, and the absence of correlation between the modified return and lagged prices for winning securities (documented in Table 2.4), on the other hand,

did allow us to conclude that initial beliefs must have been biased. It underscores the power of our methodology.

As a final note, we should caution the reader against overinterpreting our evidence against rational expectations. Return measures do not always accurately reflect potential profitability. Our returns are computed on the basis of *transaction prices*. Transactions could take place either at the bid or at the ask. If today's closing transactions took place at the bid, and tomorrow's at the ask, the return one computes from the closing prices is an incorrect measure of the payoff on a strategy of purchasing the security at today's close and selling at tomorrow's close.⁶

Among other things, the regression results reported in Table 2.5 could be affected by this bid-ask bounce, if closes with a low price predominantly take place at the bid. The subsequent return would be biased upward. This bias would translate into a negative correlation between the error in the regression of returns onto previous closing prices and the regressor. The OLS slope coefficient would be negatively biased. Absent quote data, however, we have no way to tell how severe this problem is.

Still, it is important to appreciate that if bid-ask bounce biased the regressions with the traditional return as regressor, it ought to have affected the ones with the modified return as regressand as well. Tables 3 and 4 documented, however, that the results for the latter conform nicely to the theory. This makes our evidence against rational expectations less ambivalent.

2.5 Conclusion

We cannot find any predictability in return sequences of winning securities in the Iowa experimental markets where the payoffs are defined with respect to U.S. exchange-listed stock. The returns were suitably modified to adjust for the selection bias caused by our focusing on securities that matured in the money. Since these modified returns are unpredictable if agents correctly apply the rules of conditional probability,

⁶Several papers in the finance literature point out the potential biases from using transaction prices, starting with Blume and Stambaugh (1983).

we conclude that we cannot reject that market prices in Iowa reflect rational learning.

In contrast, we find evidence of in-sample predictability of returns, in the form of a negative correlation between the daily price change and the level of the previous closing price. Since in-sample predictability violates rational expectations but agents appear to have correctly implemented Bayes' law, we conclude that their priors must have been biased. Under rational expectations, the priors had to coincide with the actual distribution of outcomes.

What does this mean in terms of the traditional notion of market efficiency (Fama (1970, 1991)), i.e., for a newcomer's ability to "make money"? Based on the evidence reported here, a newcomer should refrain from betting against the Iowa markets. There is no money to be made from mistakes against the rules of conditional probability. Still, the newcomer may want to trade because her priors differ from those of the market. The bottom line is: the marginal bet in the Iowa market is pure speculation.

Table 2.1: Descriptive Statistics

		All Markets	Microsoft Markets	Computer Industry Markets
<i>N</i> :		2380	795	1585
Daily return:	Average	0.083	0.081	0.084
	St. dev.	1.563	1.169	1.728
	Skewness	27.7	16.8	28.2
	Kurtosis	980	325	953
Average Daily Volume (Units):			80.0	266.3
Measure of Nonsynchrony:	Average		0.006	0.019
	St. Dev.		0.049	0.062

Notes: The information is based on daily closing prices from the Iowa Experimental Markets, 8/95 till 12/96 (16 months). The closing price is the price of the last transaction or, if no transaction took place during the day, the previous closing price. Sometimes, contracts continued to be traded after their payoff became known; the corresponding observations were deleted. Also, the following were discarded because they clearly represented aberrations: All computer markets closing prices on 8/28/95 and 8/29/95; Microsoft Low closing price on 5/20/96; Computer industry markets closing prices on 6/20/96. *N* denotes sample size. Nonsynchrony is measured by the average absolute deviation of the sum of the closing prices of complementary contracts from unity.

Table 2.2: Analysis Of Payoffs On Winning Contracts

		All Markets	Microsoft Markets	Computer Industry Markets
<u>Winners</u>	N :	32	16	16
	$\bar{r}$: Average	0.115** (0.030)	0.097** (0.037)	0.173** (0.047)
	Skewness	2.4	3.5	2.0
	Kurtosis	9.9	16.2	9.4
	$\bar{x}$: Average	-0.033 (0.022)	-0.032 (0.022)	-0.035 (0.038)
	Skewness	-1.7	-1.6	-1.6
	Kurtosis	5.8	4.5	5.1
<u>Complements to Losers</u>	N :	64	16	48
	$\bar{r}$: Average	0.046** (0.008)	0.086** (0.026)	0.032** (0.006)
	Skewness	4.0	3.2	2.2
	Kurtosis	25.2	14.7	8.8
	$\bar{x}$: Average	-0.016 (0.010)	-0.028 (0.020)	-0.011 (0.011)
	Skewness	-3.9	-1.4	-4.9
	Kurtosis	21.8	3.7	30.0

Notes: N denotes sample size. $\bar{r}$ denotes the average across all month-long histories of the time series average daily traditional return. $\bar{x}$ denotes the average across all month-long histories of the time series average daily modified return, where the payoff is divided by the end-of-period closing price. Standard error in parentheses. * and ** indicate that the statistic is more than 1.8 and 2 standard errors from zero, respectively. The closing prices of complements to losers is obtained by subtracting the closing price of a loser from the price at which the market portfolio can be bought from the system (i.e., \$1). In the Microsoft case, the price history of the complements to the losers is not identical to that of the winners, because of nonsynchrony.

Table 2.3: Analysis Of Payoffs On Winning Contracts: Correlation With Lagged Information

		All Markets	Microsoft Markets	Computer Industry Markets
<u>Lagged Price</u>	N :	32	16	16
	$\overline{rp}$: Average	0.086** (0.004)	0.097** (0.004)	0.035** (0.007)
	Skewness	1.1	-0.4	0.9
	Kurtosis	6.2	2.4	5.4
	$\overline{xp}$: Average	-0.011 (0.008)	-0.019 (0.013)	-0.003 (0.009)
	Skewness	-1.8	-2.3	-0.2
	Kurtosis	8.3	8.4	3.4
<u>Lagged Price Squared</u>	N :	32	16	16
	$\overline{rp^2}$: Average	0.009** (0.002)	0.006** (0.003)	0.013** (0.004)
	Skewness	0.6	-0.8	1.5
	Kurtosis	5.3	2.4	5.9
	$\overline{xp^2}$: Average	-0.005 (0.005)	-0.011 (0.008)	0.000 (0.005)
	Skewness	-2.4	-2.5	-0.0
	Kurtosis	12.0	10.3	3.6

Notes: N denotes sample size. $\overline{rp}$ denotes the average across all month-long histories of the time series average daily traditional return multiplied by the lagged price level. $\overline{xp}$ denotes the average across all month-long histories of the time series average daily modified return multiplied by the lagged price level. $\overline{rp^2}$ denotes the average across all month-long histories of the time series average daily traditional return multiplied by the square of the lagged price level. $\overline{xp^2}$ denotes the average across all month-long histories of the time series average daily modified return multiplied by the square of the lagged price level. Standard error in parentheses. * and ** indicate that the statistic is more than 1.8 and 2 standard errors from zero, respectively.

Table 2.4: Analysis Of Payoffs On Winning Contracts: OLS Projections Of Returns Times Lagged Prices Onto Lagged Prices

	<u>All Markets</u>		<u>Microsoft Markets</u>		<u>Computer Industry Markets</u>	
	Return	Modified Return	Return	Modified Return	Return	Modified Return
N	601	601	299	299	301	301
Intercept	0.044** (0.010)	-0.014 (0.014)	0.056** (0.018)	-0.032 (0.030)	0.034** (0.013)	-0.011 (0.014)
Slope	-0.044** (0.017)	0.003 (0.024)	-0.067** (0.026)	0.021 (0.045)	-0.015 (0.028)	0.011 (0.031)
R^2	0.011	0.000	0.021	0.001	0.001	0.000

Notes: To compute the modified return, the payoff is divided by the end-of-period price instead of the beginning-of-period price. N denotes sample size. * and ** indicate that the statistic is more than 1.8 and 2 standard errors from zero, respectively.

Table 2.5: OLS Projections Of Changes In Prices Onto Lagged Price Levels

	All Markets	Microsoft Markets	Computer Industry Markets
N	1792	591	1200
Intercept	0.050** (0.006)	0.058** (0.013)	0.048** (0.007)
Slope	-0.053** (0.009)	-0.069** (0.019)	-0.048** (0.010)
R^2	0.020	0.022	0.019

Notes: N denotes sample size. * and ** indicate that the statistic is more than 1.8 and 2 standard errors from zero, respectively.

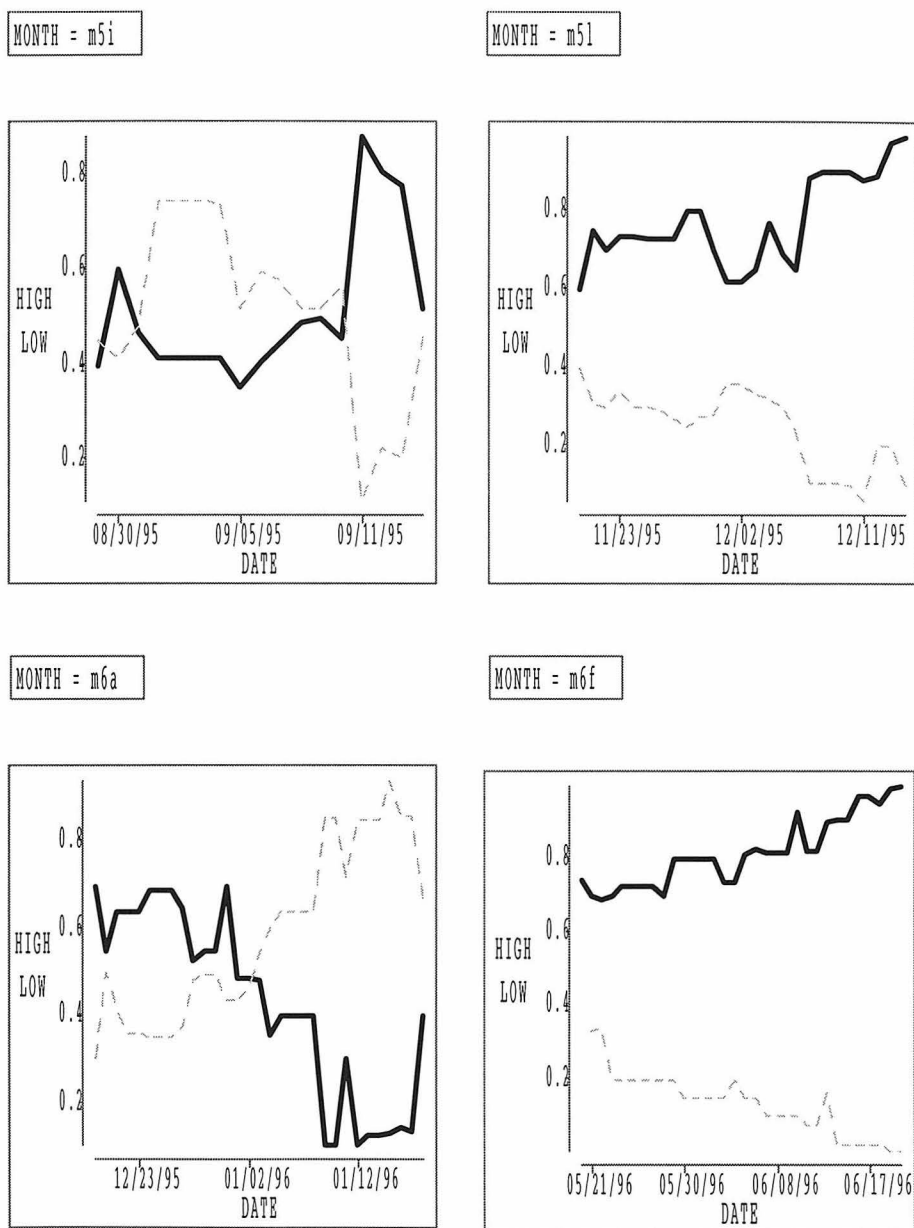


Figure 2.1: Time series plots of daily closing prices for contracts Microsoft High (bold line) and Microsoft Low (dashed line). One-month histories for four expiration months are shown: 9/95 (month m5i), 12/95 (month m5l), 1/96 (month m6a) and 6/96 (month m6f).

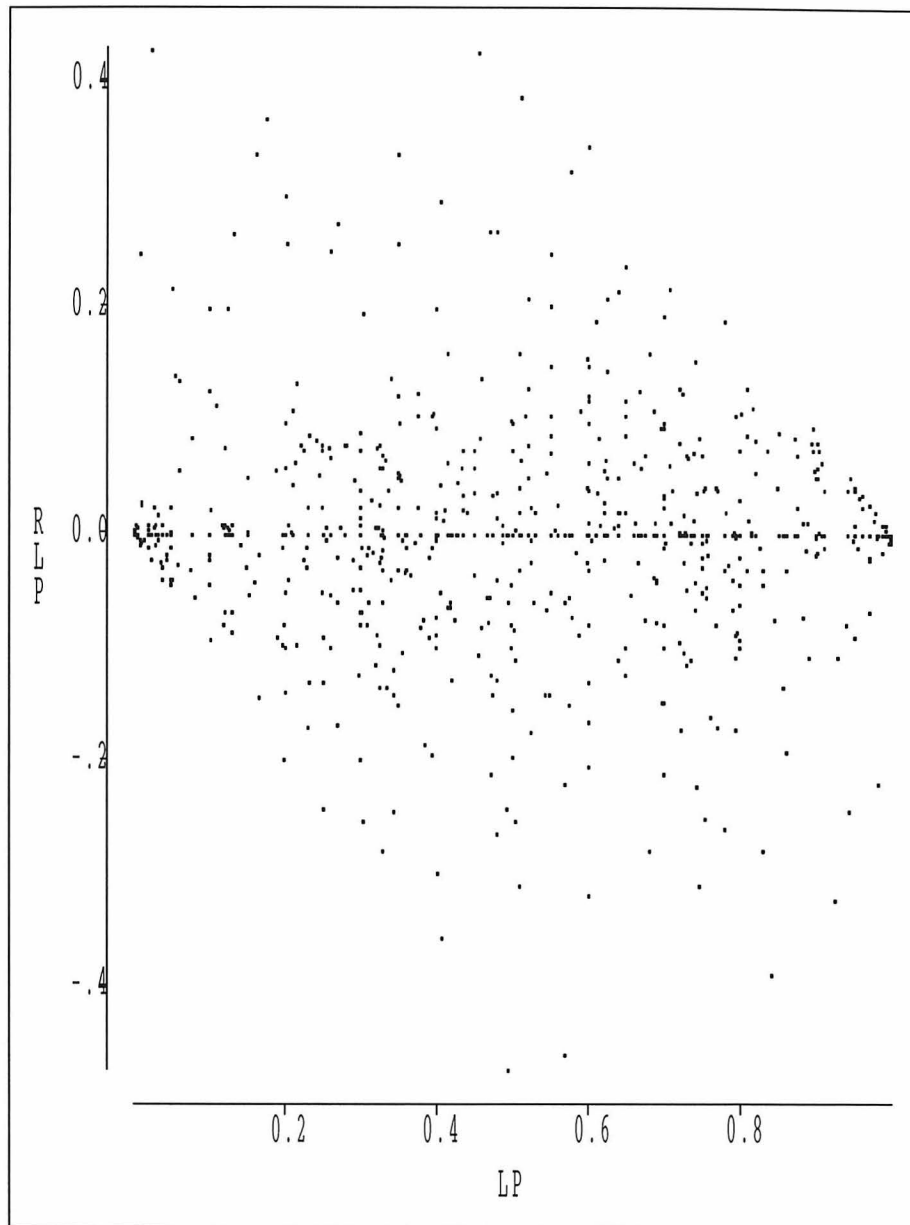


Figure 2.2: Scatter plot of daily changes in closing prices (RLP) onto previous-day closing prices (LP), Microsoft contracts.

Chapter 3 Testing Rationality of S&P 500 Index Options

3.1 Introduction

The purpose of this chapter is to demonstrate how rationality of learning can be tested in real financial markets, where no Arrow-Debreu (AD) securities are traded. The approach is implemented for the S&P 500 Index Options (SPX). We use prices of the S&P 500 Index Options to estimate prices of particular Arrow-Debreu securities, namely *digital options*. A *digital call* with exercise price X is a contract that pays \$1 when the maturity price of an underlying asset S_T exceeds X . Otherwise, it pays nothing. Thus, digital options (simply digitals) amount to straight bets on the terminal value of the underlying asset S_T .

Prices of digital options are implicit in prices of corresponding puts and calls and closely related to cumulative *risk-neutral densities* (CRND).¹ For any exercise price X , asset's risk-neutral density (RND) (or *state-price density* (SPD)) defines the risk-neutral probability that the asset's maturity price falls between X and $X + dX$, while the CRND determines the risk-neutral probability that the asset's maturity price does not exceed X . The idea to estimate the RND from prices of traded financial securities has been suggested by Banz and Miller (1978), Breeden and Litzenberger (1978) and Ross (1976). Both *parametric* and *nonparametric* approaches have been implemented in the empirical literature.

In the parametric approaches, either the underlying asset's price dynamics or the RND are prespecified in some parametric form.² The problem with parametric techniques, however, is that their success (or failure) critically depends on the ability

¹The concept is based on risk-neutral valuation of options (See, Cox and Ross (1976)).

²See Jarrow and Rudd (1982), Shimko (1993), Madan and Milne (1994). Parametric specifications for stochastic processes other than geometric Brownian motion can be found in Goldenberg (1991), Hull and White (1987), and Heston (1993).

to capture the underlying asset's price dynamics: a misspecification of the dynamics leads to systematic errors.

One recent response to these difficulties is to estimate the RND nonparametrically. These approaches use neural networks (Hutchinson, Lo, and Poggio (1994)), kernel regressions (Aït-Sahalia and Lo (1995)), or optimization methods (Rubinstein (1994), Jackwerth and Rubinstein (1996), Stutzer (1996)) to fit functional “black boxes” that map market option data (e.g., time to maturity, exercise price, underlying asset price, dividend yields, interest rates, etc.) into option prices. With no parametric restrictions on the underlying asset's price process, these approaches are “distribution-unrestricted.” Consequently, they are free of the classical *joint hypothesis* on asset's price dynamics *and* risk premium that is typical of parametric arbitrage models.

The new estimation procedure that we use to recover prices of digital options is in the spirit of nonparametric techniques and has several interesting features.

First, and most importantly, the estimation procedure does not require *stationarity* of an underlying asset's price dynamics. Evidence against the stationarity assumption abounds in the empirical finance literature.³ Macroeconomic variables (for example, level of stock market prices, volatilities, interests rates, etc.) may affect stock and options prices in a way that is yet not fully understood. The functional dependence on these variables, unaccounted in parametric or nonparametric specifications, causes the nonstationarity of the underlying asset's process (and, consequently, the option pricing formula). That is why we estimate the option pricing formula *separately* for each trading day using only most recent options prices. Since we study how investors change their beliefs over time, the empirical option formula is allowed to vary over time as well.

Second, the estimation procedure enforces rudimentary theoretical shape restrictions on options prices (such as monotonicity, convexity and boundary conditions). An iterative algorithm is suggested to solve the resulting nonlinear Constrained Con-

³See, for example, empirical study of options prices in Rubinstein (1985). He found that options prices exhibited different biases for the two studied subperiods. For the first period from August 23, 1976, to October 21, 1977, implied volatilities tended to be relatively high for in-the-money options and relatively low for out-of-the-money options. For the second period from October 22, 1977 to August 31, 1978, the exactly opposite results were obtained.

vex Regression (CCR).

Finally, the estimation procedure uses an alternative parametric option specification which is *parsimonious*, computationally *very simple*, and which significantly outperforms the Black-Scholes specification for the studied dataset.

The analysis of time series of the estimated digital options supports the hypothesis of rational learning. Several tests based on the martingale restriction under Rational Learning Equilibrium (RLE) indicate that investors of the S&P 500 Index Options process new information to the market in an efficient manner. This means that investors' reaction is consistent with revisions prescribed by Bayes' rule: they do not *over-* or *under-react* to new information.

Still, strong evidence against rational expectations surfaces, implying that investors *must* have started with biased priors. The dataset allows one to characterize the investors' expectational errors. In particular, it is shown that, over the five-year period from 1991 to 1995, investors repeatedly underestimated the S&P 500 Index expected return and overestimated its volatility.

The remainder of the chapter is organized as follows. Section 3.2 reviews the relation between prices of digitals and traditional options. Section 3.3 describes our dataset, the test specification, and the estimation procedure to recover prices of digitals. Section 3.4 presents the empirical results of the tests for rationality of learning and rational expectations. Section 3.5 concludes.

3.2 Constructing and Pricing Digital Options

Let $C_t(X)=C(S_t, \mu_{t,T}, \delta_{t,T}, t; X, T)$ and $P_t(X)=P(S_t, \mu_{t,T}, \delta_{t,T}, t; X, T)$ denote the time- t values of European call and put with maturity date T and exercise price X written on an underlying asset (index) with a current price S_t , where

- $\delta_{t,T}$ = the dividend rate between t and T ;
- $\mu_{t,T}$ = the risk-free rate of return for maturity T (continuously compounded).

As usual, we assume that both the underlying asset and the options are traded in perfect markets with no transactions costs or tax considerations and that unlimited risk-free borrowing and lending is available at the same instantaneous rate $\mu_{t,T}$. Throughout the chapter we often use the following notations:

- $\tau \equiv T - t =$ time to maturity;
- $m \equiv S/X =$ moneyness;
- $F_t \equiv S_t e^{-\delta_{t,T}(T-t)} =$ “dividend-adjusted price”,⁴
- $G_t(X) \equiv X e^{-\mu_{t,T}(T-t)} =$ value of the risk-free bond that pays X dollars at time T . We can think of $G_t(X)$ as “interest rate adjusted” exercise price, and, when it is unambiguous, we suppress the dependence on X and simply write G_t .

Options are derivative securities whose payoff depend on the maturity price of the underlying asset, S_T . For example, the payoff of a call option:

$$C_T(X) = \begin{cases} S_T - X, & \text{if } S_T > X \\ 0, & \text{if } S_T \leq X \end{cases} \quad (3.1)$$

Consider another derivative security that has a payoff structure similar to (3.1), a *digital (or binary) call*. A European digital call with exercise price X is a contract that pays \$1 when the maturity price of the underlying asset S_T exceeds X and zero dollars otherwise. Thus, a digital option amounts to a straight bet on whether the underlying asset terminal value S_T will be above the exercise price X . Its payoff is independent of how far above the exercise price the asset value is at the time of maturity.

The value of the digital call is denoted $DC_t(X) = DC(S_t, \mu_{t,T}, \delta_{t,T}, t; X, T)$ and its payoff is given by:

$$DC_T(X) = \begin{cases} 1, & \text{if } S_T > X \\ 0, & \text{if } S_T \leq X \end{cases} \quad (3.2)$$

⁴More precisely, since $S_t e^{(\mu_{t,T} - \delta_{t,T})(T-t)}$ is equal to the asset’s futures price at time T (or investors’ expectations about S_T), $S_t e^{-\delta_{t,T}(T-t)}$ represents “*present value* of the asset’s futures price.”

Also referred to as *cash-or-nothing* or *binary* options (they are discussed in most standard references on options, such as Hull (1997), Wilmott et al. (1993), or Ritchken (1996)), digitals are simple, but powerful tools for pricing complex contingent claims. And, as discussed before, they can be used to test rationality of learning in financial markets, the primary focus of this chapter.

When prices of digitals are unobservable, they can be inferred from prices of traded options written on the same underlying asset. Suppose there exists a collection of calls and puts with an infinitely dense set of exercise prices from zero to infinity. Then, in a *dynamically complete market*, the digital call $DC_t(X)$ can be priced by arbitrage arguments.⁵

Consider a replicating portfolio constructed by buying one call with exercise price X , $C_t(X)$, and selling one call with exercise price $X + \varepsilon$, $C_t(X + \varepsilon)$. The payoff of $1/\varepsilon$ shares of the portfolio at time T is given by:

$$U_T(X) = \begin{cases} 1, & \text{if } S_T > X + \varepsilon \\ 1/\varepsilon (S_T - X), & \text{if } X \leq S_T \leq X + \varepsilon \\ 0, & \text{if } S_T < X \end{cases}$$

As ε tends to 0, the limit of the last expression is equal to the payoff of the digital call $DC_t(X)$ in (3.2), so

$$\lim_{\varepsilon \rightarrow 0} U_t(X) = DC_t(X).$$

At time t , the portfolio price is given by

$$\frac{1}{\varepsilon}(C_t(X) - C_t(X + \varepsilon)),$$

which, in the limit, equals

$$-\frac{\partial C_t(X)}{\partial X}.$$

⁵The following derivation is based on the approach suggested by Breeden and Litzenberger (1978) for pricing state contingent claims implicit in options prices.

Consequently, we obtain that

$$DC_t(X) = -\frac{\partial C_t(X)}{\partial X}. \quad (3.3)$$

We can also use the put-call parity relation

$$C_t(X) + G_t = P_t(X) + F_t \quad (3.4)$$

to rewrite $DC_t(X)$ as follows

$$DC_t(X) = e^{-\mu_t, T(T-t)} - \frac{\partial P_t(X)}{\partial X} = e^{-\mu_t, T(T-t)} \left(1 - \frac{\partial P_t(X)}{\partial G_t} \right). \quad (3.5)$$

For instance, imagine that the Black-Scholes option model holds, where the risk-free and dividend rates are constant and the underlying asset has a log-normal distribution with constant variance σ^2 . The call price is given by

$$C_t^{BS}(X) = F_t \Phi(d_1) - G_t \Phi(d_2),$$

where

$$d_{1,2} = \frac{\ln(F_t/G_t) \pm \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}},$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function, $F_t = S_t e^{-\delta(T-t)}$ and $G_t = G_t(X) = X e^{-\mu(T-t)}$. From (3.3) we compute the price of the corresponding digital call, DC_t^{BS} :

$$DC_t^{BS} = e^{-\mu(T-t)} \Phi(d_2).$$

In practice, the theoretical closed-form option formula is not known. In this case, options with available exercise prices X can be used to estimate the derivative $-\frac{\partial C_t(X)}{\partial X}$. Then, using (3.3) (or (3.5)) we can determine prices of the corresponding digitals and test the rationality constraints from Chapter 1. The next section describes the dataset and the estimation procedure.

3.3 The Dataset and Test Specification

3.3.1 S&P 500 Index Options

The data for our empirical analysis are daily closing prices of the S&P 500 Index Options for the 5-year period from January 1991 to December 1995. This sample contains a total of 129,264 pairs of call and put option prices.⁶

The S&P 500 Index (formally known as the Standard & Poor's 500 Composite Stock Price Index) is a capitalization-weighted index of 500 stocks from a broad range of industries. It represents about 70 percent of all of the U.S.-traded common stocks. The S&P 500 Index Options (traded on the Chicago Board Options Exchange (CBOE), the ticker symbol: SPX) are the world's most actively traded European-style index options.⁷ Because of high liquidity, the market of the S&P 500 Index Options presents a particularly interesting application for the new methodology.

Over the 5-year period from January 1991 to December 1995, the mean and standard deviation of continuously-compounded daily returns of the S&P 500 Index was 13.20% and 10.32%, respectively.⁸ The S&P 500 Index prices are shown in Figure 3.1. Short-term interest rates during the period ranged from 2.86 to 6.22 with the mean equal to 4.31.⁹ Short-term interest rates are shown in Figure 3.2. Simple descriptive statistics of the dataset are presented in Table 3.1.

Since our choice of test specification has been partially affected by the actual rules that the CBOE uses to introduce options to the market, we briefly summarize the most important features of these rules.¹⁰ At any one time, the S&P 500 Index Options have six unique expiration months: the three near-term months and the

⁶Some of these prices were missing due to zero trading volume of the corresponding options. Of all 129,264 options pairs only 69,082 puts and 46,123 calls had positive volume.

⁷In 1995, the total volume of the SPX Options was 26,726,023 contracts and the average daily volume was more than 105,000 contracts.

⁸We follow the common convention in equity markets to report annual returns and their means calculated on the basis of *calendar* days (365 days per year) as opposed to standard deviations calculated on the basis of *trading* days (the standard assumption that there are 252 trading days per year).

⁹The risk-free rate $\mu_{t,T}$ is approximated as the yield of the 3-month Treasury bill that matures closest to the option expiration date. Since in what follows we analyze options with short lifetimes, small measurement errors in interest rates should have very little.

¹⁰See, for example, Hull (1997) or Ritchken (1996) for more details.

following three months from the March quarterly schedule (March, June, September, December). The expiration date is Saturday immediately following the third Friday of the expiration month. The minimum price tick is $1/16$ for options trading below 3 and $1/8$ for options trading above 3.¹¹ The exercise prices are multiples of 5 for the near-term months and 25 for far months. When options expire and new series are introduced, the CBOE sets the two exercise prices closest to the current S&P 500 Index price. If the current price is very close to one of those exercise prices, a third exercise price is listed to better bracket the current price. If the index trades through the highest or lowest exercise price currently available, another exercise is generally introduced for all expiration dates, to bracket that price.

The closing prices for S&P 500 Index Options data present a number of practical problems that must be taken into account in any empirical research.

First, the underlying index prices and the options prices are not necessarily perfectly synchronous. There is no guarantee that the closing transaction for each option is recorded at the same time as the closing index price. In fact, for thinly traded options this mismatch can be significant and, occasionally, can indicate that pure arbitrage opportunities are available. For example, the closing price of a deep-in-the-money option might correspond to a trade at 12:00 p.m. (and when the index was, say, 665), while the closing index price is recorded at 5:00 p.m. (when the index was 657). In this case, the reported option price does not reflect the actual price at which an option could be bought or sold for when the underlying asset stopped trading at the end of the day. Since most option trading takes place in at-the-money or out-of-the-money short-term contracts, the closing prices of in-the-money and long-term options are particularly unreliable. A related problem in using closing prices is the irregularities caused by the bid-ask spread.¹²

Second, the S&P 500 Index usually pays a dividend. At the time of trade, the expected future dividend stream is not known with certainty since it depends on

¹¹One option point is equal to \$100. Thus, the minimum price ticks are \$6.25 and \$12.5, respectively.

¹²Several papers in the finance literature have studied the potential biases from using transaction prices, starting with Blume and Stambaugh (1983).

the exact amount and timing of the dividends of the stocks that compose the index. In principle, future index dividends can be estimated in two ways. They may be reasonably well extrapolated from the past. For example, the dividends paid in January last year are assumed to be equal to the dividends paid in January this year. Alternatively, future index dividends may be measured as though they have been forecast perfectly. In this case, the dividends are assumed to be equal to the dividends actually paid by the S&P 500 Index. Neither approach, however, yields investors' *true* expectations about the future index dividends at the time of trade. Due to the variability of dividends over time, the *implied future price*, $S_t e^{(\mu_t, T - \delta_{t,T})(T-t)}$, could sometimes be above the spot index price, while at other times it could be below the spot price S_t .

The other real-world considerations relevant to our analysis include:

a) Asymmetry between options with below- ($m < 1$) and above-the-money ($m > 1$) exercise prices: Over the studied period, the number of listed below-the-money options, on average, exceeded substantially the number of listed above-the-money options. Only 46,272 out of 129,964 observations (or less than 36 percent) had an above-the-money exercise price. Of course, this just reflects the fact that the S&P 500 Index rose at about 13.2% annual rate over the five-year period and the way the CBOE sets exercise prices. See also Table 3.1.

b) Higher liquidity of out-the-money options compared to in-the-money options: The average daily trading volume for puts that were 25 points out-of-the-money ($m > 1$) was 597 contracts, while the average daily trading volume for puts that were 25 points in-the-money ($m < 1$) was 49 contracts, more than ten times less. The average trading volumes for calls that were 25 points out-of-the-money ($m < 1$) and 25 points in-the-money ($m > 1$) were 389 and 116 contracts, respectively. See also Table 3.1.

c) Higher trading activity of puts compared to that of calls: The total trading volume of put contracts was about 30% higher than that of call contracts. This must reflect the higher demand by portfolio managers for protective puts.

d) Correlation of options time series: Since we study the options written on the

same underlying asset, the returns of overlapping options are strongly correlated.

e) Discreteness of options prices: Since there is a minimum price tick, sometimes several out-of-the-money options with different exercise prices and one maturity date may have the same closing price. For example, on 01/08/92 the index was 418.10 and the put contracts with exercises 355, 365, 370 and 375 and the time-to-maturity of 7 days had the same closing price of 0.1875 points. One could then incorrectly conclude that the market did not price the additional value of the put with the exercise 375 (compared to the put with the exercise 355) or, equivalently, that the market put zero probability for the index to close at $[355, 375]$ interval on the expiration date.

To address these issues we proceed as follows. We divide the option data into separate groups corresponding to different expiration dates. In each group, we discard observations with time-to-maturity $\tau = T - t$ of more than 25 or less than 6 trading days. This yields a cross section of 58 nonoverlapping subperiods.¹³ Each of these subperiods has closing options prices for 20 trading days which, approximately, covers one calendar month ending one calendar week before the maturity.¹⁴ Short-lived options ($\tau \leq 5$) are dropped because of extreme variation of their returns which lead to even more extreme returns of the estimated digitals.¹⁵

The resulting sample represents a quite broad cross section of most liquid options and contains a total of 16,374 puts and 13,375 calls with positive trading volume for $58 \times 20 = 1160$ trading days. We denote by $t_{i,j}$ ($i = 1, \dots, 58$ and $j = 1, \dots, 20$) a trading day from subperiod i (corresponding to expiration date T_i) with time-to-maturity $\tau_j = (26 - j)$.

3.3.2 Estimation Procedure

For each trading day $t_{i,j}$ ($i = 1, \dots, 58$ and $j = 1, \dots, 20$) we approximate the empirical put formula $P_t(X)$ in the following three steps.

¹³Too short series of options with expiration in January, 1991, has been dropped. Also, there were no option series with expiration in November, 1992.

¹⁴Time to maturity is measured in trading days. See also footnote 8.

¹⁵Note also that the discreteness of options prices is particularly relevant for out-of-the-money options that are close to maturity.

In the *first step* we obtain *reliable* empirical put prices. The idea here is to exploit the fact that out-of-the-money options are more liquid compared to in-the-money options. This step is a slight modification of the procedure that has been used in Ait-Sahalia and Lo (1995). More specifically, define

$$\bar{P}_t(X) \equiv G_t(X) + C_t(X) - F_t. \quad (3.6)$$

The put-call parity relation (3.4) implies that, at least in theory, $P_t(X) = \bar{P}_t(X)$.¹⁶ When the put option $P_t(X)$ is in-the-money and illiquid, the corresponding call, $C_t(X)$, is out-of-the-money and relatively more liquid. Therefore, $\bar{P}_t(X)$ calculated from $C_t(X)$ is more reliable than $P_t(X)$. Similarly, out-of-the-money puts are more liquid than the corresponding in-the-money calls. One difficulty in using the relation (3.6) is that the term $F_t = S_t e^{-\delta_{t,T}(T-t)}$ is not known since the expected dividend rate $\delta_{t,T}$ is not observable. However, since F_t is invariant across all exercise prices X , we can estimate it from actively traded and reliable at-the-money options pairs. Specifically, let X^k ($k = 1, 2, 3, 4$) denote the four exercise prices closest to at-the-money. For each of these exercise prices, we define $F_t^k = (C_t(X^k) + G_t(X^k) - P_t(X^k))$ and set F_t to be the average of F_t^k ($k = 1, 2, 3, 4$).¹⁷

After that, we replace every illiquid in-the-money put $P_t(X)$ with $\bar{P}_t(X)$ computed

¹⁶Several empirical studies have confirmed that put-call parity is very rarely violated in practice when *synchronous* quotes are studied. Since it requires just a few rational agents in the market to eliminate any pure arbitrage opportunity, these opportunities can only exist for very short time and, oftentimes, there is no way of exploiting them. See, for example, Bhattacharya (1983), Galai (1977, 1978), Gould and Galai (1974), Klemkosky and Resnick (1979, 1980), Rubinstein (1985).

Recently Kamara and Miller (1995) have carried out the empirical study of put-call parity for the S&P 500 Index Options. They have found that violations of put-call parity for European-style options are even less frequent and smaller than for American-style options.

When *nonsynchronous* closing prices are used, we may expect $P_t(X)$ to differ from $\bar{P}_t(X)$, especially for far-from-the-money exercise prices.

¹⁷Note that in this step we essentially estimate the S&P *implied* futures price, $S_t e^{(\mu_{t,T} - \delta_{t,T})(T-t)} = F_t e^{\mu_{t,T}(T-t)}$, from observable option prices. As an alternative, one could use quotes of the S&P futures *actually* traded on the Chicago Mercantile Exchange (CME) rather than implied futures prices. However, the latter approach is not necessarily preferable. First, quotes from the two markets, CBOE and CME, may be slightly nonsynchronous. Second, when we compute the implied futures price as an average across different exercise prices, we effectively “smooth” or “cancel-out” irregularities caused by the bid-ask spread and price discreteness. Those irregularities can be substantial when one uses actual futures prices. Finally, the put-call parity relation (3.4) is typically satisfied more accurately when the implied futures price is used as opposed to the actual futures price.

from the liquid call $C_t(X)$ using (3.6). For the four closest to at-the-money exercise prices X^k , we replace $P_t(X)$ with $1/2(P_t(X) + \bar{P}_t(X))$. After all reliable information has been extracted from liquid call prices, we discard them and from now on work with put prices only.

Even though we have obtained reliable empirical put prices, they do not necessarily satisfy theoretical restrictions on the option formula. In particular, for any option pricing model,

- (i) put price $P_t(X)$ is an increasing convex function of exercise price X ;
- (ii) $0 \leq \frac{\partial P_t(X)}{\partial X} \leq e^{-\mu_{t,T}(T-t)}$ for all exercise prices X .¹⁸

Nonsimultaneity of stock and options prices, the bid-ask spread, the minimum price tick and other real-world imperfections may lead to local violations of these conditions. So, in the *second step*, we correct this by running the *constrained convex regression* (CCR). Specifically, let $\{P^i, X^i\}_{i=1}^n$ represent a collection of put and exercise prices corresponding to some trading day t and expiration day T , where $P^i \equiv P_t(X^i)$ and $X^1 < X^2 < \dots < X^n$. Our goal is to find a function $\{Y^i\}_{i=1}^n$ which is the closest (in the sense of least squares) to $\{P^i\}_{i=1}^n$ and which satisfies the shape restrictions (i) and (ii).

Mathematically,

$$\text{Minimize}_{\{Y^i\}_{i=1}^n} \sum_{i=1}^n (Y^i - P^i)^2 \quad (3.7)$$

¹⁸These restrictions are equivalent to existence of the asset's risk-neutral density. Indeed, let $\Psi_t(X)$ denote the asset's cumulative risk-neutral density (CRND), i.e., it determines *risk neutral* probability that the asset's maturity price S_T does not exceed X . In view of (3.5), $\Psi_t(X)$ is equal to $\frac{\partial P_t(X)}{\partial G_t}$. Then, the second condition simply states that CNRD must be between 0 and 1, i.e.,

$$0 \leq \Psi_t(X) \leq 1, \quad \text{for all exercises } X.$$

The first condition states that the asset's CNRD is a nondecreasing function of X , i.e.,

$$\Psi_t(X) \leq \Psi_t(X + \varepsilon), \quad \text{for } \varepsilon \geq 0,$$

and, consequently, the asset's risk-neutral density (RND) $\psi_t(X)$ is nonnegative:

$$\psi \equiv \frac{\Psi_t(X)}{\partial X} \geq 0.$$

subject to

$$Y^2 - Y^1 \geq 0, \quad (3.8)$$

$$\frac{Y^{i+1} - Y^i}{X^{i+1} - X^i} \geq \frac{Y^i - Y^{i-1}}{X^i - X^{i-1}}, \quad i = 2, \dots, n-1, \quad (3.9)$$

$$\frac{Y^n - Y^{n-1}}{X^n - X^{n-1}} \leq e^{-\mu_t, T(T-t)}. \quad (3.10)$$

To solve the constrained least squares problem (3.7)-(3.10), we use a modification of the algorithm by Dykstra (1983). The details on the implemented realization can be found in the Appendix.

Up to this point, the estimation procedure is completely nonparametric because *any* put function that satisfies rudimentary shape conditions (i) and (ii) is admissible.

The put prices obtained as a result of the CRR (3.7)-(3.10) are free of noise and arbitrage violations and represent a discrete approximation of the true put function $P_t(X)$. Of all discrete approximations that satisfy the shape restriction (i) and (ii), the approximation $\{\hat{P}^i\}_{i=1}^n$ is closest to the empirical data $\{P^i\}_{i=1}^n$ (in the sense of the least squares objective functional).

To compute prices of digitals, we need to approximate the derivative $\frac{\partial P_t(X)}{\partial X}$. One may be tempted to use put spreads with consecutive exercise prices to approximate the derivative $\frac{\partial P_t(X)}{\partial X}$. For example, when $X \in (X^j, X^{j+1})$ one can use the first-order finite difference approximation

$$\frac{\partial P_t(X)}{\partial X} \approx \frac{\hat{P}^{j+1} - \hat{P}^j}{X^{j+1} - X^j}, \quad (3.11)$$

or the second-order finite difference approximation

$$\frac{\partial P_t(X)}{\partial X} \approx \quad (3.12)$$

$$\frac{(X^{j+1} + X^{j+2} - 2X)(\hat{P}^j - \hat{P}^{j-1}) + (2X - X^{j-1} - X^j)(\hat{P}^{j+2} - \hat{P}^{j+1})}{X^{j+2} + X^{j+1} - X^j - X^{j-1}}.$$

Unfortunately, when the grid step (the difference between two consecutive exercise

prices, $\Delta X = X^{j+1} - X^j$) is relatively large, the finite difference schemes (3.11)-(3.12) may yield very rough approximations of the derivative $\frac{\partial P_t(X)}{\partial X}$ and may lead to systematic errors.¹⁹ This is particularly critical for near-to-maturity options when $\Delta X = X^{j+1} - X^j = \5 is comparable to total subperiod movements of the Index. In this case, the derivative $\frac{\partial P_t(X)}{\partial X}$ changes significantly from X^j to X^{j+1} for near-the-money exercise prices (i.e., even though ΔX is equal to only 1 to 2 percent of the Index value, the probability of S_T ending in the interval $[X^j, X^{j+1}]$ is substantial).

Another popular approach to estimate the derivative $\frac{\partial P_t(X)}{\partial X}$ is nonparametrical where the function $\{\hat{P}^i\}_{i=1}^n$ is smoothed by some kernel regression and then differentiated with respect to exercise price X .²⁰ However, this approach is also not suited for small samples and when the grid step is large.²¹

It turns out, that the best fit for $\frac{\partial P_t(X)}{\partial X}$ obtains when we use continuous *parametric* approximations of discrete put prices $\{\hat{P}^i\}_{i=1}^n$. Therefore, in the *third* step, we specify some smooth parametric put formula, $P(X, \xi)$, with a vector of unknown parameters ξ to be estimated in the following nonlinear least squares problem:

$$\text{Minimize}_{\xi} \sum_{i=1}^n (\hat{P}^i - P(X^i, \xi))^2. \quad (3.13)$$

We have tried different specifications $P(X, \xi)$, including the Black-Scholes put formula,

$$P_t^{BS}(X) := -F_t \Phi(-d_1) + G_t \Phi(-d_2), \quad (3.14)$$

for which the vector of unknown parameters $\xi = (\sigma_t^2, F_t)$ is estimated to solve the minimization problem (3.13). In this case, the resulting Black-Scholes approximation

¹⁹In particular, the RND resulting from discrete approximations (3.11)-(3.12) will have standard deviation biased upward.

²⁰Note that *any* uniformly positive kernel smoother preserves the shape restrictions (i) and (ii).

²¹Estimating a derivative from an empirical function (as well as recovering a density from an empirical distributional function) is an ill-posed problem. This means that, when $\hat{f}(X)$ is a tight smooth approximation to $f(X)$, its derivative $\frac{d\hat{f}(X)}{dX}$ does not necessarily approximate $\frac{df(X)}{dX}$ well. See, for example, Vapnik (1982, p.22). Solutions to ill-posed problems can be obtained using regularization methods (Tikhonov (1963)). Regularization can be also implemented in some nonparametric estimators. Still, accuracy of nonparametric estimators which require smoothing of empirical data is often poor for small samples because of the bias caused by the smoothing.

exhibits the well known *volatility smile* bias, i.e., it consistently underprices far-from-the-money options and overprices near-the-money options.²² For estimated digital options, this bias means that prices of far-of-the-money digitals are less extreme than true ones, i.e., the Black-Scholes approximation consistently overprices out-of-the-money ($m < 1$) and underprices in-the-money ($m > 1$) digital options. The effect of the volatility smile bias is quite pronounced in our analysis but it is not substantial enough to overturn the main conclusions of the following sections.²³

Nevertheless, after some experimentation, we have decided to use an alternative parametric specification that considerably outperforms the Black-Scholes one for the studied dataset of the S&P 500 Index Options.

More specifically, let us postulate that true function $P_t(X)$ satisfies the following relationship. For any exercise price X ,

$$P_t(X)(F_t - G_t + P_t(X)) = \alpha_t F_t G_t, \quad (3.15)$$

where α_t is a constant independent of G_t or F_t . The only positive solution to the quadratic equation (3.15),

$$P_t^Q(X) = \frac{1}{2} \left((G_t - F_t) + \sqrt{(G_t - F_t)^2 + 4\alpha_t F_t G_t} \right), \quad (3.16)$$

²²The volatility smile effect persists even when we allow for maximum flexibility by estimating the vector of unknown parameters $\xi = (\sigma_t^2, F_t, \mu_{t,T})$. This gives the Black-Scholes formula the best chance because *all* independent parameters in formula (3.14) (including the risk-free rate $\mu_{t,T}$) are implied simultaneously.

²³Pricing biases associated with the Black-Scholes option pricing model have been well documented. Studies by Black (1975), Emanuel and MacBeth (1982), MacBeth and Merville (1979) and Rubinstein (1985, 1994) have all reported that the Black-Scholes model tends to systematically misprice deep in-the-money and deep out-of-the-money options. Rubinstein (1994) has also pointed out that exercise price biases associated with the Black-Scholes model have been especially severe for S&P500 Index Options since the October 1987 market crash and that the magnitude of these biases has tended to increase since then.

These documented biases are generally thought to result from the parsimonious assumptions of the Black-Scholes model such as geometric Brownian motion, constant volatility, frictionless markets, etc. They, in turn, imply that stock prices are lognormally distributed. More complex parametric specification can be used to overcome the limitations of the Black-Scholes model. For example, Jarrow and Rudd (1982) have proposed an approximate option valuation formula which has additional degrees of freedom which allow one to account for skewness and kurtosis deviations from lognormality in stock price distributions.

satisfies the shape restrictions (i) and (ii). In addition, the “quadratic” parametric specification in (3.16) has several desirable properties. In particular, it satisfies the following two boundary conditions:

$$\begin{aligned} P_t^Q(0) &= 0, \\ \lim_{X \rightarrow \infty} \frac{\partial P_t^Q(X)}{\partial G_t} &= 1, \end{aligned}$$

which one expects from *any* parametric or nonparametric put formula. Other interesting properties of $P^Q(X)$ are discussed in the Appendix.

We solve the nonlinear least squares problem (3.13) for a vector of unknown parameters $\xi = (\alpha_t, F_t)$ to find the best approximation $P_t^Q(X)$ in the class of functions defined in (3.15). The quadratic approximation, parsimonious and computationally very simple, captures the empirical data surprisingly well.²⁴ For comparison, it outperforms the Black-Scholes approximation in 96 percent of cases (or, in other words, only accidentally does it perform worse than the Black-Scholes approximation). The average standard error (across 1160 trading days) is 0.205 for the quadratic approximation and 1.242 for the Black-Scholes approximation, respectively.²⁵ The quadratic approximation merits separate investigation and we leave it for future research.

The estimation procedure of this section is discussed in more detail in the Appendix where we report intermediate results of the above three steps for one trading day of the dataset. When illustrating the third step of the procedure we also compare results for the quadratic approximation with those for the Black-Scholes approximation. After the parametric put formula $P_t(X)$ and its derivative with respect to exercise price $\frac{\partial P_t(X)}{\partial X}$ have been estimated for each trading day, we can now define digital options and compute their prices.

²⁴A multivariate Newton-Raphson search is implemented to solve the nonlinear least squares Problem (3.13) for both the quadratic and Black-Scholes specifications. However, the procedure is much simpler for the quadratic approximation. Since Equation (3.15) is *linear* in both F_t and $\alpha_t F_t$, “good” *initial* values for α_t and F_t can be easily estimated. See the Appendix for details.

²⁵In both specifications, less reliable puts with exercise prices very far from index current price S_t ($|S_t - X| > 25$) have been eliminated to improve accuracy of the approximations near to at-the-money. This also reduces the bias caused by larger number of out-of-the-money puts ($m > 1$) compared to in-the-money puts ($m < 1$).

3.3.3 Digital Options and Their Prices

For every subperiod i ($i = 1, \dots, 58$) we define a set of five digital calls, $DC1_i - DC5_i$, with exercise prices $X1_i - X5_i$ related to the “initial” subperiod- i index price $S_{i,1}$ as follows:

$$X1 = 0.98 S_1,$$

$$X2 = 0.99 S_1,$$

$$X3 = 1.00 S_1,$$

$$X4 = 1.01 S_1,$$

$$X5 = 1.02 S_1,$$

for $i = 1, \dots, 58$.

To simplify exposition, we refer to all digitals $DC1_i$ with exercise prices $X1_i$ ($i = 1, \dots, 58$) as one security $DC1$, a digital option with exercise price $X1$ which is always 2%-in-the-money at the beginning of subperiod- i . The idea behind this definition is that the digital $DC1$ follows similar underlying dynamics for different subperiods and, thus, the cross section of prices of the digital $DC1$ can be used to test the rationality restrictions. Similarly, at the beginning of each subperiod, $DC2$ is 1%-in-the-money, $DC3$ is at-the-money, and $DC4$ and $DC5$ are 1%- and 2%-out-of-the-money. Figure 3.3 displays S&P 500 Index prices and the exercise prices $X1 - X5$ of the constructed digitals $DC1 - DC5$ for 12 subperiods of 1993. For most trading days, the exercise prices $X1 - X5$ bracket the index fairly well. Also, the exercise prices are not too far from at-the-money which assures that prices of digitals $DC1 - DC5$ can be reliably estimated by liquid puts and calls.

We use put formulae $P_t(X)$ estimated in the previous section for all trading days and equation (3.5) to compute the equilibrium prices of the digitals $DC1 - DC5$. Figure 3.4 plots the price paths of the digitals $DC1 - DC5$ for four sample subperiods.

Figure 3.4 also presents an additional evidence of *nonstationarity* of the index price dynamics. Note that, at the beginning of each subperiod, corresponding digitals have

the same maturity ($\tau = 25$ trading days) and moneyness m (by definition of the exercise prices $X1 - X5$). Therefore, almost all stationary option models would yield the same put and call prices (and, consequently, same prices of digitals) on first trading day of each subperiod. Inspection of four different subperiods in Figure 3.4 shows that this is clearly not the case. Initial prices of the digital options vary considerably across subperiods. For example, they range from 0.71 to 0.84 for $DC1$ and from 0.22 to 0.38 for $DC5$.

3.4 Empirical Results

In this section we use the martingale restrictions on prices of Arrow-Debreu securities of Chapter 1 to distinguish between rational expectations with correct beliefs and more general rational learning when investors beliefs may be biased.

Let μ_t denote the risk-free rate over the period $[t, t + 1]$. In the empirical analysis we use the traditional return r_{t+1} and modified return x_{t+1} defined *in excess* of risk-free rate μ_t , i.e.,

$$r_{t+1} = \frac{p_{t+1} - p_t}{p_t} - \mu_t \quad (3.17)$$

and

$$x_{t+1} = \frac{p_{t+1} - p_t}{p_{t+1}} - \frac{\mu_t}{1 + \mu_t}.^{26} \quad (3.18)$$

The rationality restrictions of Chapter 1 that we are going to use in the tests can be summarized as follows.

Under rational expectations, the traditional return r_{t+1} of *any* security is unpredictable from investors' past information:

$$E[r_{t+1} | \mathcal{F}_t] = 0. \quad (3.19)$$

Under more general rationality of learning, the traditional return r_{t+1} and modified

²⁶In our empirical analysis, the correction to the martingale restrictions due to discounting is usually much smaller than other effects (μ_t is less than 0.0002 when the annual riskless rate is 5%).

return x_{t+1} of *winning* digital options must satisfy the restrictions:

$$E[r_{t+1}|\mathcal{F}_t, \bar{\theta}] > 0, \quad (3.20)$$

$$E[x_{t+1}|\mathcal{F}_t, \bar{\theta}] = 0. \quad (3.21)$$

3.4.1 Testing Rationality of Learning

We begin with the tests for rationality of learning when investors' beliefs may be biased. The tests are based on the restrictions in (3.20) and (3.21) that state that for learning to be rational, the modified return x_{t+1} of the digital options that matured in-the-money should be zero on average. In contrast, the average traditional return of the winning digitals must be positive.

Table 3.2 reports the results for the winning digital options $DC1 - DC5$. For each digital options, we first determine subperiods for which the digital option matured in-the-money (N in Table 3.2 denotes the number of such subperiods; for instance, in 40 out of 58 subperiods $DC3$ was a winner). Then, for each of these subperiods we compute the average daily traditional and modified returns and test whether the cross sectional averages $\bar{r}$ and $\bar{x}$ (of subperiod averages) are different from zero.²⁷

The results in Table 3.2 are consistent with the hypothesis of rational learning.²⁸ For all contracts $DC1 - DC5$, the average traditional return $\bar{r}$ is always positive and statistically significant at the 0.01 one-tail level. In contrast, the average modified return $\bar{x}$ is insignificantly different from zero at the 0.20 two-tail level. Note also that the average traditional return $\bar{r}$ monotonically increases as moneyness decreases (from $DC1$ to $DC5$) as should be expected: the lower initial prices (smaller initial m), the stronger the upward drift in prices of winners. Nevertheless, the average modified return $\bar{x}$ "cancels" the upward drift (or the selection bias) for all contracts $DC1 - DC5$. All t -statistics in Table 3.2 are truly impressive.

The restriction (3.21) not only claims that the average modified return is zero

²⁷The results for nonoverlapping subperiods can be safely assumed independent.

²⁸The returns in all tables are calculated per period (which, in this case, corresponds to one trading day); they are *not* annualized returns.

for winning digitals but also that the modified return x_{t+1} ought to be unpredictable from past information. We verify this for a popular in finance regression, whereby the return is projected onto the lagged price p_t using OLS. In Table 3.3 we report the results of OLS projections for the daily traditional and modified returns in the time-series analysis of pooled winning digital options.

Again, the results are pronounced. For all digitals $DC1 - DC5$, both the intercept and slope coefficient are several standard errors from zero for the traditional return r_{t+1} , but they become insignificant when the modified return is used. The positive intercept and negative slope coefficient in the regressions involving the traditional return is reminiscent of similar findings on prices of common stock on organized exchanges. The insignificance of the intercept and slope coefficients when using the modified return is comforting: since the regressor is a lagged, endogenous variable, the OLS estimators are biased, and one expects to reject often, at least in small samples.

Figure 3.5 presents the scatter plots of the projections of the traditional and modified returns for the digital options $DC3$, contracts that start at-the-money at the beginning of each subperiod. As can be seen from these plots, the results of the OLS regressions are robust. They are not due to a few outliers.

Figure 3.5 also reveals pronounced heteroscedasticity in the error term: its variance decreases with price level. This, of course, affects the statistical inference. One way to correct for heteroscedasticity is to run a GLS regression. For example, the variance of the error term can be specified as some function of the lagged price p_t and time to maturity $\tau = (T - t)$. Alternatively, the variance can be estimated in a two-step GLS regression. We tried several parametric models of heteroscedasticity and the results were not altered qualitatively. It turns out that even most general GLS models do not purge the regression error of all heteroscedasticity. That is why, in what follows, we report the results that are based on the *cross sectional* analysis for nonoverlapping subperiods. Those results are not affected by heteroscedasticity.

The martingale restriction (3.21) has many other interesting implications. Let $\varphi(\mathcal{F}_t)$ be a function of the investors' information. Then, the following consequence of

the restriction (3.21) must hold:

$$E[x_{t+1}\varphi(\mathcal{F}_t)|\mathcal{F}_t, \bar{\theta}] = 0. \quad (3.22)$$

In particular, for real γ ,

$$E[x_{t+1}p_t^\gamma|\mathcal{F}_t, \bar{\theta}] = 0. \quad (3.23)$$

One way to interpret results in (3.23) and (3.22) is the following. Imagine an investor whose portfolio consists of a number of same digital options DC . At the end of each period, the digital options DC are bought and or sold such that the portfolio value is equal to $\varphi(\mathcal{F}_t)$ dollars (i.e., the portfolio consists of $\varphi(\mathcal{F}_t)/p_t$ contracts).

Assume that function $\varphi(\mathcal{F}_t)$ is uniformly positive and consider a portfolio of the digital options DC which eventually matured in-the-money (i.e., we again introduce the same selection bias). The net profit of this portfolio will be positive on average:

$$E[x_{t+1}\varphi(\mathcal{F}_t)|\mathcal{F}_t, \bar{\theta}] = 0. \quad (3.24)$$

In contrast, the *modified net profit* $x_{t+1}\varphi(\mathcal{F}_t)$ for winning digital options will form a martingale difference sequence 3.23. And as before, the modified return x_{t+1} corrects the bias caused by selecting only “winners.” In other words, when the investment $\varphi(\mathcal{F}_t)$ “earns” the return x_{t+1} (instead of r_{t+1}), the winning digitals give rise to zero average net profit.

We test the restriction (3.23) for values of $\gamma = 1, 2, -1$ and -2 .²⁹ Note the following interpretation for γ . When $\gamma > 1$, the investor uses the “follow market” strategy: she sells the securities when their price goes down and buys them when price goes up. When $\gamma < 1$, the investor follows the “contrarian” strategy: she sells the securities when their price goes up and buys them when price goes down. Finally, the case $\gamma = 1$ corresponds to the “buy and hold” trading strategy for which the composition of the portfolio remains constant over time.

The results are presented in Table 3.4. The corresponding statistics for traditional

²⁹The case $\gamma = 0$ is reported in Table 3.2.

return r_{t+1} based on (3.24) are also included for completeness. The results strongly support the hypothesis of rationality of learning. The statistics for the traditional return are always positive and statistically significant at the 0.01 one-tail level (except for two occasions when they are significant at the 0.05 one-tail level). In contrast, all statistics for the modified return x_{t+1} (except for two occasions when $\gamma = -2$) are insignificantly different from zero at the 0.20 two-tail level.³⁰

It is worth mentioning that analysis of the trading strategies in (3.23) represent a rather severe test of the martingale property of the modified return x_{t+1} . Different trading strategies can be viewed as different rules to compute weighted average returns. Extreme weights amplify any model misspecifications. In particular, the trading strategies with negative γ put unacceptably large weights on observations with unusually small p_t (these observations often correspond to abnormally high traditional returns r_{t+1}). Even a small error in the estimation of the price p_t affects the average statistics substantially. However, without specific assumptions on the error distribution, the effect of this bias cannot be accounted for.

We now consider implication of the martingale restrictions (3.20) and (3.21) for returns over periods *larger* than one trading day (of course, the rationality restrictions must hold regardless of period length). Let $r_{t,t+l}$ denote the traditional return in excess of risk-free rate $\mu_{t,t+l}$ over the period $[t, t+l]$. The modified return $x_{t,t+l}$ is defined similarly to (3.18). The restrictions (3.20) and (3.21) then imply that

$$E[r_{t,t+l}|\mathcal{F}_t, \bar{\theta}] > 0, \quad (3.25)$$

$$E[x_{t,t+l}|\mathcal{F}_t, \bar{\theta}] = 0. \quad (3.26)$$

Table 3.5 reports the tests based on (3.25) and (3.26) for the cases $l = 2$ and $l = 4$. The results are similar to those in Table 3.2 and confirm the hypothesis of rational learning nicely.

So far all tests in this section have been based on prices of digital *calls*. However,

³⁰Ironically, exactly 2 out of 20 statistics in Table 3.4 have significant t -values which literally corresponds to the 10 percent significance level.

we can as well analyze returns of contracts *complimentary* to digital calls, namely, digital puts. A digital put $DP_t(X)$ with exercise price X and time to maturity T pays \$1 when $DC_t(X)$ pays zero and vice versa, i.e., $DP_T(X) = 1 - DC_T(X)$. Of course, when a digital call matures out-of-the-money, the corresponding digital put matures in-the-money and its modified return must satisfy the martingale restriction (3.21). We have implemented all the previous tests for winning digital puts. The results are very similar to those based on digital calls and not reported to conserve space.

In summary, the empirical results of this section surprisingly consistently support the hypothesis of rational learning. Of course, we have only tested a few of *infinitely* many implications of the result (3.21).³¹ Still, the variety of implemented tests and unusually clean empirical results for the modified return lend some credence to the rationality hypothesis.

3.4.2 Testing For Rational Expectations

We now turn to tests for rational expectations when investors' beliefs must coincide with the actual distribution of outcomes. This implies that assets traditional returns must be unpredictable, i.e., the historical average return is zero and projections of returns onto past information ought to produce only insignificant coefficients – a property also known as *market efficiency* (Fama (1970), (1991)).

Tests for rational expectations are based on the restriction (3.19) which must hold for the traditional return r_{t+1} (in excess of the risk-free rate) of *any* security (not only winning digitals). We first examine returns of the S&P 500 Index and notice that, over the period from 1991 to 1995, the average daily excess return on the Index is extremely large, annualized 8.9% (the average absolute return of 13.2% net the average risk-free rate of 4.3%). The corresponding return is even large for 1995 for which it is 24.2% (29.7% net 5.5%, respectively). And this is without taking into account dividend payments. Can the excess return of that magnitude be consistent

³¹One can never prove the statement that past information cannot be used to predict future returns since there are an infinite number of ways that the sequence of past prices can be used to forecast future prices. All that can be done is to test particular ways of combining past information to predict future prices.

with the hypothesis of rational expectations? Clearly not if investors are risk-neutral. Even if they are risk-averse, still, such a huge excess return cannot be adequately rationalized by the (time-varying) risk premia.³² It must reflect the fact that the investors' beliefs were biased (they were too pessimistic about the expected return of the Index).

We can also look at traditional excess returns of other securities. Table 3.6 reports the average returns $\bar{r}$ on different derivative contracts: digitals, puts, and calls with exercise prices $X1 - X5$ (for each contract, return $\bar{r}$ is computed as a cross sectional average of 58 subperiod averages).³³ Several of these statistics are significantly different from zero rejecting the hypothesis of rational expectations. Note that for puts with exercise prices $X1 - X5$ and for calls with exercise price $X5$ average return is negative and significant at the 0.02 two-tail level. This is particularly damaging for the traditional hypothesis of rational expectations because this implies a *negative* risk premium.

Additional evidence against rational expectations is presented in Table 3.7 which reports the results of the OLS projection of the daily traditional return r_{t+1} onto the closing price of the previous day p_t in the time-series analysis for all digital options (winners and losers). The intercept and/or slope coefficient are significant for contracts $DC1 - DC3$.

What is the source of the findings in Tables 3.6 and 3.7? Because we concluded in the previous section that prices of the S&P 500 Index Options reflect rational learning, it cannot be over-reaction or under-reaction (the explanation usually advanced for this phenomenon in the stock markets). It must be investors' beliefs. They were biased, in violation of one of the assumptions behind rational expectations.

³²The equity premium puzzle has been posed by Mehra and Prescott (1985) who have observed that *ex post* historical premium provided by stocks over bonds is much higher than can be justified by reasonable specifications of investors' risk aversion. As follows from the discussion in Mehra and Prescott (1988) (in response to Rietz (1988)), to explain this phenomena in the framework of the standard competitive theory, the required investors' risk-aversion parameter must be no less than 10, which is far larger value than used in other applied general equilibrium studies. Besides, there is not any historical support to justify especially abnormal risk premium for 1995 compared to other years.

³³In Table 3.6 prices of puts and calls for the artificial exercise prices $X1 - X5$ are calculated from the empirical put formulae estimated in Section 3.3.2.

3.4.3 Analysis Of Investors' Expectational Errors

To investigate in what sense the investors' expectations were biased, let us take a closer look at the results. In Table 3.6, the average traditional return $\bar{r}$ are positive for all digital options $DC1 - DC5$ (at different significance levels). This supports the previous observation that the investors were pessimistic about the expected return of the Index. The same conclusion is also consistent with the results for puts: their average return $\bar{r}$ is significantly negative for all exercise prices $X1 - X5$. However, the negative average returns of calls (significant for exercise prices $X5$) may seem counterintuitive. After all, calls must have matured in the money more often than investors expected (the conclusions based on returns of digital options and S&P 500 Index).

The explanation to this surprising phenomenon lies in the volatility. The average historical volatility of the S&P 500 Index (annualized with a 252-day year) was 10.32% over the five-year period and 7.78% for 1995. The Index average volatility *implied* in options prices was 13.33% and 11.43%, respectively. In other words, investors considerably overestimated the volatility of the index over the studied period, which contributed to overpricing calls and puts. Even though the calls matured in the money more often than expected, their realized payouts turned out to be *lower* than what investors anticipated.

Another way to characterize investors' biases is to note that investors on average overestimated the time value of calls and puts; they underestimated the intrinsic value of calls and overestimated the intrinsic value of puts. Because of this, out-of-the-money calls (exercise prices $X4$ and $X5$) with zero intrinsic values were significantly overpriced. Overpricing decreases as moneyness increases, and far-in-the-money calls were underpriced (i.e., they had positive average returns; not reported). For put options, the two biases *added* to overpricing. Still, out-of-the-money puts (exercise prices $X1$ and $X2$) were overpriced more significantly compared to in-the-money puts (exercise prices $X4$ and $X5$).³⁴

³⁴Also, average returns become negative for far-out-of-the-money digitals (i.e., they were overpriced; not reported) again supporting the conclusion that the historical volatility of the index

It is important to keep in mind the limitations of our analysis. In interpreting the empirical finding we have relied on several assumptions. Some of them are standard, like perfect markets (no transaction costs, not tax considerations, unlimited risk-free borrowing and lending, etc.), informational efficiency (absence of riskless arbitrage opportunities), and existence of the representative agent.

The critical assumption that we have used is that investors are risk-neutral. In the above analysis of investors' biases we, in a sense, compare the *actual* distribution of the Index returns with the distribution returns *implied* in options prices. The latter reflects information about *both* the implied Index price dynamics *and* investors' preferences. If investors are *not* risk-neutral, then one wonders whether some of the previous conclusions may be, at least in part, due to investors' risk aversion. For example, one can argue that the too high implied volatility (compared to the historical volatility) may reflect the insurance feature of options (i.e., options sellers require extra premium to accept higher risk of adverse price movements of an underlying asset and option buyers are willing to pay the premium to hedge against this risk). While it may be true for puts and calls, it is very unlikely that the conclusions based on the analysis of returns of digitals may be explained in term of risk.

First, there are theoretical reasons which suggest that the rationality restrictions (3.19) and (3.21) on prices of digitals must obtain under very general risk preferences of investors. Formal characterization of investors' preferences for which the theory still holds is beyond the scope of this study, but the intuition is simple. Loosely stated, the riskiness of a digital call is similar to that of the complimentary digital put and their combination is risk-free (by the no-arbitrage argument). This fact, along with binary payoff structure, guarantees that prices of digitals satisfy the martingale restrictions (3.19) and (3.21) for a wide class of investors' preferences.³⁵

The second argument is empirical. The results for the modified returns of winning returns was considerably lower than the volatility implied by the options prices.

³⁵This property of Arrow-Debreu securities has been recognized by experimental economists who often use the Binary Lottery Payoff procedure to induce risk-neutral preferences in the laboratory. In this two-stage procedure, payoffs to a primary research problem are denominated in terms of "tickets" or "points," which determine the probability of winning a prize in a binary lottery that is played in the second stage. See, for example, Kagel and Roth (1995).

digitals in Tables 3.2-3.5 conform to the theoretical model nicely which suggests that if there is a risk premium in prices of digitals, it is much smaller than the estimation errors. This means that the risk premium should be also negligible for the traditional returns of digitals. Therefore, the analysis based on prices of digitals is robust to investors' preferences which is, of course, extremely attractive.

Another important concern in interpreting the empirical results is the power of the tests. The analysis in the previous section is based on *estimated* prices of securities and, thus, we, essentially, test a *joint* hypothesis of rationality of asset pricing *and* correctness of the implemented estimation procedure. Since we use no special assumptions on the underlying stochastic process, we cannot derive statistical properties of the estimation errors and, consequently, cannot formally state the power of the tests. We can address this potential criticism in three ways.

First, let us again take a look at results for winning digitals in Tables 3.2-3.5. The statistics for the traditional return are always very significant (several standard errors from zero) while they are insignificant for the modified return. This gives some indication on the power of the tests. Indeed, the difference between the traditional and modified returns is relatively small; it is of second order of magnitude

$$r_{t+1} - x_{t+1} = \frac{(p_{t+1} - p_t)^2}{p_{t+1}p_t} - \frac{\mu_t^2}{1 + \mu_t}.$$

Thus, it is very comforting to know that this small correction leads to impressive difference in *t*-statistics for the traditional and modified returns.

Second, the reported empirical findings seem to be robust to the choice of estimation procedure. We have tried several other estimation procedures (including different parametric and nonparametric techniques), but the main conclusions have not changed.³⁶

Finally, we interpret the unusually clean empirical results for the modified return

³⁶We realize that results based on a nonstandard estimation procedure (especially if it uses a new parametric specification) may be less convincing than results based on more traditional approaches, but since the procedure described in this chapter seems to consistently outperform other methods in fitting empirical data, we report the result that are based on it.

of winning digitals as the evidence that the estimation procedure works well enough, at least for the studied dataset.

3.5 Conclusion

This chapter implements a new approach to test rationality of asset pricing. Based on the theory of Chapter 1, it distinguishes between traditional rational expectations and more general rational learning. Several empirical tests are presented which demonstrate that there is no predictability in the modified returns of prices of winning digital options estimated from the S&P 500 Index Options. Since these modified returns are unpredictable if investors correctly apply the rules of conditional probability, we conclude that prices of the S&P 500 Index Options are in accordance with the hypothesis of rational learning.

In contrast, we find evidence against rational expectations. The evidence is in the form of significant average traditional returns of the Index, digitals, puts and calls and a negative correlation between the returns and the level of the previous closing price for digitals. Since prices of the S&P 500 Index Options reflect rational learning but violate rational expectations, we conclude that investors' beliefs must have been biased. We argue that, over the 5-year period, the investors repeatedly underestimated the expected return of the S&P 500 Index but overestimated its volatility.

The main message of this chapter is that there is a simple and parsimonious model that surprisingly accurately describes actual behavior of the S&P 500 Index Options prices. In this model, financial markets are populated with investors who may have biased beliefs but who are otherwise rational decision makers that efficiently process new information. The model is a generalization of the concept of rational expectations, but it is not *too* general in the sense that "it can explain anything." On the contrary, the model of rational investors with expectational errors suggests precise testable restrictions which *turn out* to be satisfied in the empirical data. Moreover, the model explains the data without having to appeal to sophisticated and sometimes

controversial theories of time-varying risk premia. In particular, the empirical findings indicate that, if there is a risk premium for digital options written on the S&P 500 Index, it must be negligible (smaller than the estimation errors), which is much more in line with our theoretical understanding of risk premia.

It deserves emphasis that reported results do not imply that there is money to be made. The beliefs of the market of the S&P 500 Index Options are as good as any others. An outsider may have different beliefs, but whether this would lead her to “make money” depends on the circumstances. She would only if her priors *happened* to be more accurate (this is a chance event). Still, her trading would amount to pure speculation.³⁷

Based on the evidence reported here, investors in the S&P 500 Index Options certainly do not leave “money on the table” by making mistakes against the rules of conditional probability. At least for the short-term horizons considered in this chapter, investors’ reactions to new information appear to be consistent with the rules of conditional probability: investors do not *over-react* or *under-react*. The newcomer may want to trade if her priors differ from those of the market. But disagreement is not sufficient. Because of the well-known result in Aumann (1976), she must be sure that her priors and those of the market are *inconsistent* (as defined by Harsanyi (1967)).

And finally, to implement the tests for the new rationality restrictions, we have developed the novel procedure to recover prices of digitals and risk-neutral densities implicit in prices of traded puts and calls. The procedure is very flexible, it estimates instantaneous implied risk-neutral densities using only the most recent option prices only and, thus, does not depend on stationarity of an underlying asset’s price dynamics. The clean empirical results indicate that the estimation procedure captures most salient features of the option data extremely well. This means that the procedure may be useful in many other contexts as well. For example, the estimation procedure

³⁷It is worth pointing out that investors’ biases may (and will) change over time. The estimated historical expectation errors cannot be used to predict the future expectation errors and, thus, historical discrepancies between the implied and factual return distributions do not necessarily indicate profit opportunities.

suggests how prices of traded puts and calls can be used to determine arbitrage-free prices of complex or illiquid contingent claims, for example, exotic options. The procedure also allows one to analyze the implied underlying asset's price dynamics and to calibrate parameters of theoretical price dynamics models.

3.6 Appendix

3.6.1 Constrained Convex Least Squares Regression

Dykstra (1983) has proposed a simple iterative technique for a least squares regression subject to the restriction that the solution must satisfy certain side conditions. Specifically, the following problem is considered. Let vectors $g, w, y \in R^n$. Consider a minimization problem:

$$\text{Minimize}_{y \in \bigcap_1^r K^i} \sum_{i=1}^n w^i (g^i - y^i)^2 \quad (3.27)$$

where $K^1, \dots, K^r$ are closed convex cones. It is assumed that the following problem

$$\text{Minimize}_{y \in K^i} \sum_{i=1}^n w^i (f^i - y^i)^2, \quad (3.28)$$

can be easily solved for any $f = (f^1, \dots, f^n) \in R^n$ and i . To find the solution to the more involved problem (3.27), the procedure iteratively solves simpler problems (3.28).

The constrained least squares problem (3.7)-(3.10) fits nicely in this framework. Indeed, let $w' = (1, 1, \dots, 1)$, $P^i = P_t(X^i)$, $X^1 < X^2, \dots, < X^n$. Define K^i as follows:

$$\begin{aligned} K^1 : \quad & Y^2 - Y^1 \geq 0, \\ K^i : \quad & \frac{Y^{i+1} - Y^i}{X^{i+1} - X^i} \geq \frac{Y^i - Y^{i-1}}{X^i - X^{i-1}}, \quad i = 2, \dots, n-1, \\ K^n : \quad & \frac{Y^n - Y^{n-1}}{X^n - X^{n-1}} \leq e^{-\mu_t, T(T-t)}. \end{aligned}$$

The constrained least squares problem (3.7)-(3.10) can be then written as

$$\text{Minimize}_{\{Y^i\}_{i=1}^n \in \bigcap_1^n K^i} \sum_{i=1}^n (P^i - Y^i)^2. \quad (3.29)$$

The corresponding problems (3.28) for K^i ($i = 1, \dots, n$) are simple, nevertheless, we cannot use the algorithm directly since the restriction K^n is *not* a closed convex cone. Fortunately, there is a way to resolve this complication.

We first observe that the restriction K^1 is not binding when put prices $\{P^i\}_{i=1}^n$ satisfy the condition

$$P^1 \leq P^i, \quad i = 1, \dots, n. \quad (3.30)$$

In this case, it can be shown that the solution to (3.29) (with excluded restriction K^1) satisfies the same property:

$$Y^1 \leq Y^i, \quad i = 1, \dots, n.$$

In the studied dataset, the condition (3.30) has been met for all 1160 trading days³⁸ and, thus, restriction K^1 in (3.29) can be omitted. After that, we define a linear transformation

$$Z^i = Y^i - X^i e^{-\mu_t, T(T-t)},$$

and rewrite the CCR problem (3.29) as follows:

$$\underset{\{Z^i\}_{i=1}^n \in \bigcap_2^n \widetilde{K}^i}{\text{Minimize}} \quad \sum_1^n (P^i - Z^i - X^i e^{-\mu_t, T(T-t)})^2, \quad (3.31)$$

with the restrictions $\widetilde{K}^i$ expressed in terms of Z^i :

$$\begin{aligned} \widetilde{K}^i : \quad & \frac{Z^{i+1} - Z^i}{X^{i+1} - X^i} \geq \frac{Z^i - Z^{i-1}}{X^i - X^{i-1}}, \quad i = 2, \dots, n-1, \\ \widetilde{K}^n : \quad & \frac{Z^n - Z^{n-1}}{X^n - X^{n-1}} \leq 0. \end{aligned}$$

The restrictions $\widetilde{K}^i$ are closed convex cones and the solution to (3.31) can be found according to the general algorithm. The projections onto $\widetilde{K}^i$ are very simple (they only involve corrections for three points, Z^{i-1} , Z^i and Z^{i+1} for $\widetilde{K}^i$, $i = 2, \dots, n-1$, and two points, Z^{n-1} and Z^n for $\widetilde{K}^n$). The average value of n for empirical put prices is about 21 and the convergence to the solution is quite rapid.

³⁸If it were not we could have enforced it, for example, by adding an “artificial” put $P^0 = P_t(X^0)$ with the exercise price $X^0 < X^1$ and the lowest possible price, $1/16$.

3.6.2 Properties of Quadratic Specification

In this appendix we discuss several theoretical and computational properties of the quadratic specification defined in equation (3.15).

Note that, in view of put-call parity (3.4), equation (3.15) can be compactly restated in the following form

$$P_t^Q(X)C_t^Q(X) = \alpha_t F_t G_t. \quad (3.32)$$

The put formula $P_t^Q(X)$ satisfies two additional properties (also shared by the Black-Scholes formula $P_t^{BS}(X)$ and most other parametric option formulae).

First, $P_t^Q(X)$ is homogeneous of degree one in both F_t and G_t . This property holds for options written on an asset which follows a *proportional* risk-neutralized stochastic process (i.e., when it is assumed that the risk-neutral return distribution is independent of price).³⁹

Second, let X^* denote the implied future price, i.e., the price for which $F_t = G_t(X^*)$ (or, alternatively, for which $P_t(X^*) = C_t(X^*)$):

$$X^* = S_t e^{(\mu_{t,T} - \delta_{t,T})(T-t)}.$$

Then, for any positive constant k ,

$$P_t^Q(kX^*) = kC_t^Q(X^*/k).$$

This property holds for assets for which the risk-neutralized stochastic process is proportional and *symmetric* in log-returns.

Note that, intuitively, the parameter α_t in the quadratic specification is closely related to the variance σ_t^2 in the Black-Scholes specification. In general, α_t monotonically increases as σ_t^2 increases (the range $[0, 1]$ of α_t corresponds to the range $[0, \infty]$

³⁹See Theorem 9 in Merton (1973b).

of σ_t^2). The two specifications coincide in the limits, i.e., when $\alpha_t = \sigma_t^2 = 0$,

$$P_t^Q(X) = P_t^{BS}(X) = \max(G_t - F_t, 0),$$

and when $\alpha_t = 1$, $\sigma_t^2 = \infty$

$$P_t^Q(X) = P_t^{BS}(X) = G_t.$$

From equation (3.3) (or (3.5)) we obtain the price of the corresponding digital option, $DC_t^Q(X)$:

$$DC_t^Q(X) = \frac{1}{2} e^{-\mu_{t,T}(T-t)} \left(1 - \frac{(G_t - F_t) + 2\alpha_t F_t}{\sqrt{(G_t - F_t)^2 + 4\alpha_t F_t G_t}} \right).$$

Let $\psi_t^Q(X) = \psi^Q(S_t, \mu_{t,T}, \delta_{t,T}, t; X, T)$ denote the risk-neutral density (RND) implied by the quadratic specification (3.15), i.e., for any exercise price X , $\psi_t^Q(X)dX$ is equal to risk-neutral probability that the asset maturity price S_T falls between X and $X+dX$. The *cumulative* risk-neutral density (CRND), $\Psi_t^Q(X) = \Psi^Q(S_t, \mu_{t,T}, \delta_{t,T}, t; X, T)$, is equal to the risk-neutral probability that the asset maturity price S_T does not exceed X .

Accordingly,

$$\psi_t^Q(X) = \frac{\partial^2 P_t^Q(X)}{\partial G_t^2} = \frac{4F_t^2 \alpha_t (1 - \alpha_t)}{\sqrt{(G_t - F_t)^2 + 4\alpha_t F_t G_t}^3}$$

and

$$\Psi_t^Q(X) = \frac{\partial P_t^Q(X)}{\partial G_t} = \frac{1}{2} \left(1 + \frac{G_t - F_t + 2\alpha_t F_t}{\sqrt{(G_t - F_t)^2 + 4\alpha_t F_t G_t}} \right).$$

Finally, let us discuss a useful trick that significantly simplifies solving the non-linear least squares problem (3.13) for the quadratic specification. Recall that in the third step of the estimation procedure in Section 3.3.2 for each trading day t and

maturity date T , we needed to solve the following problem:

$$\text{Minimize}_{\alpha, F} \sum_{i=1}^n (\hat{P}^i - P^Q(X^i; \alpha, F))^2, \quad (3.33)$$

where $\{\hat{P}^i\}_{i=1}^n$ is a collection of empirical put prices corresponding to exercise prices $\{X^i\}_{i=1}^n$. The parametric function $P^Q(X^i; \alpha, F) \equiv P^Q(X^i)$ is such that

$$P^Q(X^i)(F - G^i + P^Q(X^i)) = \alpha F G^i, \quad (3.34)$$

where $G^i = G(X^i) = X^i e^{-\mu t, T(T-t)}$.

Since $P^Q(X^i; \alpha, F)$ is a *nonlinear* function of parameters α and F , we use the multivariate Newton-Raphson method to search for the optimal solution (α^*, F^*) of (3.33). The ability of the iterative Newton-Raphson method to find the functional global minimum (as well as its computational efficiency) critically depends on availability of a “good” initial approximation for the solution (α^*, F^*) . In order to find initial values (α, F) close enough to (α^*, F^*) , we consider the following *weighted* least squares problem:

$$\text{Minimize}_{\alpha, F} \sum_{i=1}^n (\hat{P}^i - P^Q(X^i))^2 (\hat{P}^i + P^Q(X^i) + F - G^i)^2. \quad (3.35)$$

The problem (3.35) differs from the problem (3.33) because *variable* weights $\{\hat{P}^i + P^Q(X^i) + F - G^i\}_{i=1}^n$ are used instead of constant ones. Otherwise, the two problems are similar, and it can be shown that, under weak assumptions on $\{\hat{P}^i\}_{i=1}^n$, the solution (α', F') of (3.35) is “very close” to (α^*, F^*) (this statement can be made precise but, intuitively, it is clear since the weights in the minimization problem (3.35) do not vary significantly). Now, the problem (3.35) may *appear* more difficult than the problem (3.33), but, in fact, it is much easier to solve (3.35) than (3.33). Indeed, using (3.34) we obtain that:

$$(\hat{P}^i - P^Q(X^i))(\hat{P}^i + P^Q(X^i) + F - G^i) = \hat{P}^i(\hat{P}^i - G^i) + F\hat{P}^i - \alpha F G^i. \quad (3.36)$$

This means that we only need to project $\hat{P}^i(\hat{P}^i - G^i)$ onto $(-\hat{P}^i)$ and G^i in an OLS regression to find the coefficients F' and $\alpha'F'$ (which amounts to solving the corresponding 2×2 system of linear algebraic equations). After that, we use the vector (α', F') as an initial approximation for (α^*, F^*) in the Newton-Raphson search. Since this approximation is very tight, the method always finds (α^*, F^*) (as opposed to some local minimum of the functional (3.33)) and requires only a few iterations.

3.6.3 One Day in the Life of the Dataset

The purpose of this appendix is to illustrate the procedure used to estimate prices of digitals. Here we report intermediate results of different steps of the procedure for a randomly chosen trading day in the dataset. Specifically, consider options that matured on April 15, 94 and that were trading on March 15, 94, corresponding to time to maturity, $\tau = 22$. At the close of March 15, 94, the S&P 500 Index was 467.01 and there were 14 puts and 14 calls with positive volume and with exercise prices from 400 to 500. Not all of these options were equally liquid. Therefore, in the first step of the estimation procedure, we use put-call parity to obtain reliable put prices.

This step is illustrated in Figure 3.6. We estimate the put-call relationship for this trading day ($F_t = 465.57$) from liquid puts and calls with exercise prices 460, 465, 470, and 475 (i.e., the four closest to at-the-money pairs). Now we know two put prices: *true*, $P_t(X)$, and *imputed* from the corresponding call, $\bar{P}_t(X)$. After that, we replace less liquid puts (exercise prices above 475) with $\bar{P}_t(X)$. Puts with exercise prices from 460 to 475 are replaced by $1/2(P_t(X) + \bar{P}_t(X))$, and more liquid puts (exercise prices below 460) are not changed. Resulting prices are represented by the solid line in Figure 3.6.

In the second step, we run the constrained convex regression to ensure that put prices are increasing and convex. This step is illustrated in Figure 3.7. Here dots and the solid line represent the input and the output of the CCR, respectively. Note that, for most exercise prices, the CCR does not change corresponding puts. It does,

however, “correct” the nonconvexity around exercise prices 445-460. This example is typical, in the sense that, for most trading days, only a few prices (if any) are needed to be changed to satisfy the shape restrictions. Nevertheless, the CCR is a very useful procedure to deal with real-world imperfections because it allows one to filter rare but sometimes significant nonconvexities/outliers locally, without affecting the rest of the dataset.

The third step of the estimation procedure, we fit discrete put prices by smooth continuous functions. This step is implemented for the quadratic specification (Figure 3.8) and for the Black-Scholes specification (Figure 3.9). The continuous approximations are shown by the solid lines. Inspection of the two graphs suggest that the quadratic specification fits the data significantly better than the Black-Scholes specification does. In particular, standard error is 0.17 for the former and 0.60 for the latter. Figure 3.9 also illustrates the famous volatility smile: the Black-Scholes approximation underprices out-of-the-money and in-the-money options but overprices at-the-money options. In contrast, Figure 3.8 does not reveal any special pricing biases for the quadratic approximation. Again, this example is typical: For most trading days in the studied dataset, the quadratic approximation consistently outperforms the Black-Scholes approximation. Finally, Figure 3.10 plots prices of digitals computed using different approximations. It shows that the Black-Scholes approximation underprices in-the-money digitals ($m > 1$) and overprices out-of-the-money digitals ($m < 1$). For comparison, we also report prices of digitals computed from the finite difference equation (3.12).

Table 3.1: Distribution of Average Daily Trading Volumes, Closing Prices, and Maturities for S&P 500 Call and Put Options: Sample Period 01/91-12/95

Moneyness (m)	<u>Call</u>				<u>Put</u>			
	N	Volume	Price	τ	N	Volume	Price	τ
$1.12 < m$	2884	313	89.1	48	11520	346	1.5	75
$1.08 < m \leq 1.12$	2110	362	43.9	40	8119	503	2.7	60
$1.04 < m \leq 1.08$	4794	464	28.2	37	13408	783	3.7	52
$1.00 < m \leq 1.04$	11434	966	14.5	40	16736	1368	6.6	53
$0.96 < m \leq 1.00$	14806	1226	5.9	45	13507	622	13.3	51
$0.92 < m \leq 0.96$	7886	658	2.6	58	4189	184	27.0	70
$0.88 < m \leq 0.92$	1840	422	2.4	93	1166	123	44.0	95
$m \leq 0.88$	369	246	1.1	108	437	98	73.6	93
All	46123	849	16.5	48	69082	736	8.3	59

Notes: Averages are taken across observation with positive trading volume of corresponding options, N denotes the number of these observations. Time to maturity, $\tau = T - t$, is measured in trading days, moneyness $m = S/X$.

Table 3.2: Analysis of Returns of Winning Digital Options

	<i>DC1</i>	<i>DC2</i>	<i>DC3</i>	<i>DC4</i>	<i>DC5</i>
<i>N</i>	54	47	40	33	24
$\bar{r}$	0.0077** (4.42)	0.0129** (5.62)	0.0248** (6.87)	0.0378** (6.60)	0.0515** (5.38)
$\bar{x}$	-0.0030 (-0.87)	-0.0024 (-0.54)	-0.0018 (-0.32)	-0.0029 (-0.35)	-0.0052 (-0.43)

Notes: *N* denotes the number of subperiods for which the corresponding digital option matured in-the-money. $\bar{r}$ denotes the average across all subperiods of the time series average daily traditional return. $\bar{x}$ denotes the average across all subperiods of the time series average daily modified return. The numbers in parentheses are *t*-statistics with * and ** indicating that the *t*-statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

Table 3.3: **OLS Regressions of Traditional and Modified Returns of Winning Digital Options on Beginning-of-period Prices**, $r_{t+1} = \alpha + \beta p_t + \varepsilon_t$ and $x_{t+1} = \alpha + \beta p_t + \varepsilon_t$

	<i>DC1</i>	<i>DC2</i>	<i>DC3</i>	<i>DC4</i>	<i>DC5</i>
<i>N</i>	1026	893	760	627	456
r_{t+1} : Intercept	0.1366** (7.00)	0.1227** (6.11)	0.1598** (6.79)	0.1397** (5.61)	0.1282** (4.50)
Slope	-0.1539** (-6.69)	-0.1405** (-5.58)	-0.1943** (-5.93)	-0.1766** (-4.36)	-0.1686** (-3.01)
x_{t+1} : Intercept	0.0124 (0.53)	0.0126 (0.54)	0.0299 (1.17)	0.0106 (0.44)	-0.0017 (-0.07)
Slope	-0.0184 (-0.67)	-0.0191 (-0.65)	-0.0456 (-1.29)	-0.0234 (-0.60)	-0.0076 (-0.16)

Notes: *N* denotes sample size. r_{t+1} denotes traditional return, x_{t+1} denotes modified return. The numbers in parentheses are *t*-statistics with * and ** indicating that the *t*-statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

Table 3.4: **Analysis of Profitability of Different Investment Strategies for Winning Digital Options, $E[r_{t+1}p_t^\gamma|\mathcal{F}_t, \bar{\theta}]$ and $E[x_{t+1}p_t^\gamma|\mathcal{F}_t, \bar{\theta}]$ for $\gamma = 1, 2, -1, -2$**

	<i>DC1</i>	<i>DC2</i>	<i>DC3</i>	<i>DC4</i>	<i>DC5</i>
<i>N</i>	54	47	40	33	24
Average $r_{t+1}p_t$	0.0036** (2.79)	0.0065** (3.81)	0.0108** (5.26)	0.0138** (5.23)	0.0149** (4.37)
Average $r_{t+1}p_t^2$	0.0020* (1.76)	0.0037** (2.80)	0.0054** (3.52)	0.0060** (3.50)	0.0058** (3.19)
Average $r_{t+1}\frac{1}{p_t}$	0.0185** (3.55)	0.0301** (5.87)	0.0619** (5.64)	0.1192** (5.51)	0.2126** (3.79)
Average $r_{t+1}\frac{1}{p_t^2}$	0.0549* (2.34)	0.0911** (3.05)	0.1802** (3.84)	0.4480** (4.31)	1.0722** (2.65)
Average $x_{t+1}p_t$	-0.0026 (-1.11)	-0.0022 (-0.75)	-0.0023 (-0.64)	-0.0021 (-0.51)	-0.0015 (-0.33)
Average $x_{t+1}p_t^2$	-0.1156 (-1.23)	-0.0799 (-0.83)	-0.0741 (-0.77)	-0.0014 (-0.58)	-0.0004 (-0.18)
Average $x_{t+1}\frac{1}{p_t}$	-0.0003 (-0.08)	0.0005 (0.09)	0.0066 (0.67)	0.0054 (0.25)	-0.0007 (-0.01)
Average $x_{t+1}\frac{1}{p_t^2}$	0.0131 (1.23)	0.0213* (2.16)	0.0437* (1.69)	0.0595 (0.72)	0.0215 (0.07)

Notes: *N* denotes the number of subperiods for which the corresponding digital option matured in-the-money. Averages are taken across all subperiods for time series average statistics. r_{t+1} denotes traditional return, x_{t+1} denotes modified return and p_t denotes price of digital options. The numbers in parentheses are *t*-statistics with * and ** indicating that the *t*-statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

Table 3.5: Two-day and Four-Day Returns of Winning Digital Options

	<i>DC1</i>	<i>DC2</i>	<i>DC3</i>	<i>DC4</i>	<i>DC5</i>
<i>N</i>	54	47	40	33	24
$\underline{l = 2}:$ $\bar{r}$	0.0162** (4.86)	0.0238** (4.58)	0.0418** (5.73)	0.0664** (5.22)	0.1033** (5.15)
$\bar{x}$	-0.0030 (-0.52)	0.0002 (0.03)	0.0017 (0.17)	-0.0027 (-0.16)	0.0005 (0.02)
$\underline{l = 4}:$ $\bar{r}$	0.0277** (3.99)	0.0567** (5.49)	0.1028** (7.79)	0.1561** (7.12)	0.2333** (5.43)
$\bar{x}$	-0.0067 (-0.52)	0.0082 (0.52)	0.0198 (1.09)	0.0180 (0.59)	0.0191 (0.42)

Notes: *N* denotes sample size for which the corresponding digital option matured in-the-money. $\bar{r}$ denotes the average across all subperiods of the time series average two-day or four-day traditional return. $\bar{x}$ denotes the average across all subperiods of the time series average two- or four-day modified return. Returns $\bar{r}$ and $\bar{x}$ are computed using every second and every fourth prices of the digital options (10 and 5 observations in each subperiod). The numbers in parentheses are *t*-statistics with * and ** indicating that the *t*-statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

Table 3.6: Analysis of Traditional Returns of Digitals, Puts, and Calls

	$X1$	$X2$	$X3$	$X4$	$X5$
Digitals, $\bar{r}$	0.0060** (2.92)	0.0071* (2.02)	0.0095* (1.83)	0.0123* (1.72)	0.0104 (1.20)
Puts, $\bar{r}$	-0.0356** (-4.08)	-0.0286** (-3.13)	-0.0233** (-2.63)	-0.0199** (-2.48)	-0.0169** (-2.43)
Calls, $\bar{r}$	-0.0012 (-0.24)	-0.0037 (-0.64)	-0.0078 (-1.18)	-0.0147* (-2.06)	-0.0247** (-3.49)

Notes: The number of subperiods is 58. $\bar{r}$ denotes the average across all subperiods of the time series average daily traditional return. The numbers in parentheses are t -statistics with * and ** indicating that the t -statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

Table 3.7: OLS Regressions of Traditional Returns of Digital Options on Beginning-of-period Prices, $r_{t+1} = \alpha + \beta p_t + \varepsilon_t$

	$DC1$	$DC2$	$DC3$	$DC4$	$DC5$
Intercept	0.0845** (5.00)	0.0602** (3.50)	0.0426** (2.53)	0.0300* (1.82)	0.0135 (0.88)
Slope	-0.0952** (-4.73)	-0.0720** (-3.19)	-0.0538* (-2.10)	-0.0375 (-1.20)	-0.0091 (-0.24)

Notes: The sample size is 1102 (=58*19). $\bar{r}$ denotes the average across all subperiods of the time series average daily traditional return. The numbers in parentheses are t -statistics with * and ** indicating that the t -statistic is significant at the 0.10 and 0.02 two-tail levels, respectively.

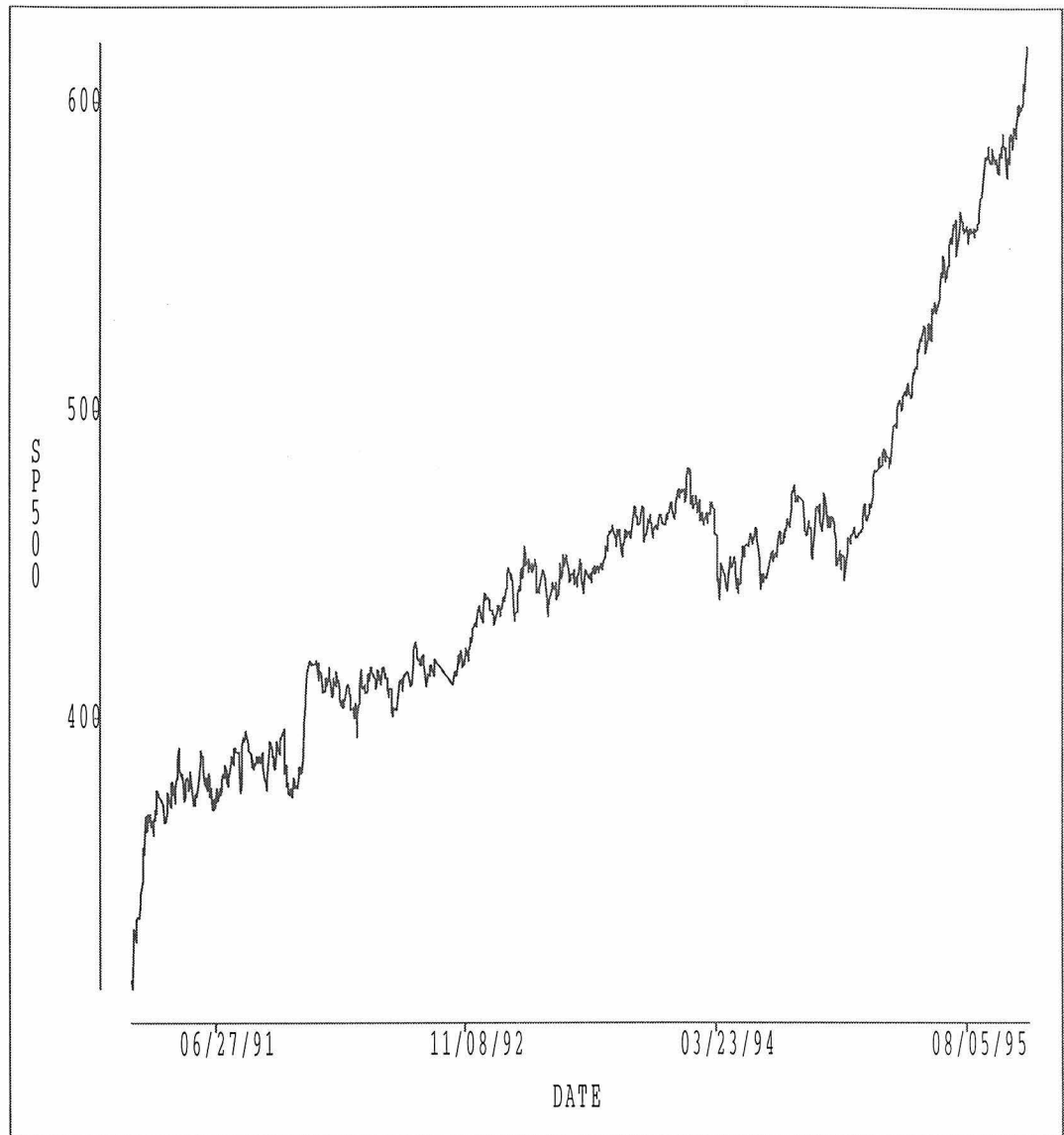


Figure 3.1: Daily closing prices of the S&P 500 index for the period from January 1991 to December 1995. The sample daily mean of continuously-compounded returns (annualized with a 365-day year) was 13.20% over the five-year period and 29.67% for 1995 year. The sample standard deviation of returns (annualized with a 252-day year) was 10.32% over the five-year period and 7.78% for 1995 year.

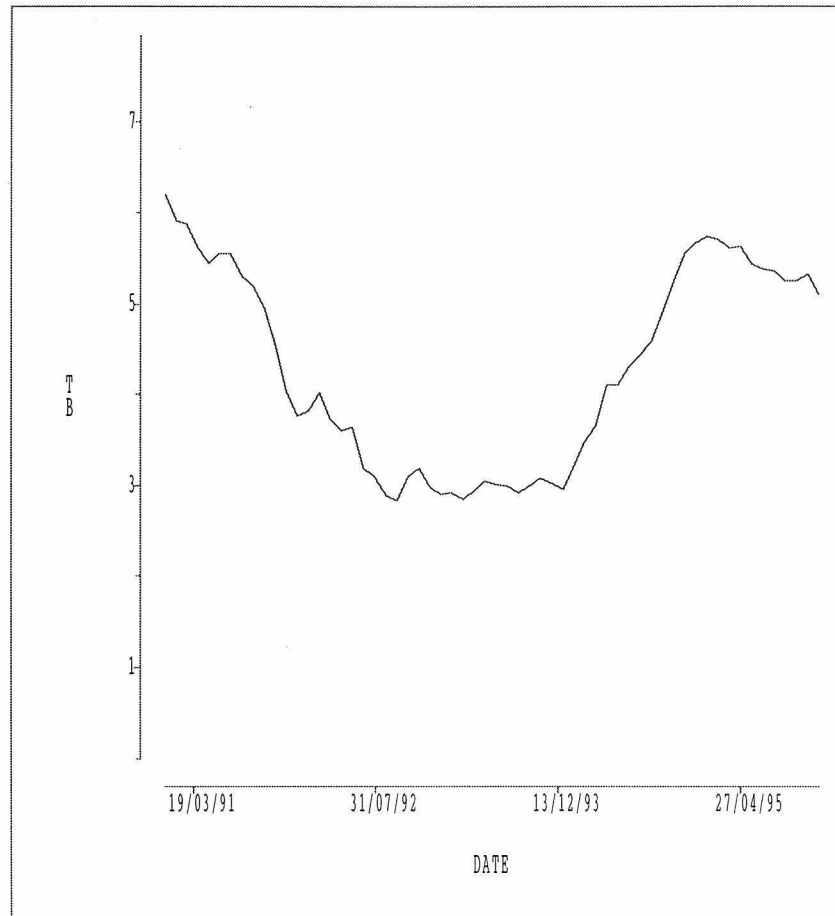


Figure 3.2: Risk-free rate μ approximated by the yield of the 3-month Treasury bill for the period from January 1991 to December 1995. The sample daily mean of risk-free rates (annualized with a 365-day year) was 4.31% over the five-year period and 5.49% for 1995 year.

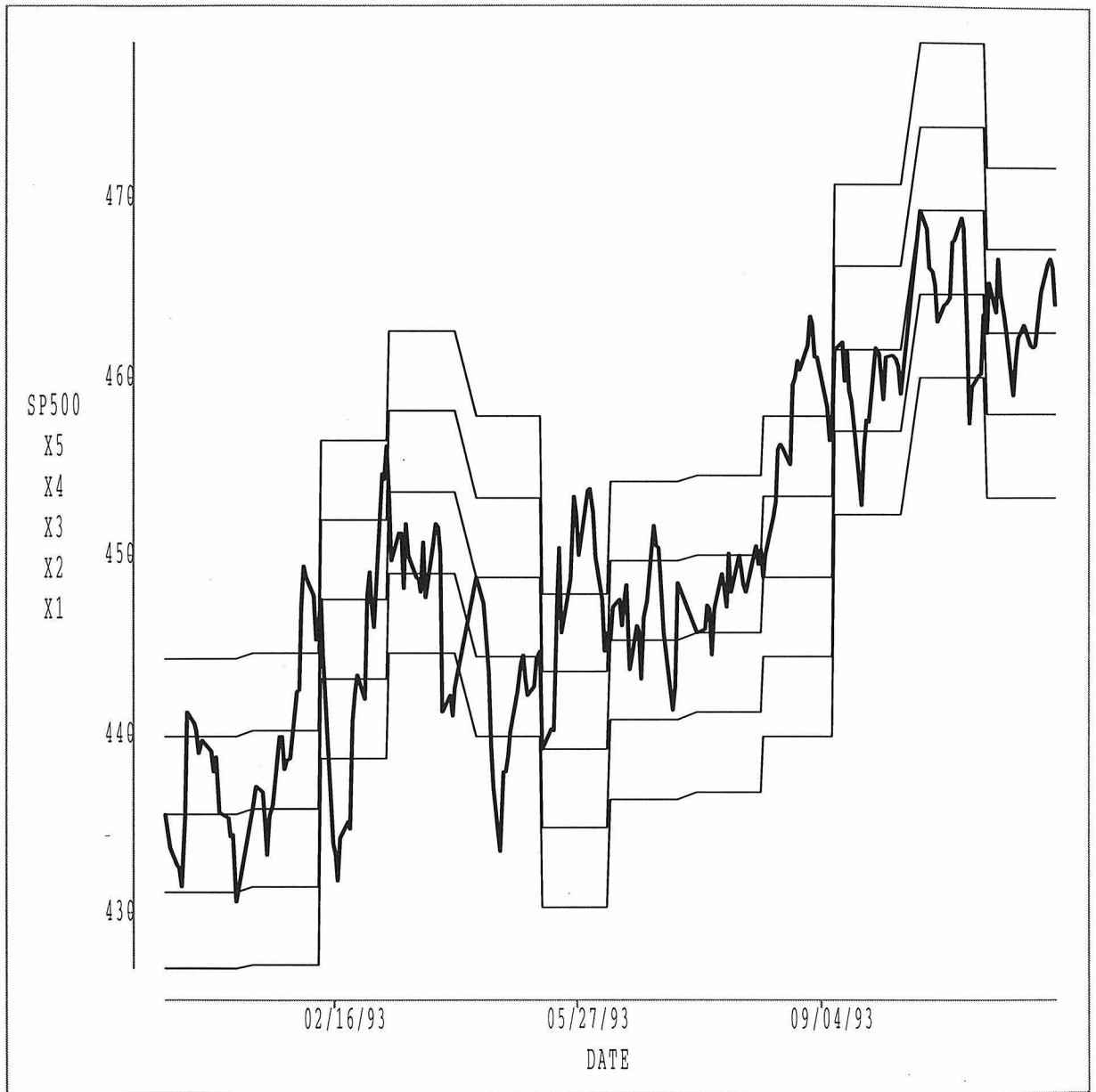


Figure 3.3: Daily closing prices of the S&P 500 index (thick line) and the exercise prices of the constructed digital options which expired in 1993 (12 months). At the beginning of each subperiod, digital options have exercise prices $X1$ - $X5$ such that $X1$ and $X2$ are 2%- and 1%-in-the-money, $X3$ is at-the-money, $X4$ and $X5$ are 1%- and 2%-out-of-the-money. See Section 3.3.3 for details.

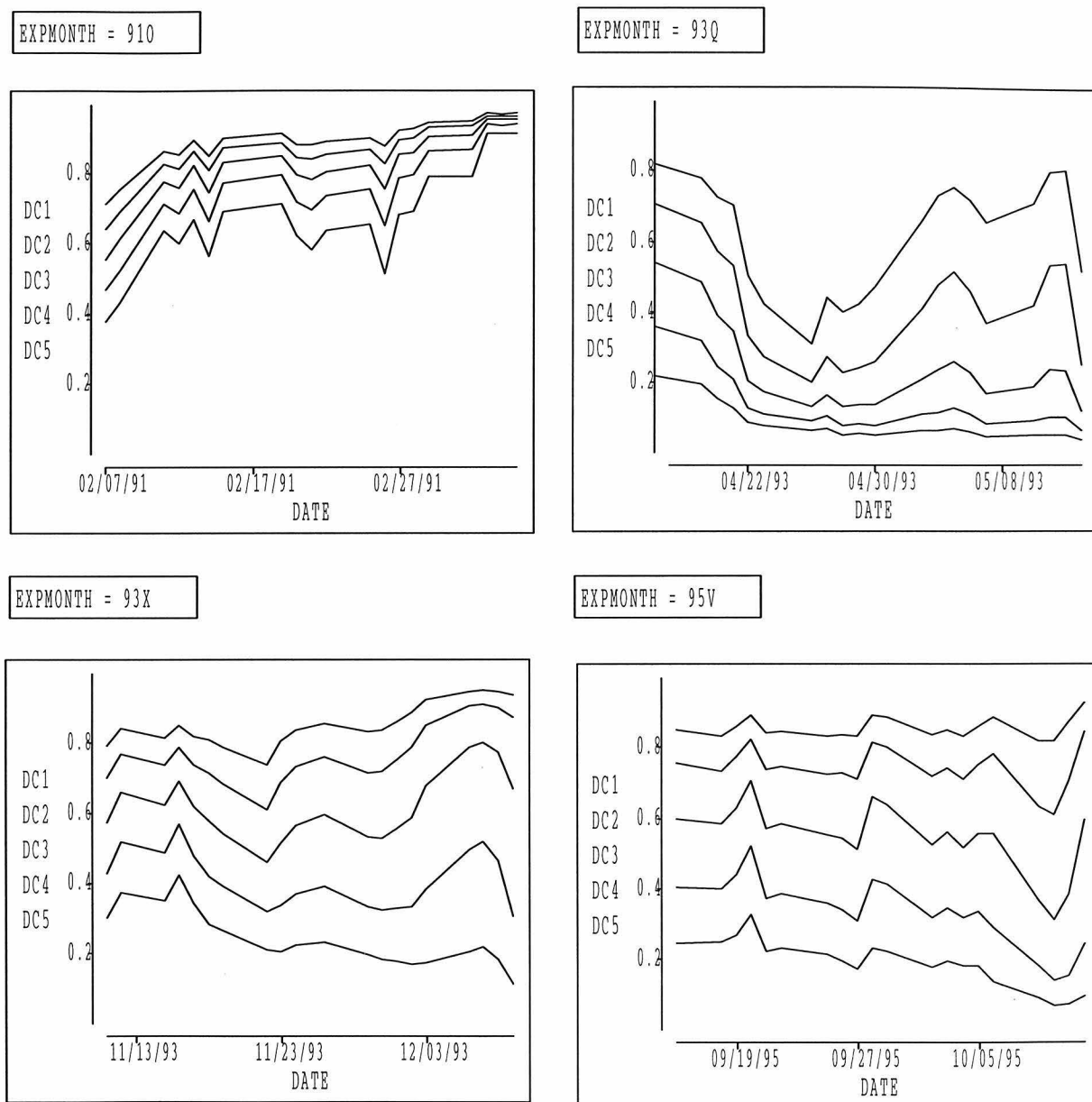
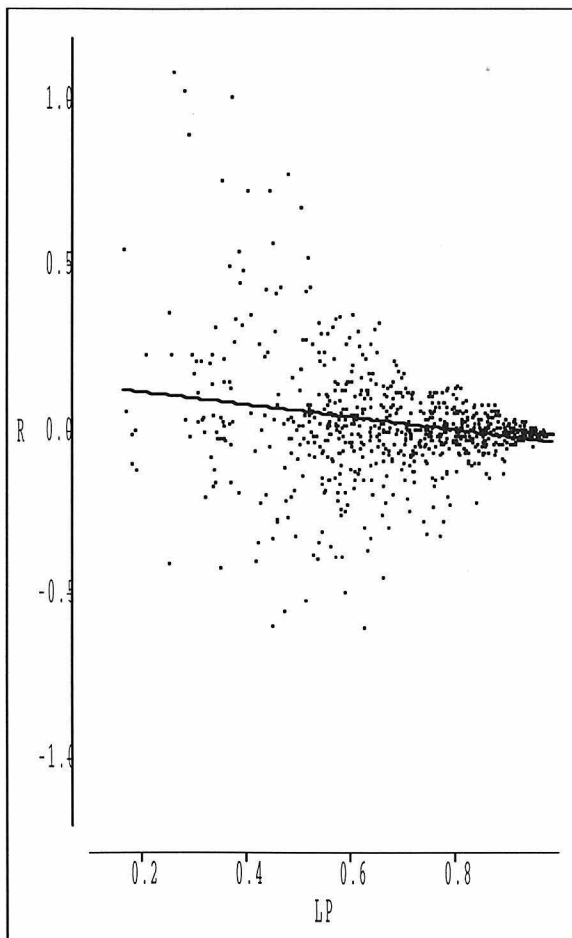


Figure 3.4: Time series plots of daily closing prices for the digital options with expiration dates in March, 91 (expmonth=91O), May, 93 (expmonth=93Q), December, 93 (expmonth=93X) and October, 95 (expmonth=95V). See Section 3.3.3 for details.

Model Equation			
R	=	0.1598	- 0.1943 LP



Model Equation			
X	=	0.0299	- 0.0456 LP

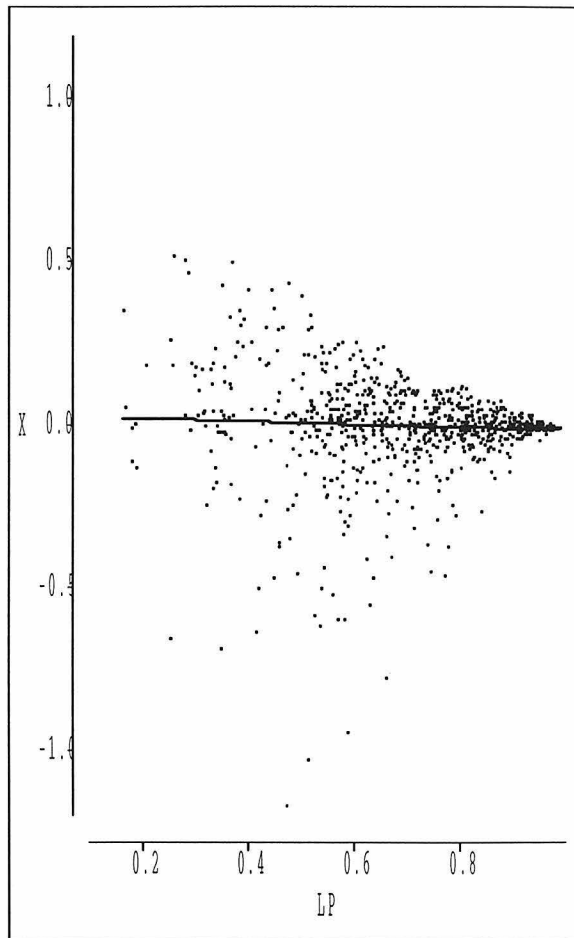


Figure 3.5: Scatter plots of traditional return r_{t+1} (left) and modified return x_{t+1} (right) against the beginning-of-the-period price p_t for winning digital options *DC3* (i.e., those that started at-the-money).

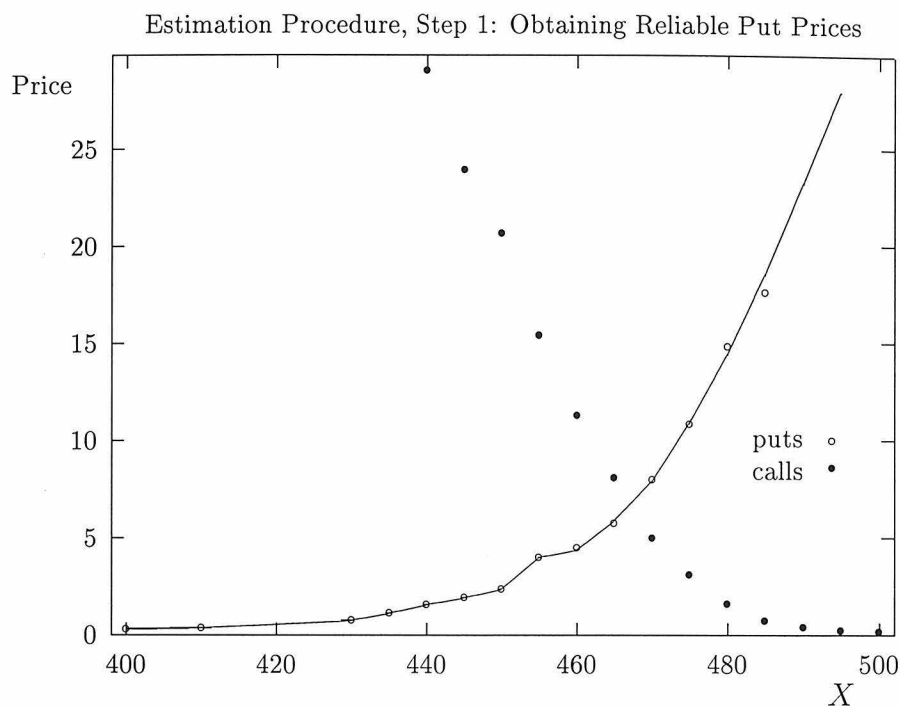


Figure 3.6: Step 1 of the estimation procedure for trading day 03/15/94 (maturity=04/15/94). Points are closing put and call prices. The solid line represents put prices estimated from the most liquid options.

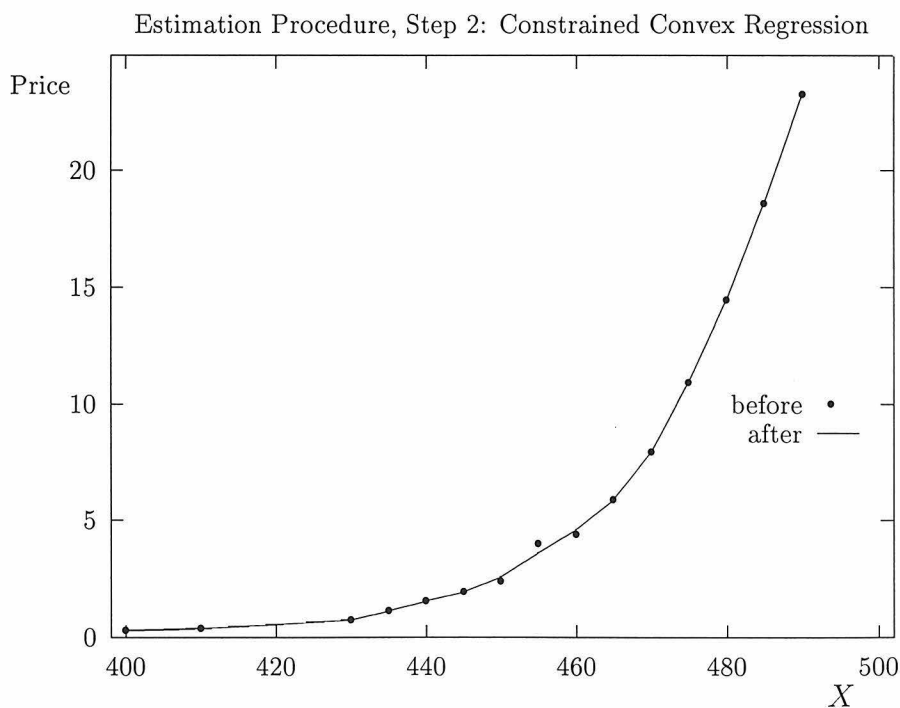


Figure 3.7: Step 2 of the estimation procedure for trading day 03/15/94 (maturity=04/15/94). Points and the solid line represent put prices *before* and *after* running the CCR, respectively.

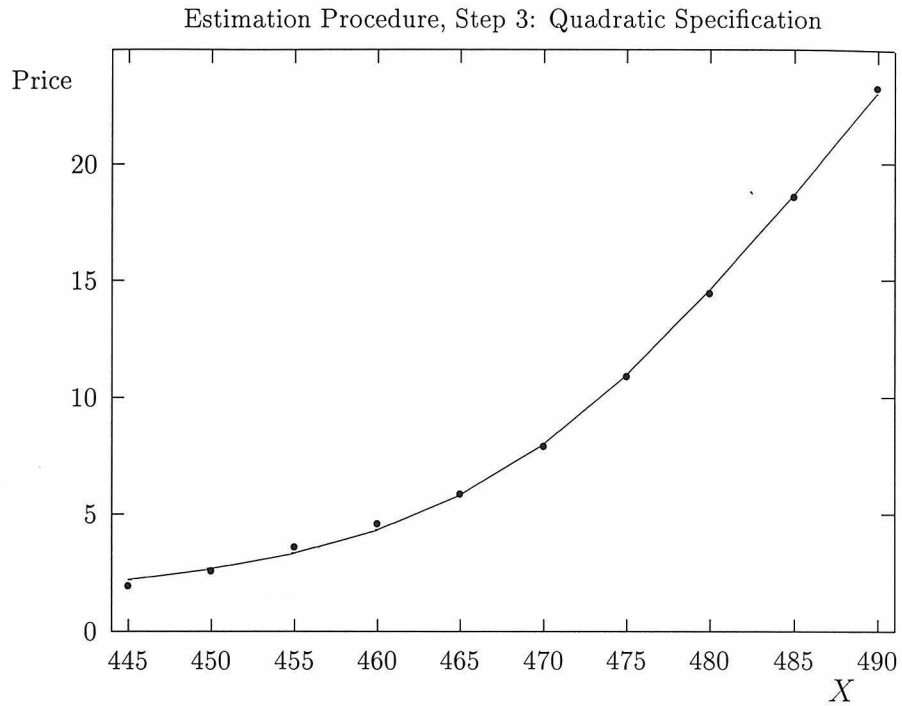


Figure 3.8: Step 3 of the estimation procedure for trading day 03/15/94 (maturity=04/15/94). The solid line is the continuous *quadratic* approximation of discrete put prices.

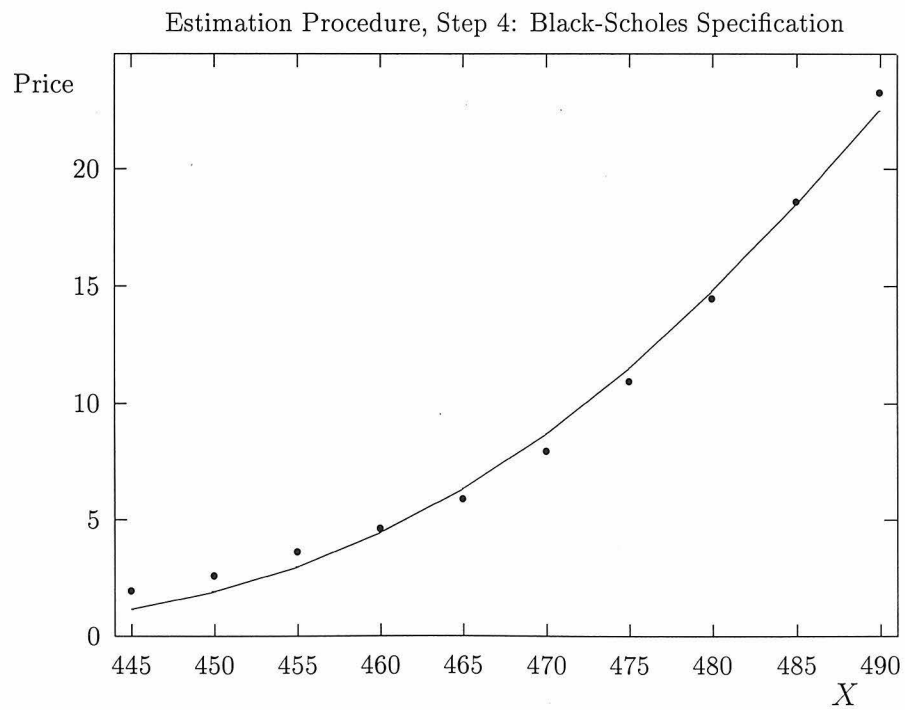


Figure 3.9: Step 3 of the estimation procedure for trading day 03/15/94 (maturity 04/15/94). The solid line is the continuous *Black-Scholes* approximation of discrete put prices.

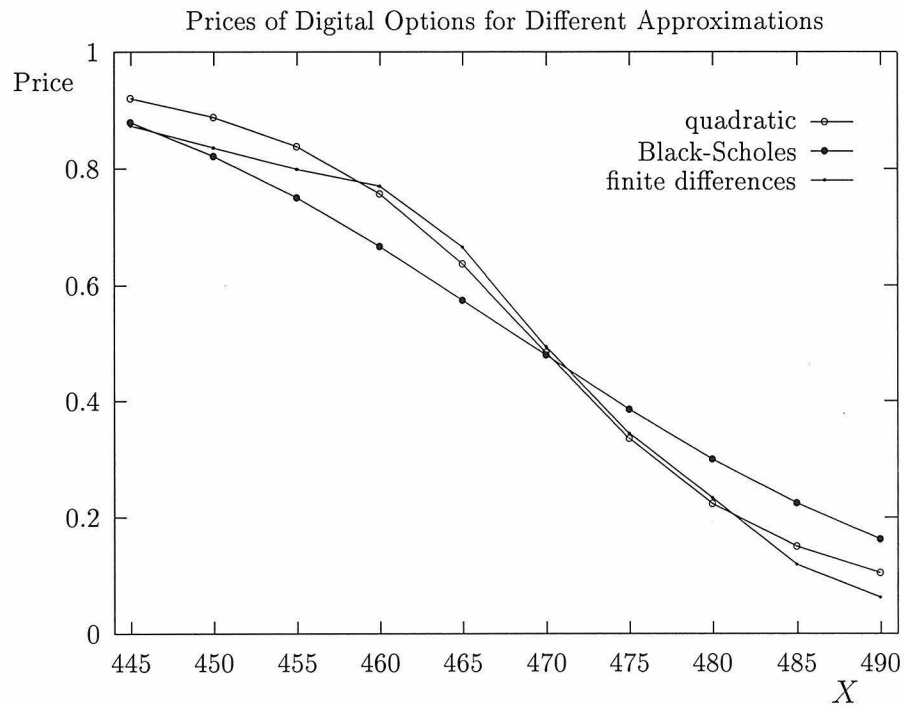


Figure 3.10: Prices of digital options for trading day 03/15/94 (maturity=04/15/94) computed using the quadratic specification, the Black-Scholes specification, and the finite difference approximation in equation (3.12).

Conclusion

This thesis provides a theoretical and an empirical study of rationality of asset pricing. We consider the relaxation of the traditional expectations model suggested in Bossaerts (1996) where investors may have potentially biased beliefs about crucial parameters of the economy. Investors remain rational, however, in the sense that they update their beliefs in a consistent and efficient manner applying the rules of conditional probability (Bayes' law). This means that of the two underlying assumptions of rational expectations, we relax the one of correct beliefs but keep the central assumption of rational learning.

The new model of Bayesian investors with biased beliefs is more general and realistic than the traditional hypothesis of rational expectations. More importantly, however, the new model offers simple and well-specified predictions which suggest that not *any* price history could be compatible with investors' rationality.

It turns out that the model of Bayesian investors with biased beliefs has more than just theoretical value. As the applications to the Iowa Experimental Markets and to the S&P 500 Index Options demonstrate, this model provides a parsimonious and fairly accurate description of empirical data. We find that prices of Arrow-Debreu type contracts traded in the IEM and prices of the digitals estimated from the S&P 500 Index Options are in accordance with rational learning but not with rational expectations.

The encouraging empirical results suggest that the new rationality restrictions can be used in many interesting applications. In particular, one can study rationality of pricing of other financial markets. One can also investigate whether errors in investors' expectations may explain other anomalies of traditional market efficiency. And, of course, one can use the new rationality restrictions to measure the extent of survivorship bias. We leave these possible applications for future research.

Bibliography

- Aït-Sahalia, Y., 1994, "Nonparametric Estimation of the Diffusion Coefficients in Ito Processes," University of Chicago working paper.
- Aït-Sahalia, Y., and A. Lo, 1995, "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," MIT Sloan School of Management working paper.
- Aumann, R.J., 1976, "Agreeing to Disagree," *Annals of Statistics* 4, 1236-39.
- Bachelier, L., 1900, "Theory of Speculation," in Cootner, P. (ed.), *The Random Character of Stock Market Prices*, MIT Press, Cambridge, MA, 1964; Reprint.
- Ball, R., 1992, "The earning-price anomaly," *Journal of Accounting and Economics* 15, 319-345.
- Ball, R., and P. Brown, 1968, "An empirical evaluation of accounting numbers," *Journal of Accounting Research* 6, 159-178.
- Ball, R., S. Kothari, and J. Shanken, 1995, "Problems in measuring portfolio performance: An application to contrarian investment strategies," *Journal of Financial Economics* 38, 79-107.
- Banz, R., and M. Miller, 1978, "Prices for State- Contingent Claims: Some Estimates and Applications," *Journal of Business* 51, 653-672.
- Basawa, I., and P. Rao, 1980, *Statistical Inference for Stochastic Processes*. New York: Academic Press.
- Biais, B., and P. Bossaerts, 1997, "Asset Prices and Trading Volume in a Beauty Contest," *Review of Economic Studies*, forthcoming.

- Black, F. 1975, "Fact and Fantasy in the Use of Options," *Financial Analysts Journal* 31, 36-72.
- Black, F., 1993, "Beta and return," *Journal of Portfolio Management* 20, 8-18.
- Blume, L., and R.F. Stambaugh, 1983, "Biases in Computed Returns: An Application to the Size Effect," *Journal of Financial Economics* 12, 387-404.
- Bondarenko, O., 1996, "Selection-Bias Induced Martingale Restrictions In Financial Markets," Caltech working paper.
- Bondarenko, O., and P. Bossaerts, 1996, "Expectations and Learning in Iowa," Caltech working paper.
- Bossaerts, P., 1995a, "The Econometrics of Learning in Financial Markets," *Econometric Theory* 11, 151-189.
- Bossaerts, P., 1995b, "Rational Price Discovery in Experimental and Field Data," Caltech working paper.
- Bossaerts, P., 1996, "Martingale Restrictions On Equilibrium Securities Prices Under Rational Expectations And Consistent Beliefs," Caltech working paper.
- Bossaerts, P., and P. Hillion, 1996, "Implementing Statistical Criteria To Select Return Forecasting Models: What Do We Learn?" *Review of Financial Studies*, forthcoming.
- Breeden, D., and R. Litzenberger, 1978, "Prices of State Contingent Claims Implicit in Options Prices," *Journal of Business* 51, 621-652.
- Brown, S., W. Goetzmann, and S. Ross, 1995, "Survival," *Journal of Finance* 50, 853-874.
- Chopra, N., J. Lakonishok, and J. Ritter, 1992, "Measuring abnormal performance: Do stock overreact?" *Journal of Financial Economics* 31, 235-268.

- Cox, J., and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics* 3, 145-166.
- De Bondt, W., and Thaler, R., 1985, "Does the Stock Market Overreact?" *Journal of Finance* 40, 793-805.
- Dykstra, R.L., 1983, "An Algorithm for Restricted Least Squares Regression," *Journal of the American Statistical Association* 78, 837-842.
- Doob, J., 1952, *Stochastic Processes*. New York: Wiley.
- Emanuel, D.C., and MacBeth, J.D. 1982, "Further Results on the Constant Elasticity of Variance Call Option Pricing Model," *Journal of Financial and Quantitative Analysis* 18, 533-554.
- Fama, E., 1970, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25, 383-417.
- Fama, E., 1991, "Efficient Capital Markets: II," *Journal of Finance* 46, 1575-1618.
- Fama, E., L. Fisher, M. Jensen, and R. Roll, 1969, "The adjustment of stock price to new information," *International Economics Review* 10, 1-31.
- Fama, E., and K. French, 1988, "Permanent and temporary components of stock prices," *Journal of Political Economy* 96, 246-273.
- Fama, E., and K. French, 1992, "The cross section of expected returns," *Journal of Finance* 47, 427-465.
- Fama, E., and K. French, 1993, "Common risk factor in the returns on stock and bonds," *Journal of Financial Economics* 33, 3-56.
- Forsythe, R., F. Nelson, G. Neumann, and J. Wright, 1991, "Forecasting the 1988 Presidential Election: A Field Experiment" in R.M. Isaac, ed., *Research in Experimental Economics* Vol. 4, Greenwich, CT: JAI Press, 1991, 1-44.

- Forsythe, R., F. Nelson, G. Neumann, and J. Wright, 1992, "Anatomy of an Experimental Political Stock Market," *American Economic Review* 82, 1142-1661.
- Foster, G., 1977, "Quarterly accounting data: Time-series properties and predictive-ability results," *The Accounting Review* 52, 1-21.
- Foster, G., C. Olsen, and T. Shevlin, 1984, "Earnings releases, anomalies, and the behavior of securities returns," *The Accounting Review* 59, 574-603.
- Goldenberger, D., 1991, "A Unified Method for Pricing Options on Diffusion Processes," *Journal of Financial Economics* 29, 3-34.
- Hall, P., and C. Heyde, 1980, *Martingale Limit Theory and Its Application*. New York: Academic Press.
- Hansen, L., and K. Singleton, 1982, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica* 50, 1269-1286.
- Harrison, J., and D. Kreps, 1978, "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics* 92, 323-36.
- Harrison, J., and D. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory* 20, 381-408.
- Harsanyi, J.C., 1967, "Games with Incomplete Information Played by 'Bayesian' Players, I: The Basic Model," *Management Science* 14, 159-82.
- Heston, S.L., 1993, "A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies* 6, 327-343.
- Hull, J., 1997, *Options, Futures, and Other Derivatives* 3rd edition, Upper Saddle River, New Jersey: Prentice Hall.
- Hull, J., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance* 42, 281-300.

- Hutchinson, J., A. Lo, and T. Poggio, 1994, "A Nonparametric Approach to Pricing and Hedging Derivative Securities Via learning Networks," *Journal of Finance* 49, 851-890.
- Ingersoll, J.E., Jr., *Theory of Financial Decision Making* Savage, Maryland: Rowman & Littlefield Publishers.
- Jarrow, R., and A. Rudd, 1982, "Approximate Option Valuation for Arbitrary Stochastic Processes," *Journal of Financial Economics* 10, 347-369.
- Jegadeesh, N., 1990, "Evidence of predictable behavior of security returns," *Journal of Finance* 45, 881-898.
- Jegadeesh, N., and S. Titman, 1993, "Returns to buying winners and selling losers," *Journal of Finance* 48, 65-91.
- Kamara, A., and T. Miller, Jr., 1995, "Daily and Intraday Tests of European Put-Call Parity," *Journal of Financial and Quantitative Analysis* 30, 519-539.
- Karatzas, I., and S. Shreve, 1987, *Brownian Motion and Stochastic Calculus*. New York: Springer Verlag.
- Keim, D., and R. Stambaugh, 1986, "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics* 17, 357-391.
- Kothari, S., J. Shanken, and R. Sloan, 1995, "Another look at the cross section of expected stock returns," *Journal of Finance* 50, 185-224.
- Kurz, M., 1994, "On the Structure and Diversity of Rational Beliefs," *Economic Theory* 4, 1-24.
- Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica* 53, 1315-1336.
- Lakonishok, J., A. Shleifer, and R. Vishny, 1994, "Contrarian investment, extrapolation, and risk," *Journal of Finance* 49, 1541-1578.

- Lo, A., and C. MacKinlay, 1988, "Stock Prices do not Follow Random Walks: Evidence From a Simple Specification Test," *Review of Financial Studies* 1, 41-66.
- Lo, A., and C. MacKinlay, 1990, "Data-snooping biases in tests of financial asset pricing models," *Review of Financial Studies* 3, 431-468.
- Lucas, R., 1977, "Understanding Business Cycles" in *Stabilization of the Domestic and International Economy*, Amsterdam: North Holland.
- Lucas, R., 1978, "Asset Prices In An Exchange Economy," *Econometrica* 46, 1429-1445.
- MacBeth, J., and L. Merville, 1979, "An Empirical Examination of the Black-Scholes Call Option Pricing Model," *Journal of Finance* 34, 1173-1186.
- Madan, D.B., and F. Milne, 1994, "Contingent Claims Valued and Hedged by Pricing and Investing in a Basis," *Mathematical Finance* 4, 223-245.
- Mehra, R., and E. Prescott, 1985, "The equity premium: A puzzle," *Journal of Monetary Economics* 15, 145-161.
- Mehra, R., and E. Prescott, 1988, "The equity premium: A solution?" *Journal of Monetary Economics* 22, 133-136.
- Merton, R.C., 1973a, "The relationship Between Put and Call Option Prices: Comment," *Journal of Finance* 28, 32-39.
- Merton, R.C., 1973b, "Rational Theory of Option Pricing," *Bell Journal of Economics and Management Science* 4, 141-183.
- Muth, J., 1961, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29, 315-35.
- Poterba, J., and L. Summers, 1988, "Mean reversion in stock prices: Evidence and implications," *Journal of Financial Economics* 22, 27-59.

- Rietz, T.A., 1988, "The equity risk premium: A solution," *Journal of Monetary Economics* 22, 117-131.
- Ritchken, P., 1996, *Derivative Markets* New York: HarperCollins College Publishers.
- Ross, S., 1976, "Options and Efficiency," *Quarterly Journal of Economics* 90, 75-89.
- Rubinstein, M., 1985, "Nonparametric Tests of Alternative Options Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Options Classes from August 23, 1976, through August 31, 1978," *Journal of Finance* 40, 455-480.
- Rubinstein, M., 1994, "Implied Binomial Trees," *Journal of Finance* 49, 771-818.
- Samuelson, P.A., 1965, "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review* 6, 41-50.
- Shimko, D., 1993, "Bounds of Probability," *RISK* 6, 33-37.
- Swamy, P., J. Barth, and P. Tinsley, 1982, "The Rational Expectations Approach to Economic Modelling," *Journal of Economic Dynamics and Control* 4, 125-148.
- Tikhonov, A.N., 1963, "Regularization of incorrectly posed problems," *Soviet Math.*, 4, 1624-1627.
- Tirole, J., 1982, "On the Possibility of Speculation Under Rational Expectations," *Econometrica* 50, 1163-81.
- Vapnik, V., 1982, *Estimation of Dependencies Based on Empirical Data*. Springer-Verlag: Heidelberg, New-York, Berlin.
- Wilmott, P., J. Dewynne, and S. Howison, 1993, *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.