# ADVANCED OPTION PRICING MODELS

## An Empirical Approach to Valuing Options

JEFFREY OWEN KATZ, Ph.D.
DONNA L. McCORMICK

# ADVANCED OPTION PRICING MODELS

## An Empirical Approach to Valuing Options

**JEFFREY OWEN KATZ, Ph.D.**
**DONNA L. McCORMICK**

**McGraw-Hill**
New York Chicago San Francisco Lisbon
London Madrid Mexico City Milan New Delhi
San Juan Seoul Singapore Sydney Toronto

The McGraw-Hill Companies

Copyright © 2005 by Scientific Consultant Services, Inc. All rights reserved. Manufactured in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

0-07-145470-5

The material in this eBook also appears in the print version of this title: 0-07-140605-0.

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps. McGraw-Hill eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please contact George Hoare, Special Sales, at george_hoare@mcgraw-hill.com or (212) 904-4069.

TERMS OF USE

This is a copyrighted work and The McGraw-Hill Companies, Inc. ("McGraw-Hill") and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill's prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED "AS IS." McGRAW-HILL AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw- Hill nor its licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

DOI: 10.1036/0071454705

*This book is dedicated to those who apply the knowledge and use their success to do some good in the world.*

# TRADEMARKS AND SERVICE MARKS

Company and product names associated with listings in this book should be considered as trademarks or service marks of the company indicated. The use of a registered trademark is not permitted for commercial purposes without the permission of the company named. In some cases, products of one company are offered by other companies and are presented in a number of different listings in this book. It is virtually impossible to identify every trademark or service mark for every product and every use, but we would like to highlight the following:

Neural-Hybrid Options Pricing Model is a trademark of Scientific Consultant Services, Inc.
Symantec is a trademark of the Symantec Corporation.
Numerical Recipes in C (book and software) is a service mark of Numerical Recipes Software.
Numerical Recipes in Fortran (book and software) is a service mark of Numerical Recipes Software.
N-Train is a trademark of Scientific Consultant Services, Inc.
Visual Basic, Visual C++, and Excel are trademarks of Microsoft Corporation.
TC2000 is a trademark of Worden Brothers, Inc.
Stricknet, stricknet.com are service marks of Stricknet.com, Inc.
GNU is a service mark of the Free Software Foundation.

Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

For more information about this title, click here

# CONTENTS

*This page intentionally left blank.*

# ACKNOWLEDGMENTS

The authors would like to express their gratitude to Stephen Isaacs for his continued encouragement, patience, and help, and to Sharon Rumbal and her production team at Alden Group for a job well done. We would also like to thank Bob Klein for moving us from articles to books, and Tom Bulkowski for his camaraderie and fine work on charting. And, as always, a special thanks to all those who have been there when we needed them.



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

# INTRODUCTION

This book is the result of years of original scientific research into the various elements that are required to accurately price options. We approached the topic in an objective and systematic manner, just as we did in our study of futures trading systems in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000). The method: a traditional labor- and data-intensive study involving thousands of hours of computer time; the result: a wealth of practical findings of direct relevance to those who use options to speculate or hedge.

An in-depth investigation was necessary because of the nature of the subject under study. As is well known, options are a fiercely competitive, zero-sum game. The amateur usually does not stand a chance and even experienced players can find it difficult to use them effectively. Therefore, to successfully speculate or hedge with options, every edge is necessary. As in almost all realms of endeavor, knowledge can provide the biggest edge. A thorough, clear-sighted understanding of the subject and the factors that influence it are critical and can only be achieved by implementing an objective, scientific approach.

When it comes to options, knowledge can mean the difference between making a profit and taking a loss. For example, there is great advantage in knowing how to identify and exploit mispriced options. We are not referring to small mispricings that



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

only the most efficient arbitrageur or market maker can exploit, but to gross mispricings that sometimes appear and, equally quickly, disappear. There is an edge in knowing what to look for and in knowing how to find it.

We know we must search for gross mispricings, but how do we find them? An option pricing model is needed. However, not just any pricing model will do. To gain an edge, a model must correctly value options under circumstances that cause standard models to break down. In addition, the model must be used with valid inputs; even the best model will not yield accurate prices if the model's inputs are in error.

In this book, we have done the research, described the logic behind it and the steps involved, and presented the results as practical solutions. We have analyzed standard option pricing models, discovered their flaws, and investigated better estimators of volatility and other model inputs. We have also explored nonstandard, rather innovative ways to achieve more accurate appraisals of option value. It is our sincere hope that this will give you the edge you need in the tough options game.

## THINKING OUT OF THE BOX

Basically, there are four kinds of books on the subject of options: (1) those that deal with the basics and the strategic use of options, (2) academic texts that discuss theoretical models of how stock returns are generated, models that are then used to construct option pricing formulae, (3) practical guides that provide advice (derived largely from personal experience) on how to grab profits from the markets, and then there is (4) the book you are holding in your hands.

*Options as a Strategic Investment* by McMillan (1993) represents a well-written, classic example of the first kind of book. Such texts provide a basic understanding of options: they cover Black-Scholes and the Greeks, discuss the effects of time decay and of price changes in the underlying stocks, and contain everything you would ever want to know about covered writes, naked puts, straddles, spreads, equivalent positions, arbitrage, and the risk profiles of various positions. Such background information can help one maneuver intelligently around the options markets, as well as tailor positions to fit individual needs and expectations. However, unless you are one of the rare few who can divine the future behavior of the markets, relying only on texts such as these will not provide the practical knowledge you need to gain a statistical edge.

The second kind of book, the academically-oriented text, is heavy in theory. Such literature mostly consists of the kind of material that has been implemented in the computer programs used daily by market makers and options traders, or of esoterica that is primarily of interest only to academic theoreticians. An example of a good book of this genre is *Black-Scholes and Beyond* by Chriss (1997). Most of these texts present detailed theoretical analyses of option pricing models like Black-Scholes and Cox-Ross-Rubinstein, as well as variations thereon. If you are like Katz, an individual who enjoys playing with theoretical models and running Monte Carlo simulations, then you will find such books a lot of fun. Active options traders concerned with their bottom lines, however, probably will not greatly benefit from such reading.

Good books of the third kind are rather rare. Perhaps the best example of this kind of publication is *McMillan on Options* (McMillan, 1996). Such books are often based on personal experience and get down to the nitty-gritty by demonstrating how to take profits from the options markets. They cover topics like how volume and implied volatility can be tip-offs to large and profitable moves, how to interpret and profit from extremes in volatility, how to recognize significant situations, and much more. In terms of providing hard-hitting, practical insight for the active options trader, such books can be extremely useful.

The last category of book is, as far as we know, occupied solely by ours. This distinction is the result of taking a unique perspective on the subject. Although, as in academic texts, we discuss distributions of returns, the Central Limit Theorem, and random walks, in our book there is a heavier-than-usual emphasis on the empirical. Many books on option pricing focus on theoretical models and only use data in an effort to test the assumptions made by these models. Rather than discuss *theoretical* distributions of price changes, we examine *actual, real-market* distributions and how they differ from the theoretical distributions assumed by such popular pricing models as Black-Scholes. By using this approach we provide far more extensive coverage of real-market behavior than do most texts on pricing models, together with a wealth of data and analyses not available elsewhere. In our examination of distributions, we search for and find practical information about pricing that anyone can use to take money out of the markets. In short, we provide a detailed exploration of standard assumptions and then demonstrate how and where they are violated by real-market behavior. Our integrative approach leads to insights, not merely of theoretical significance, but of practical value for the trader trying to pull money from the options markets. The information presented is otherwise hard to come by, but essential to anyone wishing to become successful in the highly competitive arena of options trading.

## IMPROVING OPTION PRICING STRATEGIES: A SCIENTIFIC INVESTIGATION

In this book, we are exploring a lot of new ground. The emphasis is on asking a wide range of questions and attempting to find the answers by studying real-market stock and options data. The ultimate goal is to find new and effective techniques for modeling the price movements of stocks and the value of the options that trade on them. At the same time, we investigate all the factors that bear upon the pricing of options, examine the standard pricing models, and discover where those models go wrong and lead to mispricing. The outcome: more effective ways to price options. We not only look at the subject from a theoretical point of view, but also study the actual movements of prices in the market, how they are distributed, and what the patterns tell us.

Our approach is empirical and analytic; our style is intuitive and practical. The work is based on continuing investigations by the authors who are themselves options traders. Much of this work involves extensive examination of long-persistent market characteristics and in-depth statistical and mathematical analyses. We present solid, research-based information of a kind often buried in academic journals, but do so in a manner that is immediately and practically useful. The analysis and results should be as valid and relevant many years from now as they are today.

## ASSUMPTIONS MADE BY POPULAR MODELS: ARE THEY CORRECT?

In the world of equity options, accurate, mathematically-based estimates of fair price and of the so-called "Greeks" are crucial for success. Many professionals, as well as amateurs, still use the traditional Black-Scholes model to price options. Likewise, most standard texts on the subject focus primarily on Black-Scholes, while occasionally discussing Cox-Ross-Rubinstein and other related models. A common feature of these models is the assumption that, on a logarithmic scale, the distribution of returns (profits or losses) in the market is normal (Black-Scholes), something close to normal, or something that approaches normal in the limit (Cox-Ross-Rubinstein). Most "random walkers"—proponents of the Efficient Market Hypothesis (EMH)—would argue that the assumption of normally distributed returns is justified by the Central Limit Theorem and the "fact" that stock returns reflect the accumulation of large numbers of equally small, random movements. However, do stock returns really follow the familiar bell-shaped curve of the normal distribution? No, they do not!

It is well known (and easy to verify) that empirical distributions of returns deviate from normal by, at the very least, having longer tails—extreme returns are more frequent than would be expected from a normal or near-normal distribution. This does not necessarily imply that price movements are following something other than a random walk. Perhaps the "small, random movements" of the walk are simply not homogeneous. If some random steps are drawn from a different underlying distribution than others, e.g., a distribution that has a larger variance, a long-tailed distribution of returns might result. Of course, the EMH itself might be in error; perhaps stock prices are not random, but have memory and move in trends. Again, the result could be a long-tailed distribution of returns. Regardless of the reason, the difference between the empirical and normal distributions has significant implications for option pricing. Frankly, models that assume normality (like Black-Scholes) cannot be trusted to consistently provide correct option prices and, therefore, can cost hedgers and traders serious money!

Some might argue that, although the assumptions underlying Black-Scholes and related models are technically violated, the prices generated by these models approximate correct values well enough for practical purposes. Indeed, in many instances, they do. However, it is easy to find conditions under which the prices, and other data that are generated by a model like Black-Scholes, dramatically miss the mark. Consider the so-called "volatility smile" that has been the subject of many academic papers. The smile appears when implied volatility is plotted against strike price: deeply in- or out-of-the-money options have higher implied volatilities than at-the-money options. If we take a somewhat different perspective, when using Black-Scholes deeply in- or out-of-the-money options appear overpriced relative to at-the-money options. This is exactly what would be expected when a model that assumes a normal, short-tailed distribution of returns is applied to markets with long-tailed distributions of returns. In addition, due to the mean-reverting nature of volatility, pricing errors become substantial when historical volatility is an input to the model and reaches either very high or very low levels. A number of other statistical features of the underlying security can also result in seriously mispriced options.

The errors mentioned above are not small and of interest only to academicians. Under certain circumstances, many of these errors reach a magnitude that is quite significant, even to the average options trader. Because of the popularity and naive use of Black-Scholes and similar models, options can often be found trading near the model's estimate of fair value when, in fact, they should be priced substantially higher or lower. A savvy options player can take advantage of the discrepancies and sell the overpriced options or buy the underpriced ones. A more sophisticated and realistic pricing model—one that takes into account the actual distributions of returns seen in the market under various conditions and that is less subject to systematic error—can be a powerful weapon for the trader or hedger seeking a decisive edge over his or her competitors. Can such a model be developed? What is involved in the construction of an improved pricing model? Finally, assuming it can be built, how will such a model perform in the competitive world of equity options?

This book examines the specific conditions under which Black-Scholes and other popular models fail to provide good estimates of fair option prices, the actual behavior of the market and how it differs from what the standard models assume, how to use such information to arrive at better option price estimates, and the steps involved in building better option pricing models.

## OPTIMAL MODEL INPUTS

In addition to the pricing model itself, certain inputs require special attention. As anyone familiar with options knows, volatility is one of the major factors that determine the value of an option.

When using an option pricing model, historical volatility is often employed as one of the inputs. In such a case, historical volatility is being used (knowingly or not) as a proxy for future volatility. It is actually future volatility, not historical volatility, that determines the worth of an option. Therefore, the direct use of historical volatility as an input to a standard model can lead to systematic and often severe mispricing. To some extent, volatility appears to be mean-reverting. If recent historical volatility is extremely high, one can expect future volatility to be lower; if recent historical volatility is extremely low, future volatility can be expected to be higher. As it is future volatility that matters, the use of historical volatility can distort option price appraisals: extremely high historical volatility can lead to overpricing and extremely low historical volatility can cause the model to underprice options. The use of estimated future volatility, rather than simple historical volatility, can improve price estimates based on any option pricing model. Fortunately, volatility is much more predictable than price movement and, as we will show, predictive models can be constructed for it.

A substantial amount of space has been dedicated to the prediction and estimation of volatility, considering its great importance to the pricing of options. Implied volatility is also examined in detail. Finally, the way in which time (another major determinant of an option's value) and volatility are related, both in theory and in actual practice, is studied.

## WHAT IS COVERED IN THE CHAPTERS?

The book begins with coverage of the basics of options, fair value, and pricing models. Chapter 1 provides a brief, but detailed, review of options and option terminology. A clear discussion of the Greeks and the use of standard option pricing models is included. In addition, the basics of speculation, one form of arbitrage, and equivalent positions are reviewed. The use of option characteristic curves, or price response charts, is also illustrated.

Chapter 2 attempts to elucidate the nature of fair value. What is fair value? How is fair value related to the efficient market hypothesis? Is fair value a unitary entity or a multiheaded beast and, if it is the latter, what are its heads or components? The chapter considers fair value in terms of both speculation on future prices and certain kinds of arbitrage. A simple Monte Carlo experiment, in which synthetic stock and option prices are generated and examined, is presented to illustrate some of the concepts developed in this chapter. The illustrative model examined in the experiment is the starting point that, with modifications, becomes a real option pricing model in the next chapter.

Chapter 3 contains an examination of the two most popular option pricing models: Black-Scholes and Cox-Ross-Rubinstein. These two models have such a pervasive presence in the world of options that their influence on prices, and the behavior of traders and hedgers, is overwhelming. The assumptions on which these models are based are investigated and the models themselves developed, illustrated, and dissected. A common thread in the assumptions underlying these models is discovered and analyzed. In this chapter, there is an extensive discussion of the log-normal distribution and its impact when used as a basis for understanding the underlying stock price movements or returns, as it indeed is used in the standard pricing models. Cox-Ross-Rubinstein (also known as the "binomial model") is fully developed and illustrated with both the Monte Carlo method and with pricing trees. The Black-Scholes equations are also presented and some interesting features of these equations (such as the fact that they are direct calculations of the expectation of future option prices, under the assumption of a log-normal distribution of returns) are demonstrated numerically with the aid of numerical quadrature. Finally, some phenomena associated with log-normally distributed stock price movements are discussed; specifically, the fact that if there is an even probability of either a win or a loss, there must be a positive net return, and if there is an average return that is neither positive nor negative (i.e., a return that is zero or breakeven), then the probability of any stock trade taking a loss must be greater than 50%, all this being true if stock prices are indeed log-normal random walks. The chapter concludes with an examination of stock price movements in the NASD and NYSE, as well as those generated in the course of a Monte Carlo experiment and designed to behave according to the log-normal random walk assumption.

After the heavily theoretical discussion in Chapter 3, the orientation becomes empirical.

Chapter 4 studies the distribution of actual stock returns by examining their statistical moments. The reason for studying stock price returns from the perspective of moments is to better characterize the distributions involved. Distributions of underlying stock price movement are a major determinant of the worth of options trading on those stocks. The first four moments of a distribution are defined and discussed. Moments are useful statistics in characterizing the shape either of a theoretical distribution or of one constructed based on sample data. Once the basics are defined, the database used in all the studies that follow in this chapter is discussed, as are the basic software tools and methodology.

Chapter 4 also contains a series of empirical studies or tests in which a variety of questions are answered on the basis of an examination of the statistical moments of the distributions of returns. The study of moments can help determine, for example, whether the underlying distribution of price movements in stocks is indeed log-normal, as popular option models assume. If the distribution is not log-normal, moments can help characterize its shape and how it differs from the log-normal baseline. In this chapter, moments are examined in relationship to holding period, day of the week, time of the year, and time with respect to option expiration.

Although it may sound strange to study statistical moments, the reader who is familiar with options has already encountered the second moment, which is, in fact, volatility. Almost every trader is familiar with the first moment, which is simply the expected gain or loss over the holding period, i.e., the trend. The following are some of the questions asked and answered in Chapter 4: Does volatility scale with the square root of time? Are successive returns independent of one another or, equivalently, is the market efficient and unpredictable? Is the distribution of returns log-normal and, if not, how does it differ from log-normal? Does volatility vary with day of the week and time of the year? Most traders and hedgers would say that the answer to these questions is yes. If volatility does vary with time, which days are the most and least volatile; which times of the year are most and least volatile? And what about the other moments, like skew, kurtosis, and expectation? Finally, what does the characterization (in terms of moments) of the distribution of real-market returns reveal about the worth of options under a variety of different conditions?

Chapter 5 is dedicated to that statistical moment dear to the heart of every options player, whether speculator or hedger: volatility. When pricing options, the volatility of concern is not historical, but future; it is future volatility that can be expected to occur during the holding period. The focus of Chapter 5 is on the estimation or prediction of future volatility for the purpose of appraising options. This chapter is probably unlike any other chapter on volatility that you have read in any other book. The discussion begins with measurement reliability, as seen from the perspective of a psychometrician. Although psychometrics may seem far afield from the world of finance, it turns out that some of the problems involved are similar when abstracted from the specific content and require similar solutions. Some of the basics of psychometrics or "test theory" are discussed, such as estimating reliability using split-half correlations. Model complexity and other issues are then examined. At this point, the chapter covers the methodology employed, including the particular databases used, software involved, and the calculation of implied volatility (required in some of the studies). Then begins a series of tests concerned with various aspects of volatility.

Study 1 in Chapter 5 examines the common use of simple measures of historical volatility in pricing options. It asks a variety of questions. How good is historical volatility as a predictor of future volatility? Under what conditions does the use of historical volatility lead to serious pricing errors? Can historical volatility somehow be adjusted to yield better option appraisals? How reliable is historical volatility as a measure of the underlying trait of volatility possessed by a given stock at a given time?

Moreover, which of the many different measures of historical volatility should the trader or hedger employ? Although most users of standard models might not be aware of it, there are indeed a number of ways in which a measure of historical volatility may be obtained. The study leads to some interesting findings regarding the relationship between historical volatility and future market behavior and, in turn, the fair prices for options. One finding of critical importance is that the use of uncorrected or "raw" historical volatility can result in appraisals that are systematically distorted. In other words, standard models applied in the standard way, using historical volatility as one of the inputs, will, under certain conditions, yield theoretical fair prices that are far from the actual worth of the option being priced.

The goal in Study 2 is to determine whether the combined use of two different measures of historical volatility can improve the estimation of future volatility and, thus, of pricing accuracy. Here the technique of multivariate regression is employed. Some interesting charts are presented depicting the relationship between short- and long-term historical volatility and future volatility.

Study 3 is an in-depth analysis of the reliability of volatility measurements and the stability of the underlying volatility being measured. Here, the ingenious use of psychometric theory appears. Several kinds of volatility measures are considered and their reliability and validity assessed. Some surprising findings emerge—findings that can provide immediate benefit to the user of options.

Study 4 in Chapter 5 attempts to construct a more sophisticated estimator of future volatility; multivariate polynomial regression is employed. Inputs to the volatility forecasting model include historical volatility for two periods (using the most reliable measures found in the previous studies), as well as cycle harmonics to capture stable seasonal variations in volatility. The results are dramatically better estimates of future volatility. Regardless of the option pricing model used, this is the kind of volatility estimate that should be employed. This chapter does not include consideration of standard approaches to forecasting volatility, e.g., GARCH; such approaches and models have received extensive coverage by other authors. Instead, Chapter 5 embodies the spirit of this book, which is to think out of the box, to apply a variety of techniques that are not in general use, and to gain an edge, in terms of both simplicity and power.

Study 5 examines implied volatility as an estimator of future volatility. Again, some interesting findings emerge. Contrary to popular belief, implied volatility is not necessarily any better than historical volatility when used in a pricing model.

Finally, in Study 6, historical and implied volatility are together used to forecast future volatility. Again a technique, Sewall Wright's path analysis, is borrowed from another discipline that might seem to be far afield. Sewall Wright was a geneticist who explored correlations of traits that were passed on through generations. Path analysis allows causal inferences to be made from correlation matrices; these inferences concern the strength of a causal influence of one variable upon another when considered in the context of a number of variables and possible configurations of paths of causation. Path analysis helps answer questions like the following: to what extent is implied volatility determined (1) by historical volatility, and (2) by future volatility, perhaps as a result of the leakage of inside information?

Chapter 6 deals with pricing options using empirically-based conditional distributions. In standard models like Black-Scholes, theoretical distributions are assumed a priori. As has been demonstrated, the distributional assumptions made by such models often appear to be violated by the price behavior of real stocks; this leads to option pricing errors. What happens if the a priori distributions are replaced with distributions determined from real-market behavior? This is the central idea behind the use of conditional distributions. Various questions regarding the use of conditional distributions to price options are investigated.

One of the problems with conditional distributions derived from market data concerns curve-fitting and degrees of freedom. Chapter 6 begins with an extensive discussion of these issues, which includes the use of rescaling as a means of reducing the degrees of freedom consumed when constructing conditional distributions. General methodology, including data and software, are then briefly discussed. A series of empirical studies follow.

Study 1 explores a simple pricing model in which raw historical volatility is the only conditioning variable. Theoretical option premiums, determined from conditional distributions, are compared to Black-Scholes, with the latter model computed both with raw historical volatility and with an improved estimate based on raw historical volatility that corrected for nonlinearity and regression to the mean. Also presented are charts of theoretical premiums from Black-Scholes and from the empirical distribution methodology for several strikes.

Study 2 in Chapter 6 is essentially a replication of Study 1, except that raw historical volatility is replaced with a high quality estimate of future volatility.

In both Studies 1 and 2, the distributions employed are not detrended; any consistent trends, volatility-related or not, were allowed to influence theoretical option prices derived from the conditional distribution methodology. In Study 3, a reanalysis is performed with detrended distributions, i.e., the effect of trend is removed by adjusting the first moment of each distribution (its mean) to zero.

In Study 4, historical skew and kurtosis are added to the model as conditioning variables; they are computed in a manner similar to that used to compute historical volatility. The effect of skew and kurtosis on the worth of puts and calls at different levels of moneyness is examined.

Study 5 examines the effect of trading venue on the distributions of returns and, in turn, on option prices. Again, puts and calls, with varying strikes and moneyness, are examined and their empirically determined prices are compared to Black-Scholes.

In Study 6, distributions conditional upon the status of a popular technical indicator are computed and used to price options. Crossovers of the stochastic oscillator at the standard thresholds are examined. Although such indicators are of little use to speculative traders dependent on directional movement (see Katz and McCormick, 2000), they may be significant when trying to characterize aspects of the distribution of returns other than trend. The chapter concludes with a general discussion of the methodology, its strengths and weaknesses.

One of the problems with the use of conditional distributions is a heavy demand for massive amounts of data because the degrees of freedom required by the methodology can be enormous. One way of making the empirical approach to pricing options more workable is to employ a general nonlinear modeling technique that can smooth out the noise while still capturing the true relationships revealed by the conditional distribution methodology. In Chapter 7, nonlinear models that potentially have the ability to accomplish this are explored; specifically, neural networks, multivariate polynomial regressions, and hybrid models.

Chapter 7 begins with a detailed discussion of neural networks, multivariate polynomial regressions, and hybrid models. We cover everything from the issues of numerical stability and the accumulation of round-off errors to the use of Chebyshev Polynomials as a means of dealing with problems of colinearity.

A general problem with neural networks and multivariate polynomial regressions is the tendency to curve-fit the data. The use of a hybrid model is one possible solution to this problem. For example, a hybrid model might incorporate a neural network in which the output neuron behaves like Black-Scholes. The intention is to build into the model as much knowledge as possible, even if it is only approximate, and to do so in such a way that the errors in the approximation can be corrected for by various elements in the model. Black-Scholes, although exhibiting systematic error that can be great under certain conditions, does yield a reasonable first approximation to the worth of an option. What if some of the inputs to Black-Scholes could be tweaked to force it to yield more accurate appraisals? This is the idea behind the hybrid model under discussion. Why take the trouble of developing a hybrid model, rather than a simple neural network or polynomial regression? Because, with a hybrid model, the number of free parameters required to obtain a good fit to the data is substantially lower and, therefore, the solution much less prone to curve-fitting and the excessive consumption of degrees of freedom.

Before attempting to use nonlinear modeling techniques to price options based on real-market data, it is important to discover whether they could accurately emulate Black-Scholes. If a general nonlinear model cannot emulate Black-Scholes, how can it be expected to capture the possibly more complex relationship between factors such as volatility, time, and strike, and the fair premium of an option that might exist in real-market data? The first two studies of Chapter 7 answer this question.

In Study 1, a neural network is trained to emulate Black-Scholes and its performance is evaluated. In Study 2, a multivariate polynomial regression is fitted to the same Black-Scholes data set and evaluated for performance. Both approaches are demonstrated to be capable of doing a good job. Attention is then turned to real-market data.

Study 3 investigates the ability of a polynomial regression to accurately capture the relationships between fair premium and the model inputs that are seen in data derived from real stock returns using a methodological equivalent to conditional distributions. The behavior of the polynomial model is compared to Black-Scholes and is evaluated in terms of its ability to accurately describe the empirical pricing data, as well as to filter out random variations that are seen in such data. Several tables and charts illustrating the pricing behavior are presented.

Study 4 repeats Study 3, but uses a neural network instead of a multivariate polynomial.

Study 5 examines a hybrid model that consists of a neural network with a special Black-Scholes output neuron and some additional processing elements.

Chapter 7 concludes with a discussion that compares and evaluates the various approaches. The following kinds of questions are addressed. Which approach best captures the relationships required to accurately price options? Which approach is most susceptible to undesirable curve-fitting and which is most resistant? What are the specific problems that must be dealt with when attempting to develop models of this kind? And, what are the respective potentials of these various techniques when the goal is to develop a coherent and sophisticated option pricing model? Although long and elaborate, Chapter 7 is rather unique when it comes to the treatment of option pricing.

Chapter 8 revisits volatility. In this chapter, a wide range of variables beyond those explored in Chapter 5 are examined as potential predictors of future volatility. These variables include measures such as historical skew and kurtosis, and the status of various technical indicators. The aim is to obtain the best possible estimate of near-future volatility. Attempts are made to answer questions about volatility that were raised in the course of the other investigations in this book. For example, in Chapter 6, skew and kurtosis are found to affect fair premium. Do skew and kurtosis have an effect on premium that is mediated through volatility; in other words, do these variables influence the expectation regarding future volatility and, therefore, have some impact on the value of options? What about the status of technical indicators? Are their effects on option prices due to differences in future volatility, rather than just to differences in the shape of the distribution of future returns? The answers to these questions are found in Chapter 8.

So far, with few exceptions, theoretical option prices based on observed movements in stock prices have been the focus of our studies. In Chapter 9, comparisons are made between these theoretical option prices and real-market option prices, i.e., the prices at which the relevant options are actually trading. The chapter contains a discussion of the data and software used and attempts to answer a variety of questions concerning the relationship between option prices computed with different models to those observed in the actual marketplace. For example, when there is a wide disparity between the two figures, do real option prices fall closer to Black-Scholes or to what one of the better models suggests? Can one profit by looking for large discrepancies between the theoretical price of an option and the price at which it is actually trading? How much does the use of Black-Scholes and other popular models influence the options market?

Finally, there is the *Conclusion*. Here we summarize our findings and provide you with information on how these insights can improve your option pricing strategies.

## WHO WILL BENEFIT?

This book is intended for everyone from the professional quant to the student who desires a better understanding of, and strategy for, pricing and trading options. Professional and institutional options players, who may be adjusting the standard models in an intuitive fashion, will find this book useful in that it may articulate their intuitive understanding of option pricing in such a way as to allow the automation or computerization of the pricing process. This could prove more effective than the intuitive approach, e.g., by leading to the inclusion of a wider range of conditions that will let hundreds, if not thousands, of options positions be quickly and repeatedly scanned for more frequent trading opportunities throughout the day. The sophisticated options player will also find this book helpful in that it will place him or her on a more level playing field with professional and institutional traders—we are giving you information about models that they may be using to obtain a closer estimate of an option's future value. The book should also be of interest to the academician or student trying to develop better theories and methods of option pricing. We expect that all readers will find within these pages at least one or two useful insights that will make their approach to option pricing more profitable.

Although this book contains a lot of mathematics and statistics, we make every effort to explain things—especially findings that are of practical use—in "plain English." We also include many tables and charts to illustrate the phenomena under discussion. Those who are mathematically challenged may want to ignore the equations and stick to the less technical text. Conversely, quants might want to skip the introductory material, like Chapter 1 on the basics.

## TOOLS AND MATERIALS USED IN THE INVESTIGATION

As in any scientific investigation, the one essential element that is required is a subject of study. In the present case, that subject is the world of option pricing, as represented by real-market data. We used data from two sources: (1) stock price and volume data were obtained from the Worden Brothers TC-2000 database (www.worden.com); (2) option pricing, volume, and open interest data were obtained from www.stricknet.com.

In our investigation, we focus primarily on short-term equity options, with some attention to index options. The main reason for using short-term options in our study is because they are the most liquid and are the kind traded by the authors and most other active traders. Secondly, by working with short-term options, we can (somewhat safely) ignore interest rates and dividends, thereby simplifying our investigations. The simplifications, and their possible impact on the findings, will be discussed whenever they arise.

The tools used to investigate the subject come in the form of software. To a great extent, the software was custom written by Katz either exclusively for this book or for his company, Scientific Consultant Services, Inc. For example, *N-Train* (his neural network development system) was used in Chapter 7 as part of the study of nonlinear pricing models. Other custom developed software included libraries containing routines for Black-Scholes pricing, calculation of statistical moments, numerical quadrature, regression, data management, volatility calculations, general mathematics, pseudo-random number generators, probability functions, utility functions, and the *Neural-Hybrid Options Model Library*. Custom code was written for each of the tests and studies, as well as for actual pricing models, based on the resultant findings. At the end of this book, for the benefit of those who would like to replicate and expand on our work, we have provided information on how to obtain the *Companion CD* and other software used in our studies.

In addition, a number of off-the-shelf products were used in the course of the investigation. The GNU C/C++ and Fortran compilers familiar to Unix users were used to compile code. Microsoft's Excel spreadsheet was used for visualization (charting) and presentation, as well as for some final analyses. Lastly, routines from *Numerical Recipes in Fortran 77* and *Numerical Recipes in C* (both books and software packages by Press et al., 1992) were employed.

## AN INVITATION

We invite you to visit our Web site: www.scientific-consultants.com. Here, you will find updates about our research and other information that you may find useful.

We also enjoy hearing from our readers; you are always welcome to send us your questions or comments by e-mail to katz@scientific-consultants.com and mccormick@scientific-consultants.com.

Jeffrey Owen Katz, Ph.D.
Donna L. McCormick

CHAPTER 1

# A Review of Options Basics

This chapter provides some background by defining options and many of the terms used when discussing them. It looks at where options are traded, examines a few of their salient characteristics, and discusses some of the factors that influence their value. The reader is introduced to option price charts, which are used to illustrate the option terminology and influential factors under discussion. The fundamentals of pricing models, the Greeks, put-call parity, synthetics, and equivalent positions are also covered. More advanced readers may wish to skip this chapter.

## BASIC OPTIONS: CALLS AND PUTS

Options are contracts that come in two primary flavors: calls and puts. A *call* is an agreement that gives the holder (owner) the right to purchase the *underlying security* at a predetermined price called the *strike price* either at the call's *expiration date* (European-style options) or anytime during the life of the option (American-style options). If and when the holder of a call actually exploits that right, it is said that the call has been *exercised*. From the options seller's point of view, being short a call means having given someone the right to purchase from him or her the underlying security at the strike price of the option, regardless



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

of that security's current market value. If and when this right is exercised, the trader who is short the option is said to have been *assigned*. While a call gives its holder the right to buy the underlying security, a *put* gives the trader the right to sell the underlying security at the strike price of the option. A trader who is short a put has an obligation to purchase the underlying security at the strike price of the option, should the put be exercised and he or she be assigned.

Options are *derivative securities* because they owe their existence and value to the underlying assets on which they trade. In this book, the underlying assets of the options examined are mostly individual stocks, although market and sector tracking entities such as the OEX index and the QQQ tracking stock are also studied. In other words, the focus here is primarily on *equity* (*stock) options* and *index options*. It should be noted that most stock options are American-style options, which may be exercised at any time prior to expiration. Index options may be either European style or American style.

Years ago, options contracts were traded "over the counter" and were customized to the contract writer's requirements. Today, while some options are still traded the old way, most have been standardized, are issued by the Option Clearing Corporation (OCC), and are traded on regulated exchanges as *listed* options. A standard, listed option is fully defined by the underlying security, expiration, strike price, and type (call or put). Standard stock and index options generally expire on the Saturday following the third Friday of the month of expiration. Consequently, when specifying the expiration, only the month and, for longer-term options such as LEAPS (long-term equity anticipation securities), the year need be stated. Standard stock options may be exercised until 5 p.m. Eastern Standard Time on the last day of trading. This can be a source of a nasty surprise, should the trader be caught off guard due to options being exercised after the close of the market. If assigned, a trader who is short will receive an assignment notice the day following the assignment. Finally, each standard stock options contract usually controls 100 shares of stock. An option on IBM that expires in May and gives the owner the right to buy 100 shares at $75 a share would be referred to as an "IBM May 75 call"; an option on IBM that expires in June and gives the holder the right to sell 100 shares at $52 a share would be specified as an "IBM June 52 put."

Options are traded on a number of exchanges. The Chicago Board Options Exchange (CBOE) was the first listed options exchange in the world. Their Web site at www.cboe.com contains a wealth of information on options that includes quotes, historical data, option analysis tools (including pricing models), and educational materials. A newer, electronic options trading venue is the International Securities Exchange (ISE). It offers excellent liquidity on many options, as well as fast executions. The authors trade the QQQ index options (often referred to as "Qubes") on the ISE. Options also trade on the American Stock Exchange (AMEX) and on several regional exchanges such as the Philadelphia Stock Exchange (PHLX) and the Pacific Stock Exchange (PSE).

## NAKED AND COVERED

Options may be traded either *naked* or *covered*. A trader who sells a *covered call*, also known as a *covered write*, already owns the underlying security. For example, the owner of 100 shares of IBM might sell an IBM call, which entitles the buyer to "call away" the options seller's stock. A trader may also cover a short option position with a long position in another option on the same underlying security; the combined long-and-short position is known as a *spread*. Selling a *naked call* happens when a call is sold without owning the underlying stock. If the call happens to be exercised and the seller assigned, he or she will be obligated to sell the stock at the strike price of the option. This will result in a short position in the stock. The trader will probably want to quickly cover the short stock position and will have to do so by purchasing shares in the open market at a price higher than the strike price at which the short option position was established. A trader who sells a *naked put* is selling a put without being short the underlying security. If assigned, the seller is obligated to purchase the stock at the put's strike price, a price that is almost certain to be higher than the stock's current market value. Of course, if no assignment occurs, and the put or call expires worthless, the options seller gets to keep the entire premium he or she was paid for the option. Having options expire worthless is just what the seller of naked options usually hopes for, except when puts are sold as a means of acquiring stock at a lower cost basis.

## ADDITIONAL OPTION TERMINOLOGY

The price of an option is its *premium*. Premium can be broken down into two components or kinds of value. One kind of value is *intrinsic value*. If you have an IBM call with a strike price of $100 and the stock is trading at $105, the option will have an intrinsic value of $5. This intrinsic value derives from the fact that if you exercise the option, you can buy the stock at $100 from the option's seller, then immediately turn around and sell the stock for $105 in the open market, pocketing a $5 profit. Options also have another kind of value that has to do with where the stock might go at some point in the future. Assume that the option has several months of life remaining before expiration. Its total worth is almost certain to be greater than $5; for instance, it may be trading at $7. In this example, the extra $2 is the so-called *time value* or *time premium* of the option. Time value derives from what might happen in the future. At some future point in the option's life, IBM stock might reach $200, in which case the profit that could be made from holding the call option would be at least $100. On the other hand, the stock could drop to $20, leaving the option holder with a nearly worthless option. But, if the option has any time remaining, it will still have some value since, at some point prior to expiration, the stock could again surge to over $100, the option's strike price.

An option is said to be *in-the-money* to the extent that it possesses intrinsic value. For a call to be in-the-money, the underlying security must be trading at a price that is greater than the strike price of the call. In such a case, the call's intrinsic value is equal to the price of the underlying asset minus the strike price of the option. Conversely, a put is in-the-money when the underlying security trades at a price lower than the option's strike price. An in-the-money put has an intrinsic value equal to the put's strike price minus the price of the underlying asset. An option is said to be *out-of-the-money* when it possesses no intrinsic value, only time value. A call is out-of-the-money when the underlying trades below the call's strike price, while a put is out-of-the-money when the asset trades above the strike. When the strike price of the option lies near the price of the underlying security, the option is said to be *at-the-money*. At-the-money options tend to have the greatest amount of time value and are often the most actively traded. The term *moneyness* is sometimes used to refer to how far up the scale, from deeply out-of-the-money to deeply in-the-money, an option lies.

Deeply in-the-money options are flush with intrinsic value but often have little or no time value, especially when close to expiration. An in-the-money option is said to be trading *at parity* when its price reflects only intrinsic value. Sometimes deeply in-the-money options will actually trade at prices that are below their intrinsic value. When an option trades at a price less than its intrinsic value, the option is said to be trading *below parity*. European-style options that are in-the-money often trade below parity; because they cannot be exercised until expiration, the arbitrageur is subject to the risk of declining value, not to mention a cost of carry. American-style options normally trade *above parity*.

## FACTORS INFLUENCING OPTION PREMIUM

Many factors influence the price, or premium, of an option. Among the more influential and better-known factors are the price of the underlying security, the strike price of the option, the time remaining before expiration, the volatility of the underlying security, the dividend payout of that security (if any), and the risk-free interest rate. Lesser-known factors, many of which will be studied in subsequent chapters, include skew, kurtosis, serial correlation, and other statistical properties of price movements, as well as cycles, seasonal effects, and the impact of impending events such as earnings reports and criminal or civil judgments involving the companies behind the underlying securities. For reasons of clarity, it will be assumed in the discussion that follows that the underlying asset is a stock unless otherwise stated.

### Well-Known Factors

Two obvious variables that influence an option's price are the strike price of the option and the price of the underlying security.

As already mentioned when discussing moneyness, the intrinsic value of a call is zero when the stock trades below the strike, and is equal to the price of the stock minus the strike price of the option when it trades above. For a put, the relationship is reversed: the intrinsic value of a put is zero when the stock price is greater than the strike price, and the strike price minus the stock price when the stock is lower than the strike. Intrinsic value is not the only value affected by the price of the underlying security and the strike price of the option. Time value is affected as well. All things being equal, the closer an option's strike price is to the price of the stock, the greater the chance the stock will move sufficiently to give the option real or intrinsic value before expiration and, consequently, the greater the time value. At-the-money options tend to have the most time value, while deeply in-the-money or deeply out-of-the-money options have the least. These relationships should be familiar to anyone who has examined option tables in a newspaper, perhaps purchasing calls in an effort to profit from rising stock prices, or puts to profit from falling ones.

Volatility and time, working together, also have a major impact on an option's price. Let us first consider volatility. *Volatility* is a measure of how much a stock characteristically moves in a given unit of time, often stated as an annualized percentage. As every trader knows, stocks differ greatly in their volatility. A technology stock that is "in play" can easily have a volatility of over 200% annualized, while a volatility of 20% might be found in a quiet utility stock. Volatility that is calculated using historical price data for the underlying security is referred to as *historical volatility*. *Implied volatility* refers to the volatility that, when entered into an option pricing model, yields an estimated, theoretical price that is equal to an option's actual price; it is the volatility implied by the current price at which the option is trading.

As might be expected, the excess over intrinsic value that is known as time value is highly influenced by volatility. Higher volatility in an underlying security implies greater time value in any option on that security. Why? Because volatility is a measure of movement, and the more a stock moves during the life of an option, the greater the potential payoff. Time value is also determined by the period remaining before expiration: the longer the time remaining, the greater that component of an option's worth. If there is a longer period ahead, there is more time for the stock to potentially reach a price level that would yield a significant profit for the option holder. In a sense, time value might be referred to as *speculative value*. It is value that derives from potential movement in a stock's price as a result of volatility occurring within the framework of time. As expiration approaches, speculative value declines; this phenomenon is known as *time decay*. At expiration, when time has finally run out, an option will have only intrinsic value remaining; its speculative value will have dwindled to zero. It is for this reason that options are sometimes referred to as *wasting assets*.

Interest rates and stock dividends are also well-understood determiners of the worth of an option. Prices for calls increase with interest rates, while prices for puts decrease. A simplified understanding can be had by considering an arbitrageur who purchases calls and hedges them with short stock and short puts. The proceeds from the short side can be invested in bonds and earn interest at the prevailing rate. To the extent that interest rates are high, the arbitrageur, wanting to capture those interest earnings, will be willing to pay more for the calls and receive less for the puts. The same arbitrageur, short the stock, will be responsible for dividend payments. Higher dividends mean greater costs, and so the arbitrageur will pay less for calls and demand more for puts. Consequently, as dividends rise, prices decrease for calls and increase for puts. Dividends have an opposite influence to that of interest rates on option prices. In addition to their effects on arbitrage, interest rates and dividends determine the so-called *cost of carry* and the relative worth of competing investments, which, in turn, impact stock and option prices. Most traders are familiar with the bullish impact on the market of declining interest rates and of the drop in a stock's price when it goes ex-dividend.

## Lesser-Known Factors

Skew and kurtosis define certain aspects of the statistical distribution of returns, just as volatility does. While volatility measures the spread or width of the distribution, *skew* and *kurtosis* measure features of its shape. In the world of statistics, skew and kurtosis are the third and fourth *moment*s of a distribution. The first moment is the familiar *mean* or average, while the second moment is the *variance*. The square root of the variance is the *standard deviation*. It is the standard deviation that, in the language of options, is referred to as volatility.

First, consider skew. Compared to the familiar bell-shaped *normal distribution*, a distribution with *negative skew* is one that has an extended left (negative) tail and compressed right (positive) tail, with the peak appearing tilted to the right. A distribution with *positive skew* has an extended right tail, shortened left tail, and a peak that leans to the left. In the language of the trader, a stock or index with negative skew would be one that exhibits occasional sharp declines that are set against a background of frequent, but relatively small, price gains. A stock with positive skew evidences just the opposite behavior: occasional large gains and frequent, but comparatively small, declines. Obviously, skew affects the potential payoff of an option investment and so may be highly relevant to a trader attempting to estimate the current worth of an option.

Kurtosis, likewise, can be assumed to have a significant impact on option prices. A *platykurtic* distribution is one with negative kurtosis. It has shorter tails and a wider, flattened peak. A *leptokurtic* distribution is one with positive kurtosis and has a sharper peak and longer tails. As demonstrated later in this book, the distribution of returns seen in stocks and stock indices tends to be leptokurtic in that there are more extreme returns (both positive and negative) than would normally be expected, as well as more instances where prices change little, than there are moderate gains and losses. The effect of a leptokurtic distribution on option prices is that deeply out-of-the-money and deeply in-the-money options will have greater value than expected based on standard option pricing models. This is equivalent to observing a *volatility smile* when plotting implied volatility, calculated using a standard model and observed option prices, against moneyness. The so-called volatility smile has been the subject of many academic papers concerned with option pricing models.

So far, the discussion has focused on volatility, skew, and kurtosis, which are the second, third, and fourth moments, respectively. What about the first moment, the *mean* or *expectation*? A mean return that differs from zero appears on a stock chart as a *trend*—a period during which prices tend either to rise (positive mean return) or fall (negative mean return). Traders often attempt to profit from a discernible trend by assuming that the trend will have *momentum*, i.e., a tendency to persist, and by acting on that assumption. In an uptrend, calls might be expected to have increased value and puts decreased value, because of the potential for continued directional movement in the underlying stock. The reverse might be expected in a downtrend. Such expectations, however, may not reflect market reality since conversion and reversal arbitrage (which will be discussed later) acts to attenuate the more obvious differential effects of trend on put and call premiums. Nevertheless, trends can be expected to have a significant influence on option values.

Finally, there are cycles, seasonal effects, and events that are expected to generate news at some point in the future. Cycles, it should be noted, can be observed not only in stock and index prices, but also in their volatility. For obvious reasons, a stock's volatility might be expected to rise when earnings reports are released. Since earnings reports are generally made available on a quarterly basis, there should be a discernible cycle in volatility, and, consequently, in option premiums, having a period of around 90 calendar days. In addition to the earnings cycle, there are other seasonal market phenomena. For whatever reason, certain months are notorious for high volatility, while other seasonal periods often exhibit exaggerated trends. Most traders are familiar with October as a month of crashes and market volatility, and with late December through early January as a frequently bullish period. These seasonal tendencies in volatility and trend should be reflected in, and perhaps even be anticipated by, the options markets. Because cycles and seasonal effects are easily quantified, incorporating these factors into a practical pricing model should not be difficult. News is another factor that influences option premiums. For one thing, news is a significant source of volatility. Consequently, rumors and impending news (e.g., regarding the outcome of a lawsuit, a drug trial, or a potential takeover), with the implication of market volatility to come, can cause implied volatility to skyrocket and options to gain greatly in price, even when the underlying stock remains relatively unchanged in price and exhibits normal, uninflated historical volatility. However, once the news becomes public knowledge, inflated option prices often collapse.

## USES OF OPTIONS

Options are versatile instruments that may be used in innumerable ways. One simple and familiar use is to speculate on directional movement in the underlying, such as when a trader buys calls to profit from an anticipated rise in a stock's price, or puts to benefit from a fall. When used for this purpose, options can provide leverage while controlling certain kinds of risk. However, as should be evident from the discussion of factors influencing option premiums, options go far beyond simply providing leverage or limiting risk. For one thing, options make it possible for the trader to speculate on variables other than price. Consider a trader expecting increased volatility, but unable to predict its direction. There are option strategies that will respond profitably to the anticipated change in volatility while minimizing the impact of any directional movement. Even the trader who expects stock prices to remain in a narrow trading range can find an option strategy to take advantage of the situation. Options are the perfect instruments for turning such expectations into potentially profitable actions. They are also the instruments of choice for transferring certain kinds of risk, for hedging, and for custom-tailoring the risk-reward characteristics of a variety of investments.

## OPTION PRICING MODELS

In the above discussion, repeated references have been made to option pricing models. An *option pricing model*, if not already clear from the context, is a mathematical algorithm or formula by which the theoretical fair value of an option may be calculated. Such a model, naturally, bases its calculations on certain assumptions regarding the nature of fair value, the behavior of price movements in the underlying security, and the effects of a variety of factors that are known to influence option prices.

An option pricing model is necessary, not only for estimating fair value, but also for calculating several other useful items. One such item, implied volatility, has already been discussed. *Implied volatility* is the volatility that, when entered into an option pricing model, produces a fair price estimate that matches the price at which the option is actually trading; it is the volatility being implied by option prices. Other useful products of an option pricing model include the so-called Greeks: Delta, Gamma, Theta, Vega (also known as Tau), and Rho.

## The Greeks

To one versed in calculus, the Greeks are just the partial derivatives of the pricing model's output (theoretical fair premium) with respect to its inputs. However, one need not be fluent in calculus to understand the Greeks.

First, consider Delta. *Delta*, also known as the *hedge ratio*, simply measures how much an option's premium is likely to change in response to a small change in the price of the underlying stock. For example, if a call option has a Delta of 0.40, and a share of the underlying stock moves up $0.80, the option can be expected to gain $0.32 (0.40 multiplied by 0.80) per share controlled (a standard options contract controls 100 shares). Since rising stock prices imply rising call prices and falling put prices, calls have positive Delta and puts have negative Delta. As an option goes deeply in-the-money, its Delta approaches +1.0 (call) or –1.0 (put). At the opposite end of the spectrum, a far out-of-the-money option will have a Delta approaching zero.

Options traders sometimes refer to a position as being Delta-neutral. Given the definition of Delta, it should be clear that a *Delta-neutral* position is one that is relatively unresponsive to small changes in the price of the underlying security. An example would be an at-the-money straddle. A *straddle* consists of a put and a call, both having the same strike and expiration, on the same underlying security. With a properly chosen strike, the negative Delta of the put will cancel out the positive Delta of the call. Since Delta changes with moneyness, a straddle will lose its Delta-neutral character, should stock prices move sufficiently away from the strike.

*Gamma* is the rate at which Delta changes with movement in the underlying stock. It is useful when constructing Delta-neutral hedges involving multiple options that will remain Delta-neutral over a wider range of stock price. Such hedges require positions designed to minimize not only Delta, but also Gamma. Positions that minimize Gamma are sometimes referred to as *Gamma-neutral*. Of course, as time passes and factors influencing option premium undergo change, adjustments to the positions will be required in order to maintain a Delta-neutral or Gamma-neutral stance.

*Theta* is the rate at which time decay erodes value. An option with a Theta of –0.05 loses 0.05 points a day to time decay. This works out to a loss of $5 a day per contract, given that a standard contract controls 100 shares of stock. Obviously, Theta is the option buyer's enemy and the seller's friend. Theta is highest for at-the-money options having significant time value and, for these options, increases rapidly with the approach of expiration.

*Vega* measures an option's sensitivity to changes in volatility. If volatility increases 10% and the option gains $0.20 per share, then Vega is approximately 0.02 (0.20 divided by 10). Longer-term options have more sensitivity to volatility and so possess higher levels of Vega. Just as Delta-neutral positions can be implemented to hedge price risk, so can *Vega-neutral* strategies be employed to hedge volatility risk. Since volatility is one of the most influential factors affecting option price, Vega is of great interest to hedgers. It is also of interest to the trader attempting to speculate on changes in volatility, something for which options are an ideal instrument.

Finally, *Rho* measures sensitivity of option prices to interest rates. Since call premiums increase with interest rates while put premiums decrease, Rho is positive for calls and negative for puts. For both puts and calls, the absolute value (negative or positive) of Rho increases with increasing time.

## Black-Scholes

Black-Scholes is undoubtedly the most popular option pricing model in use today. It is readily available in the form of spreadsheet add-ins, stand-alone programs, and as part of many online trading platforms. Numerous Web sites offer Black-Scholes calculators. Needless to say, the Black-Scholes model is used by almost every serious trader of options, even those having access to newer and more complex pricing models. The reason is that, although not perfect by any means, Black-Scholes is familiar, easily understood, and does provide reasonable estimates of an option's worth under normal conditions. In addition, it is very easy to calculate implied volatility and the Greeks using the Black-Scholes formula. Black-Scholes uses, for its inputs, the well-known factors that influence an option's value: the stock's price, the option's strike price, the volatility, the time remaining before expiration, the risk-free interest rate, and the dividend rate. The output, of course, is a theoretical fair price consonant with the model's assumptions.

Roughly speaking, Black-Scholes assumes that the price behavior of the underlying security is accurately described as a geometric random walk having a log-normal distribution of returns, that the options markets are perfectly liquid and efficient, and that arbitrage can always be carried out. In its standard form, it also assumes European-style options and continuous interest and dividend payouts. As with most theoretical models, it disregards the realities of transaction costs, wide ask-bid spreads, and many other disturbing "minutiae" familiar to most traders. What trader has not observed stale last price figures, options that have not traded for days at a time, and other telltale signs of extremely poor efficiency and liquidity in the options markets? Yet, despite obvious violations of at least some of the model's assumptions, Black-Scholes generally provides useful, if not surprisingly good, approximations to fair value.

It should be mentioned that Black-Scholes is not the only popular option pricing model. The Cox-Ross-Rubinstein "binomial" model also has a large following. Although computationally more involved and time-consuming than Black-Scholes, this model may actually be easier to understand. Cox-Ross-Rubinstein is also more flexible in its application. For example, it can easily be adapted to handle discrete dividends and American-style options. Because of this flexibility, Cox-Ross-Rubinstein might yield somewhat more accurate fair value estimates in many circumstances.

Finally, the correctness of prices generated by a model like Black-Scholes is very dependent on the accuracy of the inputs. Some inputs, like time left to expiration, stock price, strike price, and interest rate, are either known precisely or change so slowly that they are easily evaluated with relative precision, at least for short-term options. These inputs present no problem. Other inputs, such as volatility, do present problems in that they are difficult to precisely determine and so must be estimated. It is with such inputs and their estimates that serious inaccuracies may enter the picture. Consider the fact that historical volatility is often used as a proxy for volatility when computing Black-Scholes when what is strictly required is future volatility. Historical volatility may sometimes provide an acceptable estimate of future volatility, but not always. There are specific circumstances under which historical volatility affords a very poor approximation to future volatility and one that will lead to serious pricing errors. To avoid the need for an independent estimate of future volatility, implied volatility is occasionally used to calculate fair premium, but this only permits comparisons amongst options. When using implied volatility, it is impossible to say whether any option on a given security is overpriced or underpriced independent of others. The bottom line is that for a model like Black-Scholes—a model with inputs that have values that can only be estimated—any effort spent on improving the accuracy of the input estimates will pay off in increased precision in the theoretical fair prices the model generates.

Exhaustive coverage is given in later chapters to random walks, distributions (including the log-normal), issues concerned with estimating volatility and other relevant input variables, the assumptions made by popular pricing models, and the conditions under which the approximations to fair value that such models generate tend to break down.

### Why Use a Pricing Model?

There are many reasons to use an option pricing model. One obvious reason is that an option model enables the calculation of a theoretical fair price. A theoretical fair price provides the trader with a rational basis that can be used to judge whether actual prices in the marketplace are reasonable, and to determine where to place bids and offers. Another good reason to use an option pricing model is that it makes possible the calculation of implied volatility. Estimates of implied volatility allow meaningful comparisons to be made amongst options with differing strikes and expirations. Relatively expensive and perhaps overpriced options will evidence higher implied volatility than those that are cheap. A trader can gain a statistical edge in the market by taking advantage of significant price discrepancies revealed by comparisons of implied volatility. In addition, high levels of implied volatility sometimes reflect insider activity and can alert an astute player to a potential earnings surprise or takeover bid. For an excellent, in-depth discussion of these phenomena, consult *McMillan on Options* (McMillan, 1996). Needless to say, an option pricing model makes it straightforward to evaluate various strategies and to estimate their payoffs under different market scenarios. With the aid of a pricing model, charts can be prepared that show the theoretical response of a position to various influences such as interest rate, volatility, time, and stock price. This is true whether the trader is contemplating a simple option position, such as a long or short put or call, or a complex multiple option position like a straddle, spread, or butterfly. Finally, the Greeks, calculated with the help of a pricing model, are essential when trying to hedge specific kinds of risk. A market maker would not survive for long if, at the very least, he or she did not hedge against the risk of significant price movement in the underlying security by establishing Delta-neutral positions. In short, trading options without a good pricing model is like flying in fog without instruments. Over the long haul, traders who intelligently employ good pricing models will profit at the expense of those who do not.

## GRAPHIC ILLUSTRATIONS

A variety of response charts was prepared using the standard Black-Scholes pricing model. These charts show how options theoretically respond to the various factors discussed earlier, such as life remaining in the option, volatility, and moneyness.

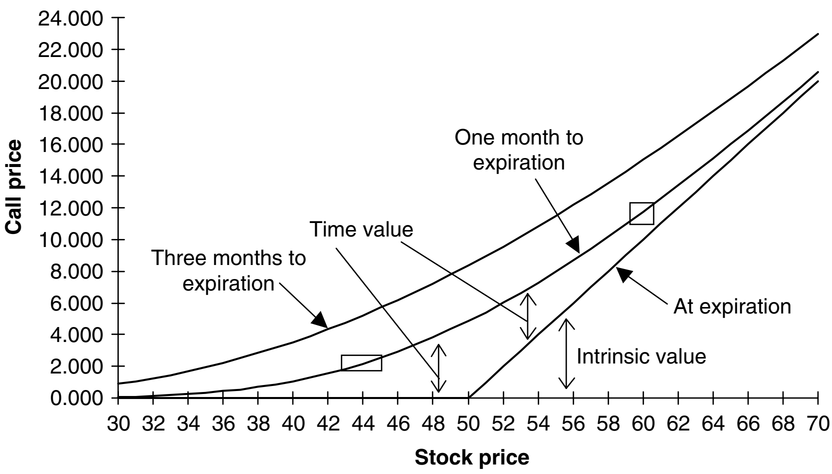
Response charts are useful for visualizing the impact of various factors on an option's premium and the related Greeks.

Figure 1–1 depicts the relationship between call option prices, stock prices, and time remaining before expiration. Annotations on the chart delineate the two kinds of value that make up an option's premium: intrinsic value and time value. Delta is illustrated by the small rectangles drawn on the middle curve. When generating data for Figure 1–1, interest and dividend rates were assumed to be zero, volatility was fixed at 85%, and the strike price was set to $50.

As shown in Figure 1–1, call option prices almost always rise with increasing stock prices. The one instance where prices do not rise is with an expiring call that is out-of-the-money. An out-of-the-money option with no time remaining has no value. In Figure 1–1, such an option is represented by the lowermost curve when stock prices are less than $50. For stock prices less than $50, the lowermost curve lies on the *x*-axis, representing a premium of zero. The same curve shows a premium that increases one-for-one with stock price when stock prices are greater than $50 and the call is in-the-money.

**FIGURE 1–1**

Influence on Call Value of Stock Price and Time to Expiration



When there is time remaining prior to expiration, call prices rise in a smooth but accelerating fashion with stock prices. This can be seen in the curves for options with one and three months of life remaining. When stock prices are low relative to a call's strike price, Delta is low and the option price curve ascends leisurely. As stock prices rise, Delta increases and the curve begins to climb more steeply, revealing an accelerating growth in option price. In general, the further out-of-the-money an option becomes, the closer its price and its Delta approach zero. At the other extreme, as an option becomes more and more in-the-money, its price begins to rise or fall one-for-one with the stock and its Delta approaches positive or negative unity. Either way, any curve that represents the theoretical fair premium of an option prior to expiration will be seen to asymptotically approach the curve that describes the option at expiration as stock prices continue to rise or fall, at least when interest and dividends are ignored.

Delta, as mentioned earlier, is the slope of the curve relating an option's theoretical price to the price of the underlying security. The small rectangles drawn on the middle curve illustrate the idea of Delta. Delta can be understood and roughly approximated as the ratio of such a rectangle's height, measured in option price units, to its width, measured in stock price units. The approximation becomes precise to the extent that the rectangle is small in relationship to the curvature of the price response. Delta is positive when the slope of the response curve is upward, normal for call options, and negative when it is downward, normal for put options. As evident in Figure 1–1, Delta changes with both stock price and time remaining before expiration.

Time value appears in Figure 1–1 as the vertical distance between the curve for the option that is about to expire and either of the other curves. From this point of view, time value is measured as the price of an option with time remaining, and hence possessing both time and intrinsic value, minus the price of an option with the same strike that is about to expire, possessing only intrinsic value. In Figure 1–1, it can be seen that time value reaches a maximum when the stock is around \$50 and, consequently, the options are at-the-money. At-the-money options generally have the greatest amount of time premium. Time value decreases as options move further in- or out-of-the-money. This appears in Figure 1–1 as a narrowing of the distance between the curves as stock prices move away from $50, the strike price of the options. Intrinsic value is the vertical distance between the price curve for the expiring option and the $x$-axis, which crosses the $y$-axis at zero.

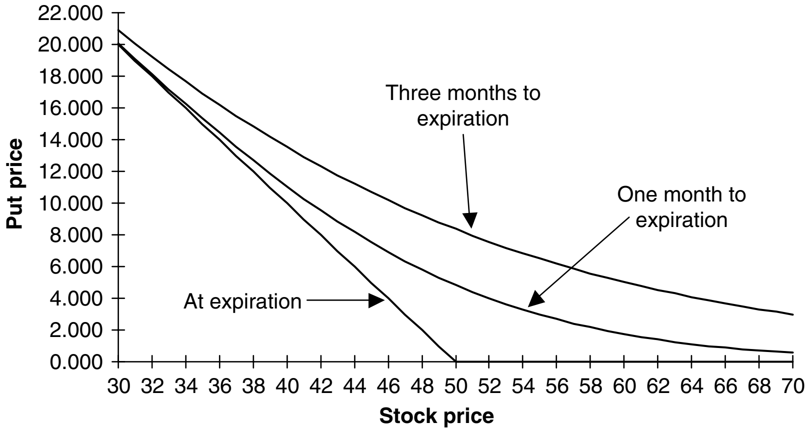
The data in Figure 1–1 demonstrate how the decay of time value can turn an option trade in which the underlying stock performs favorably into a loser. For example, had a trader purchased the three-month call (uppermost curve) with the stock at $40, it would have cost about $3.55. Assume that the stock moves up to $45 in a period of two months. The option now has one month remaining (middle curve) and is worth only about $2.55. Despite the stock gaining in price, the trader has suffered a loss. Perhaps the stock continues to rise, ultimately reaching $49 after the passage of another month. In that case, the expiring option is worthless (lowermost curve) and the trade is a total loss despite the stock's substantial appreciation.

Time erosion can be very damaging when options are purchased and held for even modest periods. However, the effect of time decay may not be very significant for the very short-term trader. For this kind of trader, there may be substantial benefits to purchasing options rather than the underlying stocks. One easily observed benefit is that an option reduces the impact of adverse movement in the price of a stock or index at the same time that it magnifies favorable movement. As can be seen in Figure 1–1, a move down causes less of a loss to the holder of a call than the same move up results in gain. This is especially noticeable with at-the-money options near expiration, when very little speculative value remains and Delta varies rapidly with stock price.

Figure 1–2 examines the behavior of put options exactly as Figure 1–1 examined the behavior of calls. As with Figure 1–1, interest and dividend rates were assumed to be zero, volatility was fixed at 85%, and the strike price was set to $50. It is readily apparent that Figure 1–2 looks like Figure 1–1 flipped left to right. This is because stock prices affect put premiums opposite to the way they affect call premiums. As stock prices fall, put

**FIGURE 1–2**

Influence on Put Value of Stock Price and Time to Expiration

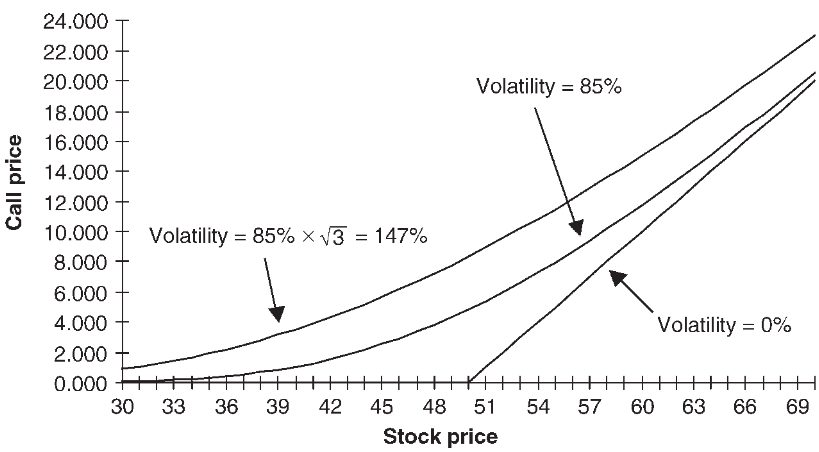



options move into the money and their premiums rise. As stock prices rise, these options move out-of-the-money and lose premium. This is visible in the way the price curves ascend as one scans from the right side of the chart to the left. When viewed in terms of moneyness, both calls and puts behave similarly: the less out-of-the-money, or more in-the-money, the greater the option premium. In Figure 1–2, the options are in-the-money for stock prices less than \$50. At all points, the response curves are either declining or flat as the chart is scanned to the right, illustrating the fact that Delta is always negative or zero for put options. As with calls, intrinsic value appears as the vertical distance between the $x$-axis, at zero, and the curve for the expiring option. The put at expiration (lowermost curve) has no value for stock prices greater than \$50, but has a value that rises one-for-one with the stock price as it drops below \$50. The vertical distance between the curve for the option at expiration and the curves for those with more life remaining is again a measure of time value. Time value is greater for the put with three months of remaining life than for the put with one month left to expiration. Finally, just as the calls in Figure 1–1 cushioned the trader against a decline in stock price and amplified an incline, puts amplify the profit from a decline in stock price, while cushioning the holder against a sudden rise. For a short-term or day trader, the purchase of a put is a good way to speculate on an anticipated decline in a stock's price, or to cash in on a crash.

Figure 1–3 shows the impact of volatility on call option premiums. Again, interest and dividend rates have been set to zero and the strike price to $50. In this instance, the time remaining until expiration has been fixed at one month, and the three curves characterize differing levels of volatility. The attentive reader will note the similarity to Figure 1–1. This similarity is not coincidental; without interest and dividends, time and volatility have precisely the same effect on option premium. This is because what influences premium is expected movement in the price of the underlying. For a random walk, the amount of movement expected can be calculated as volatility multiplied by the square root of time. Quadruple the time and you have exactly the same effect on expected movement, and thus option premium, as you would have had you left time unchanged but doubled volatility. Naturally, when interest rates and dividends are included in the model, time will have additional effects on option premium not shared by volatility and not resulting from differing expectations regarding movements in the underlying.

**FIGURE 1–3**

Effect of Volatility and Stock Price on Call Value



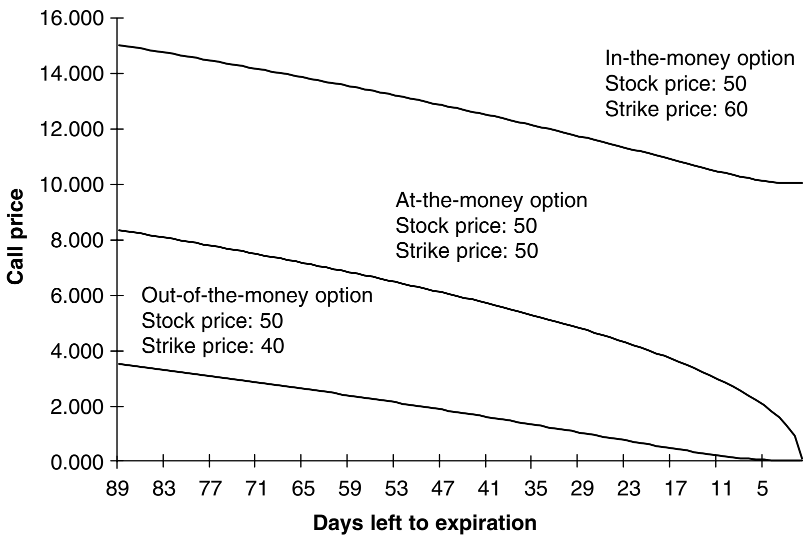
Nevertheless, interest rates and dividends generally have little impact on short-term options. They become much more significant for longer-term options, especially LEAPS, which may have one or more years of life remaining.

The effect of volatility on option premium depicted in Figure 1–3 can be taken advantage of by purchasing options when volatility is expected to increase and selling options when it is expected to decrease. This can be done in the context of a Delta-neutral strategy when it is desirable to minimize directional price risk.

Figure 1–4 demonstrates the effect of time on option premiums and illustrates Theta. In preparing the data for Figure 1–4, interest and dividend rates were set to zero, the strike price was fixed at $50 and volatility was assumed to be 85%. Only call options are shown because, when interest rates and dividends are zero, the influence of time on put options is virtually identical to its effect on call options.

**FIGURE 1–4**

Effect of Time Decay on Option Prices



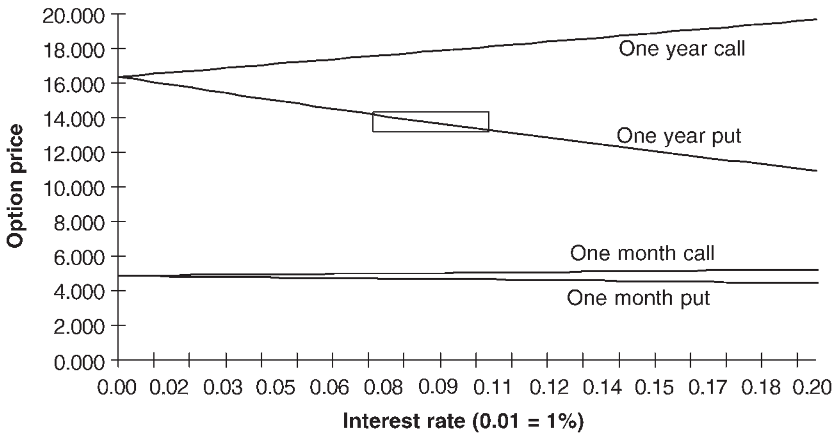
For the out-of-the-money call (lowermost curve), the loss of premium to time decay is gradual, with prices dropping below \$0.05 (the standard option tick size) when about 7 days of life remain. The rate at which premium is lost, the option's Theta, is about –\$0.036 per day near the left end of the chart when there are 87 days to expiration. Although not readily visible to the naked eye, the rate of decline accelerates slightly, reaching a peak when the option has 24 days of life remaining. At that point, Theta is –0.048, indicating a loss of \$0.048 per day to time decay. Although the proportion of remaining time value that is lost each day increases, since there is less value left to lose every day, the absolute loss per day declines. With 10 days left to expiration, the option is worth only about \$0.15 and is losing about \$0.037 per day to time decay. The behavior of the in-the-money option (uppermost curve) is similar to that of the out-of-the-money option, except that all premiums have been shifted up by \$10 as a result of the option having \$10 worth of intrinsic value. Time value is lost at a rate of about \$0.046 per day near the left end of the chart. With the in-the-money option, the rate of decay reaches a peak of \$0.072 per day at 16 days and declines thereafter.

For an at-the-money option (middle curve), the picture is different. Initially, the loss of premium to time decay is gradual and not unlike what was seen with the other options. Near the left end of the chart, when much time remains, Theta is about –0.046 for the at-the-money option. As the days go by, however, the erosion of premium accelerates steadily, reaching a maximum right at the end. In fact, nearly \$0.80 is lost just in the last day! Because at-the-money options can easily become in-the-money over a very short interval, there will be substantial amounts of time value remaining right until the very end. Substantial time value with little time left means abundant time decay.

Figure 1–5 shows the relationship of call and put prices to time and interest rates, and illustrates Rho. As can be seen, call prices increase with interest rates while put prices decrease. For both calls and puts, longer-term options are more affected by interest rates than shorter-term ones. Rho is illustrated by the small rectangle drawn on one of the curves, in the same way that Delta was illustrated in Figure 1–1. Rho is approximated by

**FIGURE 1–5**

Option Prices as a Function of Interest Rates



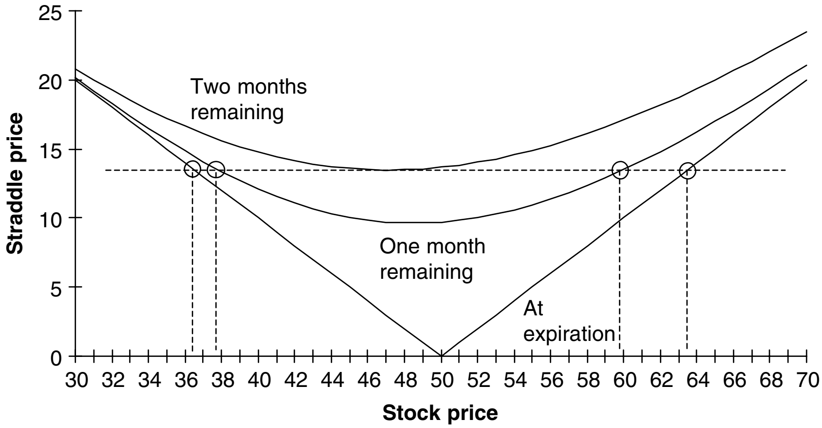

the height of the rectangle divided by the width, the approximation becoming exact in the limit as the size of the rectangle is reduced. The sign of Rho is positive for rising curves and negative for falling ones. Note that the effect of interest rates on option prices shown in Figure 1–5 appears to be linear, i.e., the curves describing the effect seem to be well-approximated by straight lines. This is in stark contrast to time, volatility, and stock price, all of which have more complex, curvilinear effects on option prices.

Graphic representations of option behavior, such as those illustrated in the figures, can help the options trader visualize the characteristics not only of single options, but also of complex strategies involving multiple options and, perhaps, the underlying stock as well. As an example, consider the next chart, which is for a straddle. A straddle—a position involving an at-the-money put and call having the same strike and expiration—was used earlier as an example of a Delta-neutral position.

Figure 1–6 shows how the premium of a typical straddle responds to stock price (the $x$-axis) and time (the three curves). This figure is similar in construction to Figure 1–1; the difference is that instead of depicting the behavior of individual options,

**FIGURE 1–6**

Straddle Price as a Function of Stock Price and Time Left




it reveals the behavior of a multiple option position. Figure 1–6 was generated under the assumption that both the call and put have a strike price of \$50, that the volatility is 85%, and that interest rates and dividends are zero.

The lowermost curve in Figure 1–6, which represents the straddle at expiration, is V-shaped. Prices rise one-for-one with the difference between the stock price and the straddle's strike price. This lowermost curve represents the intrinsic value of the straddle. When stock prices are above \$50, the Delta of the position is one. In this case, the put has no intrinsic value, but the call has an intrinsic value equal to the stock price minus the strike. When stock prices are below \$50, the Delta of the straddle is minus one, the call has no intrinsic value, and the put has an intrinsic value equal to the strike minus the stock price. Delta is undefined when the stock price is equal to the strike price, and because neither option has value, nor does the straddle. The middle curve illustrates a straddle with one month left to expiration. Here, the presence of time value is in evidence as the vertical distance from the lowermost curve (the straddle at expiration) to the middle curve. The uppermost curve, which represents the straddle with three months left to expiration, shows even greater amounts of time value.

In contrast to the sharp, V-shaped curve, the two other curves are bowl-shaped. The bottoms of all three curves occur at or slightly below a stock price of $50. At prices slightly below $50, the uppermost two curves reach bottom and become horizontal, indicating there is little change in option price in response to small changes in stock price. In other words, when the stock is appropriately priced relative to the strike, a straddle with time remaining has a near-zero Delta. Such a straddle represents a Delta-neutral position. As stock prices drop well below $50, Delta becomes increasingly negative as the two uppermost curves rise at an accelerating rate, asymptotically approaching each other and the lowermost curve. As stock prices rise above $50, the curves also accelerate upward, asymptotically approaching each other and the lowermost curve, while Delta becomes increasingly positive, reaching unity in the limit.

Because two options are losing time value, the stock must move significantly if a profit sufficient to compensate for the loss is to be made. For instance, had the straddle been purchased with the stock at $47 and three months of life remaining in the options, it would have cost $13.48. Two months later, with only one month of life remaining in the straddle, the stock would have had to move either below $37.85 or above $59.98 for the straddle to have yielded a profit. This can be seen by drawing horizontal lines from the bottom of the uppermost (three-month) curve to where they intersect with the middle (one-month) curve, and then determining the stock prices at the points of intersection. To be profitable at expiration, the stock would have had to move even further, the position taking a loss for any stock price between $36.52 and $63.48. Again, the breakeven points can be located by extending horizontal lines from the bottom of the uppermost curve, which represents the entry point, to where they intersect the lowermost curve at the breakeven points. The asymmetry observed in the breakeven points is due to the geometric (proportional) rather than arithmetic scaling of stock price movements. A stock that drops to half its original price has made a move having essentially the same magnitude and probability as a stock that doubles, although the arithmetic or dollar value of the downside move is less than that of the upside move.

Figure 1–6 demonstrates the value of generating families of response curves, as was done in these charts. With charts like those shown above, a trader can easily visualize alternative positions and how they are likely to perform under various market scenarios. A model, like Black-Scholes, which can be implemented in a spreadsheet or other easy-to-use software application, makes the preparation of such charts a simple and routine process.

## PUT-CALL PARITY, CONVERSIONS, AND REVERSALS

A specific relationship, know as *put-call parity*, exists between puts and calls of the same strike and expiration. For the sake of simplicity, let us assume that interest rates and dividends are zero. Ignoring interest rates and dividends, the put-call parity relationship can be precisely expressed by the following equation:

$$C - P = U - S \tag{1.1}$$

In this formula, $C$ is the call price, $P$ is the put price, $U$ is the price of the underlying stock, and $S$ is the strike price of the options. The formula basically says that when both options have the same strike and expiration, the difference between the price of a call and the price of a put must equal the difference between their intrinsic values, and, consequently, that the time value of the call must equal the time value of the put.

The relationship between call and put prices expressed in the formula derives from the following facts. At expiration, options have only intrinsic value. Thus, at expiration, a position consisting of a long call and short put with similar terms will have a premium of zero when the stock price is equal to the strike price, a premium that rises one-for-one with stock price when the stock price is above the strike, and a value that declines one-for-one with stock price when the stock price is below the strike. This follows because, for a stock price above the strike price, the long expiring call rises one-for-one with the stock, while the put remains valueless; for stock prices below the strike, the short expiring put rises one-for-one with declines in the stock, while the call premium remains at zero. Analyzed in this way, it is easy to understand why put-call parity must always exist at expiration.

To the extent that the market is efficient, however, this simple relationship between the prices of puts and calls with the same terms will be enforced, not just at expiration, but at all times prior to expiration. The enforcing agent is arbitrage, and deviation from put-call parity at any point in time represents an inefficiency that can be exploited for profit and driven from the market. To see how this works, imagine that the calls become relatively more expensive than the puts in terms of the stated relationship. In such a circumstance, an arbitrageur can sell the expensive calls and purchase the relatively cheaper puts, creating a synthetic short stock. This synthetic stock is represented by the part of the equation that lies to the left of the equal sign. The arbitrageur can then purchase the actual underlying stock to offset the synthetic short stock for a total position cost that is less than the strike price of the options. A profit will result because at expiration put-call parity will reassert itself, and the entire position will have a value precisely equal to the strike price of the options. Of course, to the degree that such arbitrage is carried out, exploitable price discrepancies will be eliminated. The selling of relatively expensive options will cause their prices to fall while the buying of relatively inexpensive options will cause their prices to rise. Put-call parity will thus be reestablished. The form of arbitrage just described is known as *conversion arbitrage*.

The converse, *reversal arbitrage*, involves the purchase of underpriced calls and the sale of the corresponding overpriced puts to establish a synthetic long stock, which is then offset by shorting the actual underlying stock. The rationale is the same as for conversion arbitrage, except that the individual option and stock positions are all reversed. Again, the effect is to maintain put-call parity.

Note that the put-call parity relationship will tend to prevail regardless of the expected (or actual) distribution of future stock price movements. To understand why this is so, imagine that options traders become very bullish and begin to aggressively purchase calls. Naively, it might be assumed that such demand will drive up call premiums while put premiums fall or remain unchanged. However, this will not happen because, as the calls are bid up, arbitrageurs will sell the calls and purchase the corresponding puts, forcing the options back into parity. As a result of conversion arbitrage, demand for calls will translate into demand for puts, and premiums (as well as implied volatilities) are likely to increase for both kinds of options.

## SYNTHETICS AND EQUIVALENT POSITIONS

In studying options, the reader is likely to encounter references to synthetics and equivalent positions. A *synthetic* or *synthetic position* is a position in one security that is simulated using a combination of positions in other securities. As an example, consider a position consisting of a long call and a short put, both having the same strike and expiration. Such a position constitutes a *synthetic stock*. It possesses profit and loss characteristics similar to those of an actual stock position. Synthetics are important to the arbitrageur since they can become overvalued or undervalued relative to the actual securities they mimic, thus providing arbitrage opportunities. As discussed earlier, conversion and reversal arbitrage is one form of arbitrage in which synthetics play a role; however, synthetics are useful not just to arbitrageurs. Sometimes it is simply easier or more advantageous to establish a synthetic position than an actual one. A trader can establish a synthetic short stock position, for instance, without having to wait for an uptick, as would be required (by law) in order to establish a short position in the stock itself; an uptick that might never come in a swiftly falling market. In other instances, a synthetic position might be established at a better price or permit greater use of margin. Finally, in the world of futures, synthetics constructed using options can help the trader cope with limit-locked markets.

An *equivalent position* is a position having profit and loss characteristics similar to those of another, different position. Consider the covered call, or covered write, as it is also known. Selling covered calls is a popular strategy that many investors regard as very conservative. However, assuming the strike and expiration are the same, being long a stock and short a call is equivalent to being short a naked put. The short, naked put has the same payoff characteristics and profit graph as does the long stock and short call. They are, in fact, equivalent positions. Moreover, in many cases it is more cost effective to sell naked puts than to establish covered writes! However, selling naked puts, strangely, is not generally perceived as a conservative strategy. Obviously, many investors have an inadequate understanding of equivalent positions.

There is a close relationship between equivalent positions and synthetics. The reader may have noticed in the above example that a covered write is actually a synthetic short put. A short put, however, would not normally be described as a synthetic covered write. In general, when one of a pair of equivalent positions involves a single security, the other of the pair is typically understood to be a synthetic version of that security. Synthetics originated in the early days of options trading. In those days, there were no puts, only calls. However, a trader could short a stock and buy a call, thereby establishing a position with the characteristics of a put. These equivalent positions became known as *synthetics*.

There are many equivalent positions and synthetics with which any savvy trader or investor should gain familiarity. A few of these are listed below. Consult *Options as a Strategic Investment* (McMillan, 1993) for a more extensive list and discussion of strategies.

| **Position** | **Equivalent Position** |
|---|---|
| Long stock + long put | Long call |
| Short stock + long call | Long put |
| Long stock + short call | Short put |
| Short stock + short put | Short call |
| Long call + short put | Long stock |
| Long put + short call | Short stock |
| Long stock + 2 long puts | Long straddle (long put + long call) |

By definition, equivalent positions possess similar profit graphs. However, they generally do not have the same capital and margin requirements. This is also true for synthetics. A synthetic will usually have different capital and margin requirements than the security it mimics. Tax consequences may also differ between otherwise equivalent positions. In estimating relative returns and choosing an optimal strategy, a trader must take tax, capital, and margin issues into consideration.

## SUMMARY

In this chapter, options were described as contracts providing certain rights and obligations. Two basic types of options contracts were discussed: the put and the call. A standardized option was fully specified by its type (put or call), strike price, and expiration date. Many of the terms used in options discourse were defined: naked versus covered, in-the-money versus out-of-the-money, intrinsic value versus time value, and so on. Standardized options were observed to trade on a number of exchanges, both traditional (AMEX, PSE, CBOE) and fully electronic (ISE). It was pointed out that options contracts, like stocks and futures, could be traded from either the long or the short side, hedged (covered) or unhedged (naked).

Options were observed to have value, reflected in their price, said price also being known as *premium*. Option premium was shown to derive from two kinds of value: intrinsic value and time value. Option premium was found to be influenced by well-known factors such as stock price and strike price, time, volatility, interest and dividends, and by lesser-known factors such as skew, kurtosis, trend, and cycles. Estimating a fair premium, based on such factors, was recognized as a problem for which an option pricing model was the solution. An option pricing model was characterized as a mathematical formula or algorithm that enables its user to calculate a theoretical fair price. An option pricing model, it was noted, not only provides fair value estimates, but also makes it possible to calculate the so-called Greeks. It was demonstrated that the Greeks are not difficult to understand, even for a student not versed in higher mathematics, and that they measure important characteristics of option price behavior.

Some of the uses of options, and how they can benefit traders and investors, were then examined. Option profit graphs or response charts were illustrated and found to be useful in gaining a visual understanding of how single options and multi-option positions respond to stock price, volatility, time, and other factors. Attention then turned to a specific relationship known as *put-call parity* that exists between the premiums of puts and calls with the same terms, i.e., with the same strike and expiration. Put-call parity was seen to be maintained by conversion and reversal arbitrage. Finally, synthetics and equivalent positions were discussed. Synthetics were described as positions involving several securities (such as options) that simulate the profit and loss characteristics of another security (e.g., a stock). Equivalent positions were defined as distinct positions that, nevertheless, have the same profit and loss behavior and, consequently, the same profit graphs.

The goal of this chapter has been to provide the reader with a working knowledge of the fundamentals and jargon of options. Much of what follows in this book assumes knowledge of such option basics.

One fundamental concept of great importance to options traders, arbitrageurs, and investors, is the notion of *fair value*. References to fair value and fair price were made frequently in this chapter, although the meaning of these terms was never precisely defined. It was merely suggested that such a thing as fair value or fair price exists, and that option pricing models were how it could be reckoned. The designations were left to be interpreted within the framework of ideas regarding pricing that all traders possess. All traders have some idea of what it means for a security to be overpriced or underpriced. Every investor looks to buy bargains and attempts to sell securities that have become overvalued. Most readers know that profits can be made buying underpriced securities and selling overpriced ones, i.e., buying those priced below fair value and selling those priced above it. In the world of options, some very precise notions of what constitutes fair price or value can be developed. The next chapter explores fair value in much greater depth and attempts to elucidate more fully its nature and appraisal. It also explores the intimate connection between fair value and market efficiency. Fair value and its estimation, of course, is the subject of this book. A good understanding of fair value is critical when trying to develop better pricing models, given that such models are intended to provide estimates of theoretical fair value.

## SUGGESTED READING

Good coverage of option basics and of various strategies employing options, as well as an excellent discussion of hedging and the Greeks, can be found in *Options as a Strategic Investment* (McMillan, 1993). Lots of practical information for traders, including a discussion of how option activity can reveal the activities of insiders and signal trading opportunities, appears in *McMillan on Options* (McMillan, 1996). Another excellent book on the basics is *Options: Essential Concepts and Trading Strategies* (The Options Institute, 1999).

CHAPTER 2

# Fair Value and Efficient Price

Fair value and pricing models were briefly discussed in the previous chapter. However, fair value was never precisely defined, nor were the inner workings of pricing models examined. It is now time to explore and define the concept of fair value. Since the purpose of an option pricing model is to appraise fair value, good knowledge of the concept is critical. In this chapter, it is shown how fair value is related to market efficiency and how it may be context dependent. Next, fair value is considered with respect to arbitrage and statistical expectation regarding future prices. Statistical (mathematical) expectation is then examined in detail, using both direct and Monte Carlo methods, since it is such a crucial concept in the context of this work. How statistical expectation and arbitrage relationships may be used to price options is demonstrated. This forms a basis for a practical understanding of how option pricing models work. In short, this chapter is intended to provide some of the technical background necessary for what is to follow in subsequent chapters.

## DEFINING FAIR VALUE

What is fair value? A good place to begin with is the intuitive or commonsense understanding that every trader and investor has of the term. Most traders and investors attach meaning to terms



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

such as "undervalued" and "overvalued." Perhaps a stock has a low price-to-earnings ratio, or the company's breakup value is greater on a per-share basis than its current stock price. Such fundamental information might suggest that the stock is priced below its true value. Perhaps all potential investors have already purchased the stock, driving its price into orbit, and there are no buyers left to keep the ball rolling. Sentiment and contrary opinion theory might then lead to the belief that the stock is overpriced and ready for a serious correction. Technical analysis might reveal important support and resistance levels and, like the analysis of fundamentals and sentiment, contribute to a sense of a stock's current value. Fair value is understood by way of contrast: intuitively, a price represents fair value if it is neither a bargain (excessively low) nor a swindle (excessively high).

Traders and investors understand that it pays to buy undervalued securities and to sell overvalued ones, and that prices sometimes become discordant with reality. A bargain is worth purchasing as it will eventually be discovered and bid up by other traders until it is no longer a bargain. The trader who bought the stock or option at a bargain price will have made a profit. Likewise, an overvalued security is one that a savvy trader may want to short. As prices fall to more appropriate levels, the trader can cover the short and take a profit. Traders and investors also grasp the fact that it is much harder to profit when a security is neither undervalued nor overvalued, i.e., when its price represents fair value. This commonsense understanding of fair value is quite good, as far as it goes. However, a more precise, explicit, and computationally useful definition is required.

For the purpose of this book, fair value is defined as that value or price which discounts all foreseeable events and all public information (fundamental, technical, or otherwise) regarding a security. It is that price which allows no exploitation by traders for a profit, and to which the market price of a security will return should that security momentarily become overpriced or underpriced. In short, a security is fairly valued when it is efficiently priced by the market.

This notion of fair value refers to the so-called *efficient market hypothesis*. The efficient market hypothesis essentially says that all publicly known information has already been factored into securities prices and, therefore, that traders cannot make profits beyond those that derive simply from taking on nondiversifiable risk.

## FAIR VALUE AND THE EFFICIENT MARKET

Obviously, there is an intimate relationship between fair value and market efficiency. When a market behaves in accord with the efficient market hypothesis, the prices of securities are such that traders and investors cannot systematically profit by taking speculative positions or engaging in arbitrage. In other words, an "efficient" market is one where all prices reflect fair value. Were securities not fairly priced, traders could profit. Traders could short overpriced securities and take long positions in underpriced securities, exploiting the mispricings. However, in an efficient market, traders cannot systematically profit. Ergo, efficient pricing is pricing at fair value.

Real markets, of course, are efficient only to the extent that the actions of traders and other participants make them so. Overall market efficiency comes about as a result of traders attempting to capitalize on momentary inefficiencies. Their very actions produce rapid adjustments in price, adjustments that eliminate any further opportunities for profit. The trader who accumulates underpriced securities represents demand: he or she will bid up the prices. Likewise, the trader who sells overpriced securities in an effort to exploit inefficient pricing adds to supply, causing prices to fall. In the process, exploitable inefficiencies are more or less banished from the market.

## THE CONTEXT DEPENDENCE OF FAIR VALUE

Although often presented as a generic characteristic, fair value is actually context dependent. Fair value or price differs depending on the angle from which the problem is examined. Only in relation to specific trading strategies, and the kinds of inefficiencies these strategies seek to exploit, does fair value exist. As mentioned earlier, pricing efficiency is maintained by the collective actions of traders as they strive to profit from inefficiencies that arise in the marketplace. Their efforts have the effect of driving the inefficiencies out; nevertheless, a trading strategy does not cast out all inefficiencies. It only tends to eliminate inefficiencies of the kind that it is designed to exploit, and this is where context dependence enters the picture.

A particular trading strategy capitalizes on specific deviations from efficient price or price relationships. To the extent that it is employed, it brings prices and price relationships into line, i.e., it forces them to fair value. Prices or price relationships become "fair" with respect to a given strategy when that strategy can no longer yield a profit. However, what is fair or efficient for one strategy may or may not be fair or efficient for another. A given strategy capitalizes on, and tends to eliminate, a specific kind of inefficiency or deviation from fair value. It does not annihilate all inefficiencies or mispricings. In fact, in the process of doing away with one kind of market inefficiency, a given strategy may actually accentuate some other kind of inefficiency. Conflicts can easily arise between the various kinds of fair value being imposed on the market by diverse trading strategies. Because of its context dependent nature, fair value must be analyzed in the context of a strategy designed to exploit a specific kind of mispricing.

As stated, different trading strategies seek to exploit different types of inefficiency in the market in their quest for profits. Conversion and reversal arbitrage, for instance, attempts to exploit inefficient pricing relationships between puts and calls with the same terms. So-called "value investing" attempts to profit from situations where companies appear to have greater or lesser fundamental value than reflected in the price of their stocks. Technical trading strategies look for predictable trends, cycles, support and resistance, and other related inefficiencies that can possibly be exploited for a profit. The effect of traders employing these strategies is to reduce or eliminate the inefficiencies from which such strategies profit. The more a particular strategy is applied, the smaller the inefficiency on which it capitalizes becomes, until the strategy is no longer profitable except to those with the most sophisticated implementations, lowest trading costs, and fastest executions.

A good example of how the widespread use of a strategy can eliminate the very market inefficiency on which it depends can be found in the realm of breakout systems. Until the late 1980s, breakout systems were incredibly profitable across a wide range of futures contracts. As breakout systems received increasing coverage in books and trade magazines (such as *Technical Analysis of Stocks and Commodities* and *Futures*), more and more traders began using them. As a result, it became progressively more difficult to profit from most breakout systems. By the middle 1990s, it was possible to profit only from the most sophisticated of such systems, and then only in certain futures markets. The extensive use of breakout-based trading methods has made it more difficult, generally, to find enduring, easily captured trends. On the other hand, breakout trading has increased the amount of noise and countertrend activity in the market, possibly making it easier to implement and profit from trading-range strategies, such as those based on oscillators or support and resistance. The impact and performance of breakout systems over time was thoroughly analyzed and discussed in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000).

One way to look at the multiplicity of fair values and strategic contexts is by treating each strategy as a force that imposes a fair value constraint, one that may be expressed in terms of some mathematical equation, on the market. As strategies are applied in the markets, such equations are imposed with more or less force. In many cases, a solution that simultaneously satisfies all equations may be possible. This is the ideal sought by academics. In other instances, however, a single solution may not exist that fully satisfies all equations. When dealing with option pricing, one can readily think of situations involving more equations than there are unknowns; algebraically, this is not surprising. Conflict might then be imagined to exist between different kinds of fair value, and some kind of least-squares solution might be necessary. Using a weighted least-squares approach, each of the many equations could be assigned a weight, representing the force with which it is imposed on the market. Options traders can find treasures buried in these ideas. How a strategy imposes a mathematical constraint on the market will be seen below in the context of conversion and reversal arbitrage.

To summarize, an option or other instrument is fairly valued if a trader cannot consistently profit by exploiting an inefficiency in the marketplace. Different kinds of fair value exist because there are different kinds of inefficiencies, generated by distinct market forces, that can arise in the markets and that traders can try to exploit for a profit using a variety of methods.

## UNDERSTANDING AND ESTIMATING FAIR VALUE

A good understanding of fair value is essential in the context of developing and improving option pricing models. The whole purpose of a pricing model is to estimate fair value. Fair value is important because it provides a solid baseline against which actual market prices can be judged, and because a model for estimating it is required to calculate the Greeks, and hence for hedging. Throughout this work, many ways to measure or estimate fair value are examined. Popular pricing models are dissected to learn how they are constructed and to determine how well they approximate fair value. An increasingly refined understanding of fair value and its estimation is the goal.

How can fair value be estimated? The seeds of the answer are in the above discussion. An estimate of fair value or price can be obtained by looking at a strategy designed to exploit some inefficiency in the marketplace and ascertaining the prices or price relationships necessary to eliminate that strategy's profit potential. In other words, if it is possible to find those prices or price relationships that eliminate the inefficiency from which a strategy can profit, then it is possible to determine a fair value or price, or a fair price relationship.

## FAIR VALUE AND ARBITRAGE

Fair value is highly related to arbitrage. In fact, the elimination of arbitrage opportunities is, in large measure, what fair value is all about. Prices are considered fair when they are at those levels that make it difficult, if not impossible, to profit by arbitrage.

One example of fair value can be found in the put-call parity relationship that is maintained by conversion and reversal arbitrage. In the context of conversion and reversal arbitrage, a fair price is one that leads to put-call parity and thus eliminates any arbitrage opportunity; i.e., if the prices of puts and calls with the same terms are in parity, no arbitrage opportunity exists and the options may be considered to be efficiently priced. Note that this kind of fair or efficient price is not absolute, but only of one option relative to another. The put-call parity relationship implies that puts and calls will have roughly the same amount of time premium, if the terms of the options are similar. Of course, interest and dividends, which make the put-call parity formula a little more complex than presented in the last chapter, have not been considered. The equation defining the put-call parity relationship shown below is essentially the same as Equation 1.1 in the previous chapter, except that interest and dividends are now included.

$$P + C = S - U - I + D \tag{2.1}$$

In the equation, $P$ and $C$ are the prices of the put and the call, respectively, with the same terms, $S$ is the strike price of the options, $U$ is the price of the underlying stock, $I$ is the interest (as a dollar amount over the time remaining until expiration), and $D$ is the dividend (also as a dollar amount over the same period). Given the terms of the options, the stock price, the interest, and the dividends, this equation indicates that a corresponding call price may be calculated for any put price, and for any call price, a corresponding put price may be determined. If plotted on a graph, where the $x$-axis represents call price and the $y$-axis put price, the relationship expressed by Equation 2.1 appears as a diagonal line. Any pair of call and put prices falling on this line satisfies the put-call parity relationship and offers no opportunity to profit from conversion or reversal arbitrage. In other words, option prices that fall on the line represent fair values with respect to this form of arbitrage. To the extent that prices deviate from this line, and thus fail to satisfy the constraint expressed in Equation 2.1, there exists an inefficiency that conversion or reversal arbitrage can exploit for a profit. In other words, as prices move away from the line, the options become less fairly priced relative to one another. Arbitrage activity will then tend to push prices back to the line, back into a fair value relationship.

In the context of conversion and reversal arbitrage, fair value is strictly relative. It applies only to a price relationship between puts and calls, and not to the prices of either the calls or puts on their own. Equation 2.1 cannot be used to calculate the fair value for a put without having a fair price for the corresponding call, or vice versa. There are, in fact, an infinite number of prices that satisfy the put-call parity relationship and are thus efficient with respect to conversion and reversal arbitrage. In other words, put-call parity defines a price relationship involving two unknowns, the put price and the call price (the strike and stock prices, interest rates and dividends are assumed known in this context).

As can be seen from the equation, and from the brief discussion of conversion and reversal arbitrage in the previous chapter, the kind of fair value involved in put-call parity in no way depends on either the current or future price behavior of the underlying stock. The equation represents a clean, deterministic relationship rather than a messy, statistical one. There is no need for any sophisticated option pricing model, involving a host of assumptions regarding distributions of returns, to calculate this kind of fair value. Equation 2.1 is the only pricing model—and it is a pricing model—that is required.

## FAIR VALUE AND SPECULATION

Another kind of fair value comes into play in the context of speculative trading, i.e., when options are bought, or sold short, in an effort to profit from changes in price over time. Speculative fair value is the kind of value that must be considered when statistical uncertainty intrudes. Again, fair value can be analyzed in terms of efficient pricing. As with the prices of any security, option prices are efficient when they fully discount the future. When an option is priced so that a trader buying it with the intention of exercising or selling it later (or shorting it with the intention of either being assigned, covering, or having the option expire worthless) will, on average, neither take a profit nor sustain a loss, the option can be said to be at fair value. Fairly priced options, in this sense, are those that offer no real speculative opportunity, at least not statistically. Options with prices that deviate from fair value do offer the trader an opportunity to profit. A trader can buy underpriced options and short overpriced ones. This activity will tend to bring prices back in line with fair value, and thus to eliminate the inefficiencies from which the speculative trader seeks to profit.

In the context of speculative or directional trading, and in contrast to the case with conversion and reversal arbitrage, some method is necessary to foresee prices when the option is expiring or exercised, and its value easily ascertained. What happens down the road defines the profit or loss that will result from taking a position. However, without a crystal ball or time machine, it is impossible to use a deterministic model, as was possible with conversion and reversal arbitrage, to assess fair value. Instead, it is necessary to resort to some form of stochastic modeling or estimation. This means that some consequential assumptions must be made regarding the statistical behavior of the underlying stock. These assumptions may then be used to determine the mathematical expectation for an option's value at some point in the future when that value is easy to directly compute, for instance, at expiration.

## ESTIMATING SPECULATIVE FAIR VALUE

To illustrate how speculative fair value can be appraised and how option pricing models work, a simple option pricing model will be constructed. For the sake of clarity, dividends and interest will be ignored and a European-style option (no early exercise rights) will be assumed. In order to appraise speculative fair value for such an option, two items are required: the set of possible outcomes (e.g., prices) and the probabilities or frequencies associated with these outcomes. How are the potential outcomes and their associated probabilities found? In the world of option pricing, these items are often calculated based on some theoretical model of stock price movements. Such a model generally involves random variables and probability distributions. The probabilities may be obtained directly (or by iteration) from some formula, or they may be estimated using a Monte Carlo technique. Both Monte Carlo and direct mathematical solutions have their place, with the direct method usually being faster and more precise. However, a well-defined mathematical solution may not exist, making a more flexible method such as Monte Carlo necessary. For this reason, and for didactic purposes, both techniques are demonstrated.

## Modeling the Underlying Stock

Because probabilities must be established, the construction of the illustrative model begins with assumptions regarding the statistical behavior of the underlying stock. Imagine a strange world in which stock prices move randomly up or down $1 every day, with the direction of movement decided by the flip of a coin. Further, imagine that the stock under scrutiny is currently trading at $50, and that an at-the-money call option expiring in 10 days is to be priced.

Given assumptions regarding the behavior of the underlying stock, one way to determine the probability distributions and potential outcomes required for pricing the option is with a Monte Carlo analysis. The first step in such an analysis involves the creation of synthetic price series for the underlying stock that behave consistently with the model's assumptions. For the current assumptions, the following algorithm (expressed in the Microsoft Excel *Visual Basic for Applications* macro language) will generate one such price series:

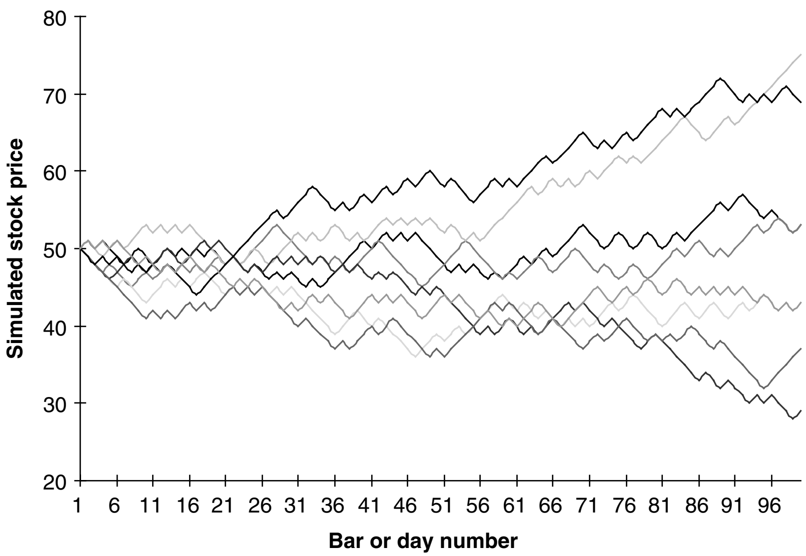
```
Price(0) = 50
NumberOfDays = 10
For I = 1 To NumberOfDays
      If Rnd() > 0.5 Then Shock = 1 Else Shock = -1
      Price(I) = Price(I-1) + Shock
Next I
```

An *algorithm* is simply a recipe for performing a calculation or solving a problem. The logic of the above algorithm is straightforward: Start with an initial price of $50. Then, for each day, determine the required price movement or "shock" (`Shock`), using a random number generator (`Rnd`), and calculate the new price by adding the shock to the previous day's price. Repeat the process until a price has been calculated for each day in the series. In the code above, price shocks take on values of +\$1 (heads) or –\$1 (tails), each with a probability of 0.5, consistent with the imagined scenario in which daily price changes are determined by the toss of a coin. The whole procedure may itself be repeated as many times as necessary to generate the required set of price series. An example of a set of eight price series constructed using the above algorithm, each starting with an initial price of \$50 and running for 100 days, is shown in Figure 2–1. Series containing 100 days of prices (rather than 10 days, as used for pricing the option) appear in the figure for the sake of visual clarity. Each price series shown in Figure 2–1 illustrates a particular kind of binomial random walk.

For pricing the option, Monte Carlo analysis involves repeatedly applying the above algorithm to generate a large number of synthetic price series, such as those shown in Figure 2–1.

**FIGURE 2–1**

Eight Synthetic Price Series Following a Binomial Random Walk



The only difference is that each series generated is 10 days long, the number of days remaining before the option expires. Essentially, the analysis involves conducting a kind of experiment or simulation with many trials. Once the experiment has been done, i.e., once the appropriate price series have been generated, the outcomes and their respective probabilities or frequencies may be tabulated. Such tabulations contain exactly the information necessary for pricing the option.

Another way to determine the required probabilities and potential outcomes for pricing the option is to use a mathematical probability distribution or, more precisely, a probability density function. For prices moving in a unit binomial random sequence, which is how prices move for the imagined stock, the binomial distribution is the appropriate choice. The density function for the binomial distribution, shown in Equation 2.2, can be used to directly calculate the probabilities associated with the various possible stock price outcomes.

$$f_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{2.2}$$

Equation 2.2 gives the probability of having exactly $k$ wins (positive shocks) appear in $n$ games (days) when the probability of a win in any single game is $p$ (here, 0.5).

Table 2–1 contains the tabulated final price outcomes, frequencies, and probabilities from a Monte Carlo experiment involving 75,000 price series of the kind shown in Figure 2–1, prepared using the algorithm described earlier. The table also contains theoretical probabilities derived from the binomial density function. The columns in Table 2–1 are as follows: The first column, *Final Stock Price*, contains the possible stock price outcomes at the end of each data series. For each trial, a series of 10 daily prices was constructed. The possible terminal stock price ranged from 40 to 60 in steps of two. The second column, *Frequency of Occurrence*, contains the number of times each price outcome was observed when considered over all 75,000 trials. The column sum is 75,000. This reflects the fact that the summed frequencies of occurrence for each of the possible outcomes must equal the total number of trials. Each number in the third column, *Monte Carlo Probability*, represents the probability

**TABLE 2–1**

Frequencies, Probabilities, and Mathematical Expectation for a Stock Exhibiting Binomial Random Walk Behavior

| Final Stock Price | Frequency of Occurrence | Monte Carlo Probability | Theoretical Probability | Monte Carlo Probability * Price Level | Theoretical Probability * Price Level |
|---|---|---|---|---|---|
| 40 | 79 | 0.001 | 0.001 | 0.042 | 0.039 |
| 42 | 730 | 0.010 | 0.010 | 0.409 | 0.410 |
| 44 | 3,286 | 0.044 | 0.044 | 1.928 | 1.934 |
| 46 | 8,731 | 0.116 | 0.117 | 5.355 | 5.391 |
| 48 | 15,469 | 0.206 | 0.205 | 9.900 | 9.844 |
| 50 | 18,308 | 0.244 | 0.246 | 12.205 | 12.305 |
| 52 | 15,376 | 0.205 | 0.205 | 10.661 | 10.664 |
| 54 | 8,926 | 0.119 | 0.117 | 6.427 | 6.328 |
| 56 | 3,300 | 0.044 | 0.044 | 2.464 | 2.461 |
| 58 | 727 | 0.010 | 0.010 | 0.562 | 0.566 |
| 60 | 68 | 0.001 | 0.001 | 0.054 | 0.059 |
| Sums | 75,000 | 1.000 | 1.000 | 50.007 | 50.000 |

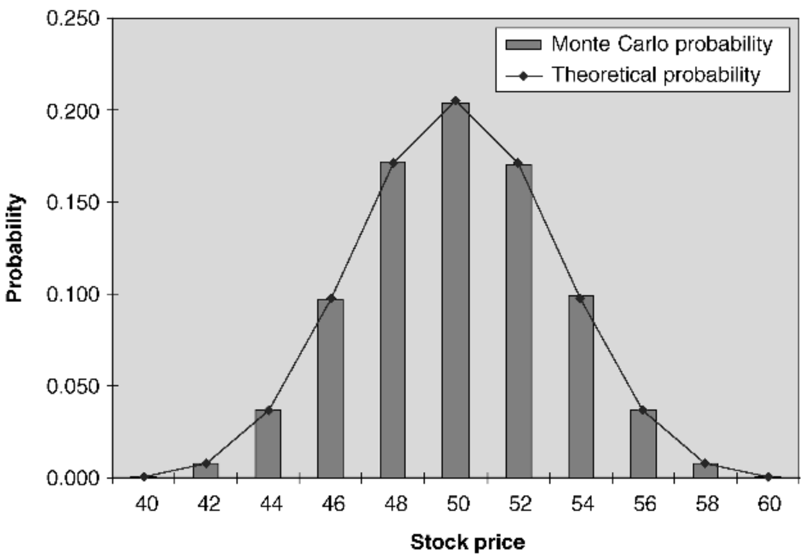
of observing a given possible outcome, i.e., the final price, in the context of the experiment. The sample probability (normalized frequency) of any outcome is obtained by dividing its frequency by the number of trials or, equivalently, the sum of frequencies over all possible outcomes. The numbers in this column add to unity. The fourth column, *Theoretical Probability*, contains probabilities derived from the binomial density function. These numbers represent the theoretical probabilities associated with each of the possible outcomes. As with the sample or Monte Carlo probabilities, the theoretical probabilities add to unity. Each number in the fifth column, *Monte Carlo Probability * Price Level*, was calculated by taking the product of the corresponding numbers in the columns "Final Stock Price" and "Monte Carlo Probability." The sixth column, *Theoretical Probability * Price Level*, is similar to the fifth column, except that stock prices are multiplied by the probabilities obtained from the binomial distribution rather than those derived from the Monte Carlo analysis. As was done for the second, third, and fourth columns, sums were also computed for the fifth and sixth columns. These later sums define the mathematical expectation for the stock price (not the option price) at the close of the tenth day.

Figure 2–2 graphically displays the probability distributions obtained from both the Monte Carlo method (column 3, plotted as bars) and the theoretical density function (column 4, plotted as markers on a line). Note the similarity of the data shown in Figure 2–2 to the familiar bell-shaped curve of the normal distribution. This similarity is not coincidental.

Comparisons reveal that, despite what may be considered a relatively small sample (75,000 trials), the probability estimates obtained from the Monte Carlo experiment are remarkably close to the exact probabilities derived from the binomial density function. Only in the third significant digit do these probabilities differ, and then only by a small amount. The differences are too small to be seen in the graphical presentation of the probability data. Mathematical expectations (sums under the last two

**FIGURE 2–2**

Probability Density for Final Stock Price



columns) for final stock price obtained from the two methods also agree exceptionally well, with differences not appearing until the fifth significant digit. These results demonstrate the power and utility of the Monte Carlo method.

What are these numbers, referred to as "mathematical expectations," that appear under the last two columns in Table 2–1? Each number represents the theoretical mean or average of the final stock price. The first of the two numbers is the average estimated by the Monte Carlo approach. The second of the two numbers is the average calculated with the binomial density function. For the Monte Carlo analysis, it is easy to verify that the sum under the column labeled *Monte Carlo Probability * Price Level* is really nothing more than a simple average of the final prices. Just add the final prices (all 75,000 of them) and divide by 75,000 to obtain the average. This average works out to be $50.007, a perfect match to the expectation computed by adding the products of price and probability. Such an alternative trial-by-trial calculation cannot be performed for the solution based on the binomial distribution since, for that solution, expectation represents a theoretical average based on an infinite number of trials. However, the products of price and probability can be used to compute the expectation. The sum of these products (as shown in Table 2–1) yields a theoretical expectation or average of $50. This is exactly what one would expect given a starting price of $50 and a symmetric random walk like the unit binomial walk that was assumed for the hypothetical stock.

In the language of the mathematician, expectation is the first moment of a distribution. The term "moment" comes from the idea of a balance point. Imagine bricks stacked at various positions along a board, each position corresponding to a price outcome, and the number of bricks stacked there corresponding to the outcome's frequency. The first moment would be the balance point of the board. Described this way, mathematical expectation sounds esoteric. However, as just demonstrated, it is familiar to anyone who has ever calculated a mean or average. The term "expectation" is sometimes employed preferentially when the mean or average under scrutiny is theoretical or intended to describe a statistical distribution of events.

In the discrete case (i.e., when the potential set of outcomes is finite or at least denumerable) expectation is defined by

$$\bar{x} = \sum_{i=1}^{n} f_i x_i \tag{2.3}$$

where $\bar{x}$ is the expectation, $i$ an index variable, $n$ the number of potential outcomes, $f_i$ the probability (or normalized frequency) of the $i$-th potential outcome, and $x_i$ the value (e.g., price) corresponding to the $i$-th potential outcome. This equation, although perhaps unfamiliar when expressed mathematically, represents just those calculations by which the last two columns and their sums in Table 2–1 were determined.

After a somewhat lengthy mathematical diversion, the problem of appraising speculative fair value can finally be brought back in focus. Numbers representing price expectation for a hypothetical stock at a specified future time are shown in Table 2–1. How is future price expectation related to speculative fair value? Recall that a security is fairly priced when a trader will, on average, break even over a sufficiently large number of trades. Assume that trades are entered on day zero, and closed out on day $n$ (in this instance, day 10). Stated in mathematical terms, $p_0$ (an unknown constant) represents a fair price if

$$0 = E(p_n - p_0) \tag{2.4}$$

where $E$ is the expectation operator, and where $p_n$ is the final stock price (a random variable) when the trade is closed out. In other words, $p_0$ (the entry price) is fair if the statistical expectation for the trade's profit or loss (the final or exit price minus the entry price) is zero. Given that the expectation of a constant is just the constant, and that the expectation operator is associative, Equation 2.4 can be rearranged as

$$0 = E(p_n) - E(p_0) = E(p_n) - p_0 \tag{2.5}$$

which finally yields

$$p_0 = E(p_n) \tag{2.6}$$

the sought-after fair price for the stock. In other words, a security is fairly valued (in a speculative sense) when its current price ($p_0$) is equal to the statistical expectation ($E$) of its final price ($p_n$).

In the case of the hypothetical stock analyzed in Table 2–1 under the assumption that trades would be closed out after 10 days, the fair price was \$50, as calculated exactly, or \$50.007, as approximated by Monte Carlo. On average, a trader entering a position at \$50 and exiting 10 days later could expect to break even (absent transaction costs). This is exactly what one might anticipate given a symmetric random walk: that profits and losses over any given interval would average out to zero on a per-trade basis.

What else (other than the stock's fair value) can be gleaned from Table 2–1 and the kind of analysis it represents? One thing easily determined from the data is the probability that the final stock price will lie within any specified interval. What is the probability that the stock will close above \$54 on the 10th day? Add the probabilities in Table 2–1 (exact or Monte Carlo) associated with prices greater than \$54. How likely is it that the stock will end with a price of \$50, unchanged from its initial price? Just locate the row that corresponds to a final stock price of \$50 and read the probability. This kind of information can be of practical use to traders of stocks and options.

### Pricing the Option

So far, only stock prices have been discussed and some mathematical background presented. Option prices have mostly been ignored. The reader may feel cheated since this section began by suggesting that it would illustrate the principles underlying the appraisal of an option's speculative fair value. However, what has been accomplished is very significant for pricing options and for understanding market movements. Imagine taking a stock at a given price, assuming that its movements behave in accord with some theoretical distribution (in this case, the binomial distribution), calculating the mathematical expectation for the stock's price at some point in the future, and demonstrating that this expectation is the stock's fair value. Most of the elements necessary for pricing an option are here and have, in fact, been illustrated. All that remains is to establish, using the same assumptions and probabilities, the mathematical expectation for the price of the option at expiration. As with the stock, future mathematical expectation defines speculative fair value.

Table 2–2 is very similar to Table 2–1. The only differences are that a *Final Option Price* column has been added, and that the last two columns, *Probability ∗ Price Level* and *Theoretical Probability ∗ Price Level*, have been modified so that they now contain products of probabilities with option prices rather than with stock prices. Each price in the column *Final Option Price* is simply the corresponding stock price minus the strike price, or zero, whichever is larger. This is the easily calculated worth of a call option about to expire and with no speculative or time value remaining. In the same way that the expectation of final stock price was determined by summing the products of probability and stock price, the expectation of final option price may be found by accumulating the products of probability and option price. In fact, the same probabilities used for pricing the stock may be used for pricing the option since the probability of the stock having, for instance, a terminal price of $54, is identical to the

**TABLE 2–2**

Frequencies, Probabilities, and Mathematical Expectation for an At-the-Money Call on a Stock Exhibiting Binomial Random Walk Behavior

| Final Stock Price | Final Option Price | Frequency of Occurrence | Monte Carlo Probability | Theoretical Probability | Monte Carlo Probability ∗ Price Level | Theoretical Probability ∗ Price Level |
|---|---|---|---|---|---|---|
| 40 | 0 | 79 | 0.001 | 0.001 | 0.000 | 0.000 |
| 42 | 0 | 730 | 0.010 | 0.010 | 0.000 | 0.000 |
| 44 | 0 | 3,286 | 0.044 | 0.044 | 0.000 | 0.000 |
| 46 | 0 | 8,731 | 0.116 | 0.117 | 0.000 | 0.000 |
| 48 | 0 | 15,469 | 0.206 | 0.205 | 0.000 | 0.000 |
| 50 | 0 | 18,308 | 0.244 | 0.246 | 0.000 | 0.000 |
| 52 | 2 | 15,376 | 0.205 | 0.205 | 0.410 | 0.410 |
| 54 | 4 | 8,926 | 0.119 | 0.117 | 0.476 | 0.469 |
| 56 | 6 | 3,300 | 0.044 | 0.044 | 0.264 | 0.264 |
| 58 | 8 | 727 | 0.010 | 0.010 | 0.078 | 0.078 |
| 60 | 10 | 68 | 0.001 | 0.001 | 0.009 | 0.010 |
| Sums | | 75,000 | 1.000 | 1.000 | 1.237 | 1.230 |

probability of the option having a terminal price of \$4. The individual probability-price products found in the last two columns of Table 2–2 can be interpreted as the contributions made by different potential outcomes to the total current value of the option. In the example, most of the option's value derives from instances where the stock finishes between \$52 and \$56. The expectations appear as sums under the last two columns and differ from one another only in the fourth significant digit, indicating a good agreement between the theoretical and Monte Carlo calculations. Both numbers represent fair value for the option.

In Table 2–2, the column *Option Price* contains the prices at expiration (on day 10) for a call option with a strike price of \$50. Note, however, that this column could just as easily hold prices for a call option with a different strike, for a put option, or even for some combination involving multiple options, e.g., a straddle. The same calculations and probabilities used to price the example call could also be used to determine the new position's fair value. Even for an option alive beyond day 10, the mathematics would still apply if acceptable fair values for that option on day 10 could be found and placed in the column. It would still be possible to obtain the fair value on day zero from such data.

In the above presentation, speculative fair value was analyzed from the viewpoint of statistical expectation. Interest and dividends were ignored. Some readers, however, may argue that the only proper approach to pricing an option is to assume a self-financing hedge involving a bond or other interest-bearing security, an option, and the underlying stock. Except for the way interest rates and dividends are handled, approaches involving self-financing hedges generally produce solutions that are equivalent to those based upon statistical expectation. In addition, whether implicitly or explicitly, both approaches require the same kinds of statistical assumptions about price movements and all of the other related mathematical paraphernalia.

Prevailing interest rates are easily handled in the context of an expectation-based approach to option pricing by assuming a risk-neutral world. A *risk-neutral world* is one in which all assets are expected to grow at the prevailing rate of risk-free interest. It makes sense to assume a risk-neutral world because, in the world of options, all risk can presumably be hedged. To correctly price an option by analysis of future expectation, the underlying stock must have an expected (in the statistical sense) rate of growth equal to the risk-free interest rate, and the expected future price of the option must be discounted for growth at that same rate to obtain the current fair value.

## SUMMARY

In this chapter, it was shown that fair value has much to do with market efficiency, that it is context dependent, and that two of the major forces that define or create fair value are arbitrage (including conversion and reversal arbitrage) and expectation regarding future prices. The relationship between conversion and reversal arbitrage and fair value was expressed in the form of a simple, deterministic equation. In the case of speculative fair value related to expectation, an example illustrated some of the elements that go into an option pricing model and how such pricing models work. The example demonstrated the Monte Carlo method and showed how one can proceed from assumptions regarding the price behavior of the underlying instrument to a model for pricing an option trading on that instrument. Some of the essential mathematical concepts required when pricing options were also covered. In the following chapter, the two most popular option pricing models will be examined and their inner workings dissected. The manner in which interest rates are handled in the context of expectation-based pricing will also be demonstrated.

## SUGGESTED READING

A good discussion of market efficiency and the Efficient Market Hypothesis can be found in *The Random Walk and Beyond* (Johnson, 1988). *A Random Walk Down Wall Street* (Malkiel, 1985) is another basic source on the subject. For those who wish to pursue the matter in greater depth, there is an endless procession of articles concerned with market efficiency in the academic journals.

CHAPTER 3

# Popular Option Pricing Models

Chapter 2 examined the concept of fair value. Fair value was demonstrated to be related to arbitrage. It was also shown to be related to speculative potential, which is a function of forward expectation, i.e., the statistical expectation (or average) of future prices. A simple option pricing model was developed and examined as a learning exercise. The model demonstrated how statistical expectation regarding future price could be calculated so as to determine the current fair value of a hypothetical option. For the sake of simplicity, the model used in the learning exercise made rather unrealistic assumptions about the statistical behavior of the underlying stock prices.

In this chapter, the two most popular models by which options are priced are examined and their inner workings revealed. One model is the Cox-Ross-Rubinstein or "binomial" model. The other is the well-known Black-Scholes model. In contrast to the crude model presented in Chapter 2, these two models make more realistic assumptions regarding the price behavior of securities. Although not flawless by any means, the assumptions made are good enough to permit these models to be of practical use to those who trade options or use them to hedge other investments. When their assumptions are satisfied, Black-Scholes and Cox-Ross-Rubinstein allow calculation of a fair price that maintains put-call parity, which consequently eliminates



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

opportunities for conversion and reversal arbitrage, and that makes it impossible to profit from speculation on future prices.

The reader should be warned that much of what follows is highly mathematical. It seemed appropriate to present this material to provide an in-depth understanding of option pricing that would serve as a sound basis for the work discussed in later chapters. Despite the mathematics, the aim is not to generate a lot of academic theory, but to work towards more trader-friendly pricing models—models that traders can use to gain an edge in the markets. Those who are mathematically challenged may wish merely to skim the next few sections or perhaps skip immediately to the Summary.

## THE COX-ROSS-RUBINSTEIN BINOMIAL MODEL

The discussion of Cox-Ross-Rubinstein begins with the simple model presented in Chapter 2. Surprisingly, that model is not unlike the real binomial model. As the reader may recall, the illustrative model presented in Chapter 2 relied on the assumption that stock prices move up or down $1 every day, the direction of the movement being randomly determined by the flip of a coin.

Is it rational to assume that stock prices move in this manner? Imagine a $10 stock. It moves up $1, which represents a reasonable 10% gain in price. Then there are several losses. It does happen. Now the stock is trading at $1. The stock moves up by $1 the next day, which is an impressive 100% gain. Then three more losses occur. The stock is now trading at –$1. But stocks neither take on negative prices nor do they generally fall from $1 to $0 in a single day. There are obviously some serious problems with the illustrative model and its assumptions.

One problem, easily remedied, is that stock price movements tend to be proportional or *geometric,* rather than additive or *arithmetic*. A $10 stock moving up $1 is roughly equivalent to a $1 stock gaining $0.10 or a $100 stock appreciating by $10. Random addition or subtraction must be replaced with random multiplication or division. Instead of adding $1 to a stock's current price to get the next price, perhaps multiply the price by

1.10; and instead of subtracting $1, divide by 1.10 (or equivalently, multiply by its reciprocal).

Another problem with the illustrative model from Chapter 2 is the assumption of one-size-fits-all price movements. Price movements in the real world differ, not only in direction but also in amplitude, from day to day and stock to stock. In the course of a typical day, some stocks fluctuate in price only by fractional amounts, while others trace out great swings. Stocks also differ in their long-term rates of growth. What this means is that the up and down ratios or multipliers must be made independently adjustable. Independently adjustable multipliers enable stocks having different forward expectations (growth rates) and volatilities to be correctly modeled. The probability of an up transition may remain fixed at 0.5, as in the example from Chapter 2, implying an equal probability for a down transition. Some simple equations that yield the correct up and down multipliers for a specified growth rate and volatility must also be provided.

Finally, should price shocks occur only once a day? Definitely not! A finer grained analysis is required. This problem is easy to rectify. Allow the time steps or intervals into which a given period is divided to be as small as desired. Of course, the finer the grain or smaller each time step, the more extensive and time consuming the calculations.

If the problems inherent in the original model and its assumptions are addressed by these relatively minor changes, a dramatically more realistic and useful model results. What results is a model of an efficient market with statistical properties that can be characterized by growth and volatility, the first and second moments of the statistical distribution of price movements. In fact, what has been arrived at is the equal probability implementation of the famous Cox-Ross-Rubinstein binomial pricing model.

### Specifying Growth and Volatility

How are growth and volatility set for the binomial pricing model? By choosing appropriate values for the up and down multipliers. The best place to start, however, is actually the other way around, with equations that yield the growth and volatility, given the up and down multipliers. It is easy to see that the annual growth rate is given by

$$r = \frac{1}{\Delta t} \ln[pu + (1-p)d] \tag{3.1}$$

where $\Delta t$ is the length of each time step (in years), $p$ the up transition probability, $u$ the up ratio, $(1-p)$ the down transition probability, and $d$ the down ratio. The growth rate may be recognized as the logarithm of the one-step forward expectation of a \$1 stock (the return achieved in a single time step) multiplied by the number of time steps in one year. Think of compound interest and you will have the right idea. Volatility (in percent per annum) can be determined as

$$\sigma = \sqrt{\frac{p(1-p)}{\Delta t}} \ln \frac{u}{d} \tag{3.2}$$

which follows from the fact that

$$\sigma^2 \Delta t = p(\ln u - \mu)^2 + (1-p)(\ln d - \mu)^2 \tag{3.3}$$

where

$$\mu = p \ln u + (1-p) \ln d \tag{3.4}$$

These equations allow growth and volatility to be expressed as functions of the up and down ratios and the transition probabilities. However, the goal is to express the up and down multipliers in terms of growth rate ($r$) and volatility ($\sigma$). Fixing the transition probability $p$ at 0.5, some algebraic manipulation of Equations 3.1 and 3.2 yields

$$d = \frac{2e^{r\Delta t}}{e^{2\sigma\sqrt{\Delta t}} + 1} \tag{3.5}$$

and

$$u = \frac{2e^{r\Delta t + 2\sigma\sqrt{\Delta t}}}{e^{2\sigma\sqrt{\Delta t}} + 1} \tag{3.6}$$

the expressions for the up and down multipliers in terms of the remaining variables.

The end of the road is almost in sight. All that is necessary is to accept the fact that, in a risk-neutral world, options can be priced using forward expectation. In a risk-neutral world, all assets grow at the same rate. This, of course, is the risk-free interest rate. If the growth is set to the rate of risk-free interest (T-bonds are a good proxy) and volatility is set to a suitable value that depends on the stock, it is possible to correctly model a stock's price behavior. Options trading on that stock can then be appraised for fair value.

## Monte Carlo Pricing

For continuity with Chapter 2, and because it is an effective learning exercise, a Monte Carlo simulation is used to price an option. In contrast to the previous effort, however, this exercise takes into account the volatility of the stock to be simulated and the risk-free rate of interest, properly handles the proportional nature of stock price variation, and employs a generally more acceptable statistical model of stock price behavior. The equal probability implementation of the Cox-Ross-Rubinstein model, on which this Monte Carlo example is based, is actually used by traders, hedgers, and market makers to price options.

The first step in setting up a Monte Carlo simulation is to construct an algorithm that can generate a series of prices that accord with the underlying theoretical model. An algorithm that generates simulated stock prices consistent with the equal probability variation of Cox-Ross-Rubinstein appears below:

```
subroutine MCBRW (s, n, t, v, r)
dimension s(n)
p = 0.5
dt = t/(n - 1)
evt = EXP (2.0 * v * SQRT (dt))
ert = 2.0 * EXP (r * dt)
d = ert/(evt + 1.0)
u = d * evt
do 10 i = 2, n
```

```
        if (RANFU ().lt. p) then
             s(i) = s(i-1) * u
        else
             s(i) = s(i-1) * d
        endif
10      continue
   return
   end
```

This time around, the algorithm is actualized as a Fortran subroutine. Popular with quants, Fortran is a language better suited to heavy-duty number crunching than Microsoft's *Visual Vasic* language. The name of the subroutine, `MCBRW`, is an acronym for Monte Carlo Binomial Random Walk. Not only is the language used for the Monte Carlo analysis now Fortran, but a high quality pseudo-random number generator is also employed. Built-in random number generators, such as `Rnd` in Microsoft's *Visual Basic*, tend to be of poor quality. These generators customarily fail many statistical tests of randomness. For serious Monte Carlo work, it is vital to use a random generator that passes such tests. The universal random number generator, proposed by Marsaglia and Zaman (in report FSU-SCRI-87-50), seems to be acceptable and is freely available over the Internet. Another generator that appears serviceable is the `RAN2` generator from *Numerical Recipies in Fortran 77* (Press et al., 1992). Function `RANFU` in the Fortran code above, which returns pseudo-random numbers between 0 and 1, employs the Marsaglia and Zaman generator.

The logic of the algorithm is as follows: First, the up and down multipliers (`u` and `d`) are calculated using Equations 3.5 and 3.6, respectively, from the volatility (`v`) and growth or interest rate (`r`) passed to the subroutine. Volatility and growth are specified in percent per annum. The probability of an up transition (`p`) is fixed at 50%, which also determines the probability of a down transition (`1–p`). The size of each time step (`dt`) is computed by dividing the time from start to finish (`t`) by the number of intervals into which it is to be partitioned (`n–1`, where `n` is the number of prices in the series `s`). In this algorithm, time is measured in years. Then comes the loop, which is very similar to the one appearing in the *Visual Vasic* code in Chapter 2. For each time step (indexed by `i`), the corresponding stock price is found by either multiplying the previous price by the up ratio (`u`) or by the down ratio (`d`). Note how multiplication by u or d has replaced addition or subtraction of 1. Just as in the former algorithm, a pseudo-random number generator decides which of the two operations is performed on each pass through the loop. Before executing the algorithm, along with the other arguments passed to the subroutine, the initial stock price must be supplied in `s(1)`; the subroutine will then be able to generate all of the remaining prices in `s(2)` through `s(n)`. The generated series will have a theoretical volatility of v and growth rate of r. It will span a period of t years.

For the Monte Carlo experiment presented below, volatility was set to 85% per annum (v = 0.85), risk-free interest to 10% per annum (r = 0.10), and elapsed time to 10 days (t = 10.0/365.25). In this instance, elapsed time is the time remaining until expiration. The parameter n was fixed at 30, making each time step equal to 8 hours. Both the initial stock price and the strike price of the call option to be valued were specified as $50. A set of 250,000 simulated price series were harvested. The final prices and their associated probabilities were then tabulated. This exactly follows the procedure employed in Chapter 2. Table 3–1 contains most of the results from the Monte Carlo experiment.

The first column in Table 3–1 lists all final stock prices that were observed at least once in the Monte Carlo sample. The final or *terminal* stock prices are the prices found in `s(n)` after calls to subroutine `MCBRW`. In the example from Chapter 2, terminal prices were evenly spaced along the real number line. Equal intervals are not a feature of prices in the current experiment. Instead of equal intervals, these prices exhibit equal proportions or ratios. The ratio of any price to the preceding lower price is always equal to the up multiplier (u) divided by the down multiplier (d). Although the prices themselves are not equally spaced, logarithms of these prices do fall at evenly spaced intervals.

For each terminal stock price in the first column, the second column in Table 3–1 has the corresponding final option price. Given that the terminal stock price is the price of the stock at option expiration, the final value of the option is easily reckoned. For a call that is about to expire, the fair premium is either zero, or the stock price minus the strike, whichever is

**TABLE 3–1**

Example of Pricing an Option with Monte Carlo Using the Equal Probability Implementation of the Cox-Ross-Rubinstein Binomial Model

| Final Stock Price | Final Option Price | Sample Frequency | Sample Probability | Probability * Stock Price | Probability * Option Price |
|---|---|---|---|---|---|
| 27.23 | 0.00 | 2 | 0.0000 | 0.00 | 0.00 |
| 28.69 | 0.00 | 6 | 0.0000 | 0.00 | 0.00 |
| 30.22 | 0.00 | 50 | 0.0002 | 0.00 | 0.00 |
| 31.85 | 0.00 | 221 | 0.0008 | 0.02 | 0.00 |
| 33.55 | 0.00 | 716 | 0.0028 | 0.09 | 0.00 |
| 35.35 | 0.00 | 1,950 | 0.0078 | 0.28 | 0.00 |
| 37.25 | 0.00 | 4,635 | 0.0185 | 0.69 | 0.00 |
| 39.25 | 0.00 | 9,308 | 0.0372 | 1.46 | 0.00 |
| 41.35 | 0.00 | 16,130 | 0.0645 | 2.67 | 0.00 |
| 43.56 | 0.00 | 24,240 | 0.0969 | 4.22 | 0.00 |
| 45.89 | 0.00 | 31,498 | 0.1260 | 5.78 | 0.00 |
| 48.37 | 0.00 | 36,100 | 0.1444 | 6.99 | 0.00 |
| 50.97 | 0.97 | 36,105 | 0.1444 | 7.36 | 0.14 |
| 53.71 | 3.71 | 31,771 | 0.1271 | 6.83 | 0.47 |
| 56.58 | 6.58 | 24,155 | 0.0966 | 5.47 | 0.64 |
| 59.61 | 9.61 | 16,041 | 0.0641 | 3.82 | 0.62 |
| 62.81 | 12.81 | 9,430 | 0.0377 | 2.37 | 0.48 |
| 66.17 | 16.17 | 4,631 | 0.0185 | 1.23 | 0.30 |
| 69.71 | 19.71 | 1,993 | 0.0079 | 0.56 | 0.16 |
| 73.45 | 23.45 | 731 | 0.0029 | 0.21 | 0.06 |
| 77.39 | 27.39 | 212 | 0.0008 | 0.06 | 0.02 |
| 81.54 | 31.54 | 60 | 0.0002 | 0.02 | 0.00 |
| 85.91 | 35.91 | 13 | 0.0000 | 0.00 | 0.00 |
| 90.52 | 40.52 | 1 | 0.0000 | 0.00 | 0.00 |
| 95.37 | 45.37 | 1 | 0.0000 | 0.00 | 0.00 |
| Column sums | | 250,000 | 1.0000 | 50.15 | 2.91 |

greater. When examining the numbers in Table 3–1, keep in mind that the call being appraised has a $50 strike price.

It should be mentioned that the option values in the second column need not be the prices of an expiring option as they are in the present instance. In fact, these prices can be the appraised values for the option at any point in time prior to expiration. What is crucial is that the option can somehow be valued at a specified point in time, and the appraised values can be placed in the second column. In such a case, the option valuation produced by the calculations in this Monte Carlo analysis would be for a period of time (`t`) prior to the time to which the appraised values correspond. This fact will become significant later in the discussion of binomial trees, when fair option values are determined recursively by working backwards through the tree, a layer at a time.

The third column in Table 3–1 contains the frequencies with which each of the stock and option prices occurred in the Monte Carlo sample. These frequencies total 250,000—the total number of trials in the experiment—confirming that every possible outcome was accounted for in the table.

Each sample probability (normalized frequency) in the fourth column of Table 3–1 was computed by dividing the corresponding frequency in the third column by the number of trials. The sample probabilities add to 1, as they should. Just as in the example from Chapter 2, these numbers can be used to assess the probability of the stock (or option) finishing in any particular range of price.

The fifth column in Table 3–1 contains the probabilities from the fourth column multiplied by the corresponding stock prices from the first column. These numbers sum to $50.15, the expected (mean) final price for the stock. The sum is slightly greater than $50, the initial price, because, in a risk-neutral world, all assets are expected to grow at the risk-free rate of interest, specified here to be 10% per annum. This price should actually be $50.14. The reason that the column sum differs from the theoretically correct value is that it is based on one particular sample of 250,000 prices and not on an infinite sample. Nevertheless, the Monte Carlo result is remarkably close to the exact, theoretical expectation for the final stock price in a 30-time-step model.

It must be pointed out that growth over any period is ordinarily measured as $\ln(b/a)$, where $b$ is the final price and $a$ the initial price. Compounding of growth is calculated as $m \ln(b/a)$, where $m$ is the number of time periods over which compounding takes place. In other words, growth is not always measured in terms of a percentage, although it is frequently referred to and treated as if it were. There is, however, a direct relationship between percentage-based and logarithmic measures of growth. A true 10% (0.10) gain is equivalent to a logarithmic growth of 0.095 or ln(1.10), while a loss of 20% (–0.20) is equivalent to a growth of –0.223 or ln(0.80). Likewise, a logarithmic growth of 0.20 is equivalent to percentage-based growth of 22.1% (0.221) or exp(0.20) – 1. Based on these examples, it should be clear how percentage-based expressions are related to logarithmic measurements of growth and vice versa. To the extent that growth is small, so is the difference between the logarithmic expression and the more familiar expression in terms of percentage gain (or loss). This is why in practical situations, involving normal interest rates and short periods of time, these two ways of defining growth are often used interchangeably. Consider the Monte Carlo experiment. A 10% yearly return from interest works out, over the 10-day period analyzed, to a return of 0.002609 in logarithmic terms or a gain of 0.2603% (0.002603) in true percentage terms. As can be seen, the difference between the two measures is rather inconsequential as far as option premiums are concerned.

The sixth and final column of Table 3–1 contains the products of probability and option price. The column sum is the expected value of the option at expiration. Once discounted for interest, this is the option's fair value according to the equal-probability variation of the Cox-Ross-Rubinstein binomial model. In a risk-neutral world, all assets, including options, grow at the risk-free rate of interest, hence the need to discount the final expected price for the option by the interest over the holding period. To discount for interest, the final expected value of $2.905 (rounded to 2.91 in Table 3–1) is divided by `EXP(r*t)`, with `r` and `t` as in the Fortran subroutine. With the specified interest rate of 10% per annum (logarithmic form) and a holding period of 10 days, the estimated fair value of the option is $2.898.

Apart from being useful for calculating option price, the numbers in the sixth column also demonstrate how certain potential outcomes contribute more to the current value of an option than do other potential outcomes. For example, much of the value in the option being priced in Table 3–1 derives from instances where the stock price finishes between \$53 and \$66. Although higher stock prices imply richer options, such prices occur with such rarity that the average payout on a relative basis is small. Lower stock prices, those between \$50 and \$53, are quite common. With lower stock prices, however, the value of the expiring option is rather small. Again, on a comparative basis, the expected or average payout from such outcomes is limited.

Although not shown, growth and volatility were also calculated for the Monte Carlo data. The observed volatility on an annual basis was 84.9%, as against the specified or theoretical value of 85%. Annual growth was 0.106 rather than 0.10 to which it was set. As with other quantities taken from the Monte Carlo experiment, the small discrepancies are due to the sampling errors present whenever a finite sample is used in an experiment. When necessary, greater accuracy can be readily achieved by employing larger samples.

## Pricing with Binomial Trees

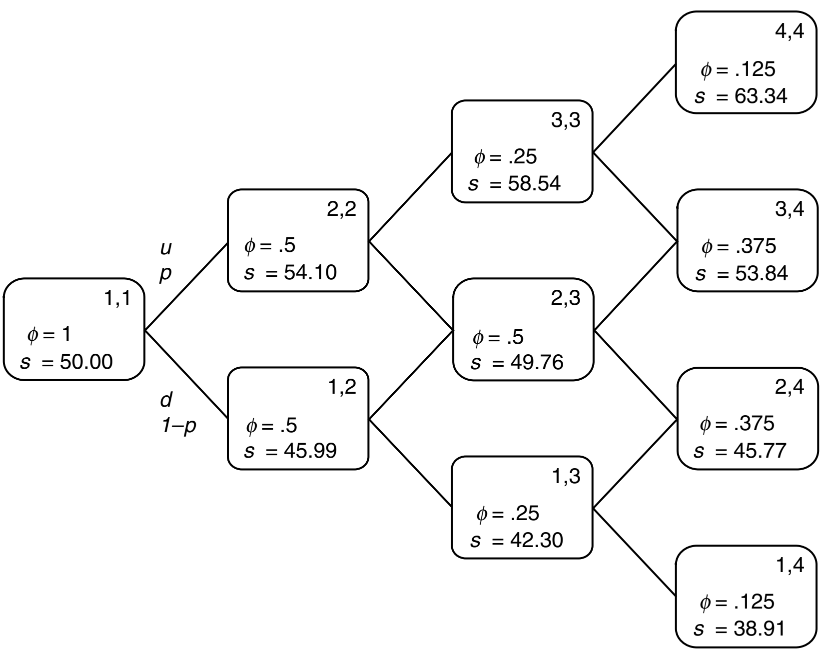
Although binomial pricing may be accomplished with Monte Carlo, there exists an alternative approach that is flexible, powerful, and quite elegant: the method of binomial pricing trees. A *binomial pricing tree* is a structure that maps all possible trajectories of stock price through time as are allowed by the model. This structure consists of nodes and branches. The nodes of a binomial tree are arranged in columns or layers, like the nodes of a neural network. Each layer corresponds to a particular moment or time step. Each node in a given layer, therefore, corresponds to a potential stock price at a particular point in time. Nodes are identified with traversal probabilities and option valuations, as well as with stock prices. Nodes, and the data items with which they are associated, are easily indexed as elements in matrices, which indeed they are. A convenient indexing scheme has the layer or time step represented by $j$ (a number between 1 and $n$, the number of layers or time steps) and the nodes within each layer (the potential stock prices) by $i$ (a number between 1 and $m$, the number of nodes in the layer). Depending on whether or not the tree is recombining, the node count $m$ for any given layer may range from $j$ to twice the number of nodes in the previous layer.

Each branch or path in a binomial pricing tree represents a possible transition from one node to another node later in the tree and has a probability and a ratio associated with it. The probabilities and ratios associated with the branches are the same as those discussed earlier. Branches to higher nodes reflect up probabilities ($p$) and multipliers ($u$), while branches to lower nodes implement the down probabilities ($1 - p$) and multipliers ($d$). In a uniform, equal probability tree, $p$ and $1 - p$ are 0.5 for all branches, all upside branches share the same $u$, and all downside branches share the same $d$. Such a tree is recombining.

To illustrate the process of pricing an option on a binomial tree, consider Figure 3–1, which depicts a binomial tree with four layers. In this figure, nodes are shown as round-cornered

**FIGURE 3–1**

A Recombining Binomial Tree with Node Indices, Stock Prices, and Traversal Probabilities



boxes and branches or transitions are shown as lines running from the boxes in one column or layer to those in the next. Node indices appear as pairs of numbers in the upper right corners of the boxes. Counting is from left to right and bottom to top.

Given such a tree, whether drawn on paper or implemented as a set of arrays in a computer program, the first step in pricing an option is to fill in the stock prices and traversal probabilities for each node. Nodes are filled in recursively, column by column from left to right, beginning with the one and only node in the first layer of the tree. The stock price for the first node ($s_{1,1}$) is simply the initial stock price for which a fair option price is desired; it is the stock price that would have appeared in `s(1)` in the Fortran code implementing Monte Carlo price series generation. The traversal probability ($\phi_{1,1}$) is 100% or 1, since this stock price is the one and only possible price at the initial time step that corresponds to the first layer in the tree.

Nodes in the second layer are filled in using the transition parameters and values from nodes in the first layer. For node 2,2, the stock price ($s_{2,2}$) is the price from the previous node ($s_{1,1}$) multiplied by $u$, and the probability ($\phi_{2,2}$) is the probability from the previous node multiplied by $p$. For node 1,2, the stock price ($s_{1,2}$) is the price from the previous node multiplied by $d$, and the probability ($\phi_{1,2}$) is the probability from the source node multiplied by $1 - p$. The same kinds of calculations are repeated to get values for the nodes in the third layer from those in the second. In this case, however, there may be a pair of nodes that have transitions to the node under scrutiny. When more than one path to a node exists, the node's traversal probability is the sum of the probabilities deriving from all possible paths. The stock price for a given node in a recombining binomial tree may be calculated either as $u$ multiplied by the stock price at the previous lower node or $d$ multiplied by the stock price at the previous higher node; both calculations yielding the same result. In fact, this is what is meant when a binomial tree is said to be *recombining*: that different paths recombine at future nodes.

In general, for a recombining tree, stock prices in one layer may be determined from those in a previous layer by

$$s_{ij} = us_{i-1,j-1} = ds_{i,j-1} \quad \text{for all } 1 < i < j \tag{3.7a}$$

with

$$s_{ij} = ds_{i,j-1} \quad \text{for } i = 1 \tag{3.7b}$$

and

$$s_{ij} = us_{i-1,j-1} \quad \text{for } i = j \tag{3.7c}$$

The probabilities of traversal may likewise be computed, layer-by-layer, as

$$\phi_{ij} = p\phi_{i-1,j-1} + (1-p)\phi_{i,j-1} \quad \text{for all } 1 < i < j \tag{3.8a}$$

with

$$\phi_{ij} = (1-p)\phi_{i,j-1} \quad \text{for } i = 1 \tag{3.8b}$$

and

$$\phi_{ij} = p\phi_{i-1,j-1} \quad \text{for } i = j \tag{3.8c}$$

In the above equations, $\phi_{ij}$ and $s_{ij}$ refer to the specific $\phi$ and $s$ of the $ij$-th node.

Traversal probabilities and stock prices were filled in for all nodes in Figure 3–1 with the aid of a small Fortran subroutine implementing the layer-by-layer approach outlined above. With the calculation of the data in the figure, the first step in the two-step process of pricing an option has been completed.

Like Table 3–1, Table 3–2 displays the distribution of terminal stock prices and their probabilities taken from the last column of nodes in a tree such as that depicted in Figure 3–1. Table 3–2 also contains the price-probability products and the expectations computed by summing these. In other words, the data in Table 3–2 are for the same transition parameters $u$, $d$, and $p$, but generated using a binomial tree and Equations 3.7 and 3.8 rather than by Monte Carlo as was the case for Table 3–1. A few rows appear in Table 3–2 that were absent in Table 3–1. This is because some low probability events simply did not occur in the Monte Carlo sample. The expected final (at option expiration) price for the stock was \$50.14. For the option it was \$2.896, which, when discounted for growth (interest), works out to be \$2.888. In these calculations, $r$ (the risk-free interest rate) was 0.10, $\sigma$ (the volatility) was 0.85,

**TABLE 3–2**

Example of Pricing an Option with a Binomial Tree Using the Equal Probability Variation of the Cox-Ross-Rubinstein Model

| Final Stock Price | Final Option Price | Terminal Node Probability | Probability * Stock Price | Probability * Option Price |
|---|---|---|---|---|
| 23.28 | 0.00 | 0.0000 | 0.00 | 0.00 |
| 24.53 | 0.00 | 0.0000 | 0.00 | 0.00 |
| 25.84 | 0.00 | 0.0000 | 0.00 | 0.00 |
| 27.23 | 0.00 | 0.0000 | 0.00 | 0.00 |
| 28.69 | 0.00 | 0.0000 | 0.00 | 0.00 |
| 30.22 | 0.00 | 0.0002 | 0.00 | 0.00 |
| 31.85 | 0.00 | 0.0008 | 0.02 | 0.00 |
| 33.55 | 0.00 | 0.0029 | 0.09 | 0.00 |
| 35.35 | 0.00 | 0.0079 | 0.28 | 0.00 |
| 37.25 | 0.00 | 0.0186 | 0.69 | 0.00 |
| 39.24 | 0.00 | 0.0373 | 1.46 | 0.00 |
| 41.35 | 0.00 | 0.0644 | 2.66 | 0.00 |
| 43.57 | 0.00 | 0.0966 | 4.21 | 0.00 |
| 45.90 | 0.00 | 0.1264 | 5.80 | 0.00 |
| 48.36 | 0.00 | 0.1445 | 6.99 | 0.00 |
| 50.96 | 0.96 | 0.1445 | 7.36 | 0.14 |
| 53.69 | 3.69 | 0.1264 | 6.79 | 0.47 |
| 56.57 | 6.57 | 0.0966 | 5.47 | 0.63 |
| 59.60 | 9.60 | 0.0644 | 3.84 | 0.62 |
| 62.80 | 12.80 | 0.0373 | 2.34 | 0.48 |
| 66.17 | 16.17 | 0.0186 | 1.23 | 0.30 |
| 69.71 | 19.71 | 0.0079 | 0.56 | 0.16 |
| 73.45 | 23.45 | 0.0029 | 0.21 | 0.06 |
| 77.39 | 27.39 | 0.0008 | 0.06 | 0.02 |
| 81.54 | 31.54 | 0.0002 | 0.01 | 0.00 |
| 85.91 | 35.91 | 0.0000 | 0.00 | 0.00 |
| 90.52 | 40.52 | 0.0000 | 0.00 | 0.00 |
| 95.37 | 45.37 | 0.0000 | 0.00 | 0.00 |
| 100.49 | 50.49 | 0.0000 | 0.00 | 0.00 |
| 105.88 | 55.88 | 0.0000 | 0.00 | 0.00 |
| Column sums | | 1.0000 | 50.14 | 2.90 |

$s_{1,1}$ (the initial stock price) was \$50, and $t$ (the time remaining) was 10 days (10.0/365.25 years). The strike price was assumed to be \$50, as was the case for the Monte Carlo analysis. There were 30 layers ($n = 30$) in the tree that was used to generate the data in Table 3–2. It can easily be seen that the results from the binomial tree are remarkably close to those obtained from the Monte Carlo method when 30 time steps were simulated.

Although an option has just been appraised directly from the terminal stock prices and their probabilities, this is not the standard or the most flexible way to work with a binomial tree. The second step in the pricing process generally involves working backwards through the tree, recursively, starting with the final column of nodes representing the terminal distribution of stock prices. As before, the calculations are performed layer by layer, with entries in the $j$-th layer being determined from the transition parameters and the entries in the $j + 1$-th column of nodes. What is calculated in this instance are the one-step forward expectations, discounted for growth. Forward expectations, of course, represent theoretical option valuations. In Figure 3–2, below, these are identified by $\xi$.

In general,

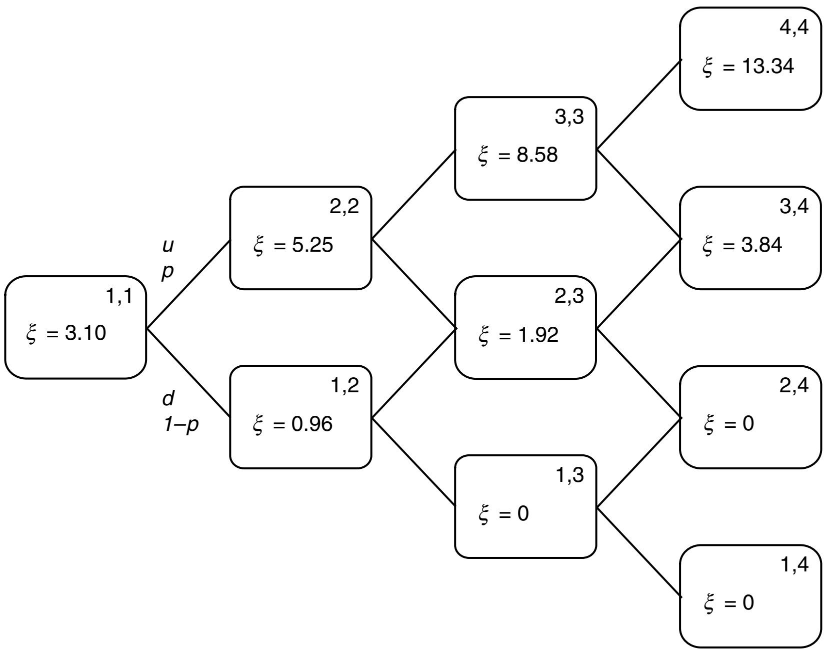
$$\xi_{ij} = e^{-r\Delta t}\left[p\xi_{i+1,j+1} + (1-p)\xi_{i,j+1}\right] \tag{3.9}$$

i.e., for each pair of nodes to which a given node has transitions, multiply the transition probabilities by the valuations of the target nodes and then discount for growth in a risk-neutral world. For the last layer of nodes, the option is expiring (usually) and the valuations ($\xi$) are easily determined. For an expiring call, $\xi$ is the greater of 0 or $s - k$, where $s$ is the stock price and $k$ the strike. For an expiring put, $\xi$ is the greater of 0 or $k - s$. In all earlier layers, $\xi$ may be determined recursively, using Equation 3.9. When the first layer is reached, the option has been priced. The fair value of the option is $\xi_{1,1}$ (the $\xi$ associated with the initial node in the binomial pricing tree).

In actual fact, there are direct formulae for finding the stock prices and traversal probabilities for any of the nodes in a binomial tree, including the terminal ones, without the need for any recursive, layer-by-layer calculations. There is also a direct

**FIGURE 3–2**

The Same Binomial Tree with Node Indices and Option Valuations




formula for the value of an option, given the prices and probabilities at the terminal nodes.

$$\phi_{in} = \binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} \tag{3.10}$$

$$s_{in} = s_{1,1}u^{i-1}d^{n-i} \tag{3.11}$$

$$\xi_{1,1} = e^{-(n-1)r\Delta t}\sum_{i=1}^{n}\phi_{in}\xi_{in} \tag{3.12}$$

These formulae assume that the tree is recombining, and that the transition parameters $p$, $u$ and $d$ are constant throughout the tree. The reader may recognize Equation 3.10 as the binomial probability density function, with parameters modified for compatibility with the indexing scheme used to identify the nodes. The results obtained from these equations will be identical (ignoring roundoff error) to those computed by the recursive approach. Tests using Fortran routines to compute binomial pricings bear this out.

Although direct formulae (such as Equations 3.10 through 3.12) are available, there is a great loss of flexibility when using them. For instance, there are no definitive direct solutions for pricing options with early exercise rights, the so-called American-style options. However, options with early exercise rights are readily priced working back layer-by-layer through a binomial pricing tree using Equation 3.9. The procedure is quite straightforward: after calculating the forward expectations for all nodes in a given layer, comparisons are performed. If the value of exercise for a given node is greater than the forward expectation, which represents the value of holding another time step, then assume that the option is exercised at that node and replace the node's $\xi$ with the option's exercise value. Do this for all nodes in the layer. Once done, each $\xi$ in that layer will represent either the value of exercise or the value of holding, whichever is greater. Once all the values have been determined, apply Equation 3.9 to obtain the forward expectations for the next layer back. Repeat the process until the initial node is reached and the American-style option is priced.

The flexibility of the method of binomial trees makes it feasible to price even exotic derivatives and to price options under assumptions involving nonconstant interest rates and volatility. No doubt, it is this power and flexibility that has contributed to the enduring popularity of the method of binomial pricing trees. An extensive and reasonably good treatment of the Cox-Ross-Rubinstein binomial pricing model and its many variations and applications can be found in *Black-Scholes and Beyond* (Chriss, 1997).

## THE BLACK-SCHOLES MODEL

Consider the binomial pricing model presented above, where $p$, $u$, and $d$ are constant at all time steps throughout the pricing tree. In such a model, as $n$, the number of time steps, approaches infinity, the distribution of prices at the terminal nodes approaches log-normal. Even for modest $n$, the approximation is quite good, as can be seen in Figure 3–3. Figure 3–3 shows the log-normal probability density (solid line) plotted along with the discrete probabilities associated with the terminal nodes of a 30-step binomial tree (dotted line with markers). The terminal node probabilities used in the figure were taken from the third column of Table 3–2. Note the great similarity between the distributions. By the time $n$ reaches 200, the curves are visually indistinguishable.

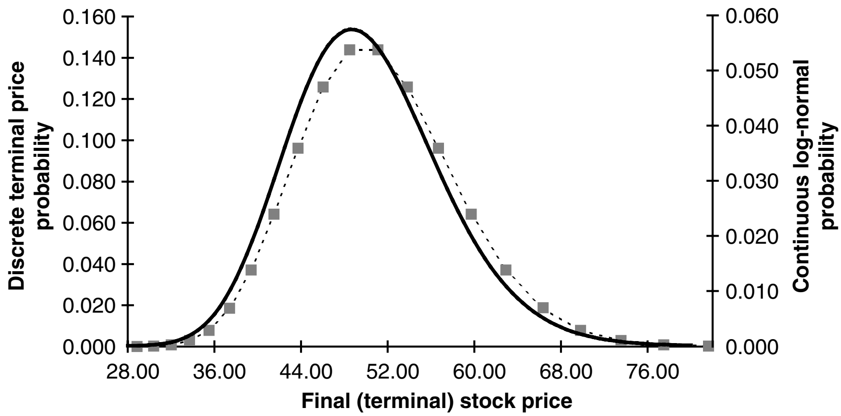
Geometric binomial random walks, such as those described by binomial pricing trees, are not the only random walks that converge to log-normal price distributions. Consider a model where

$$R_n = \frac{s_n}{s_0} = \prod_{i=1}^{n} \exp\left(\frac{\sigma\, x_i}{\sqrt{n}} + \frac{u}{n}\right) \tag{3.13}$$

and where $s_n$ and $s_0$ are the final and initial stock prices, respectively. If $\sigma$ and $u$ are finite, and if $\{x_1 \ldots x_2\}$ is a sequence of independent, identically distributed random variables with a

**FIGURE 3–3**

Discrete Terminal Price Probabilities from 30-Step Binomial Tree versus Continuous Log-Normal Probabilities from Probability Density Function



common expectation of 0 and variance of 1, then it can be shown that

$$F_{R_n}(x) \to \frac{1}{\sigma\sqrt{2\pi}} \int_0^x t^{-1} \exp\left[-\frac{1}{2}\left(\frac{\ln t - u}{\sigma}\right)^2\right] \mathrm{d}t \tag{3.14}$$

as $n \to \infty$. The right hand side of Equation 3.14 will be recognized as the log-normal probability integral. Equation 3.13 describes the taking of $n$ random proportional steps, the multiplier at each step being the exponential of a random variable with a standard deviation of $\sigma/\sqrt{n}$ and mean of $u/n$ and having an arbitrary distribution. The values $\sigma$ and $u$ are the volatility and growth, respectively.

The proof is actually quite simple. Taking the logarithms of both sides of Equation 3.13 gives

$$\begin{aligned} \ln R_n &= \ln \frac{s_n}{s_1} = \ln \frac{s_n}{s_{n-1}} + \ln \frac{s_{n-1}}{s_{n-2}} + \ldots + \ln \frac{s_2}{s_1} \\ &= \sum_{t=1}^{n} \ln\left(\frac{\sigma x_i}{\sqrt{n}} + \frac{u}{n}\right) \end{aligned} \tag{3.15}$$

In other words, the return (expressed as a logarithm) is the sum of $n$ random shocks. As a consequence of the Central Limit Theorem, the right hand side of Equation 3.15 will approach a normal distribution with mean $u$ and standard deviation $\sigma$ as $n$ approaches infinity. What this means is that, if enough small shocks are added together, then the distribution of their sum will be normal in the limit as $n \to \infty$, even if their individual distributions are not normal. By definition, if $R_n$ is a log-normal random variable, then $\ln R_n$ is normally distributed, and vice versa. It has just been shown that, indeed, $\ln R_n$ is normally distributed, thus demonstrating that $R_n$ has a log-normal distribution.

In more trader-friendly language, what this means is that, if the moment-to-moment proportional movements in a stock's price are random, and unpredictable on the basis of previous movements, and if each moment's movement is essentially like any other moment's movement in a statistical sense, then, given an initial price, the probability of any final price being reached after some specified interval will be arbitrarily well approximated by the log-normal probability density function. Since, in an efficient market, moment-to-moment movements are presumably close to random, the price movements over tradable intervals (which are made up of many such moment-to-moment movements) might be expected to be close to log-normal in their distribution. Even in a market that is only moderately efficient, a trader might expect prices to exhibit behavior that is at least a reasonable approximation to log-normal. Given such theoretical observations, the trader might even consider pricing options on the basis of geometric, log-normal random walks.

## The Black-Scholes Formula

A closed form solution actually exists for pricing European-style options on stocks having log-normal terminal price distributions. It is the famous Black-Scholes formula, which is shown below for both calls and puts. The fair value for a call is given by

$$\xi_{\text{call}} = sN(d_1) - ke^{-rt}N(d_2) \tag{3.16}$$

and the fair value for a put is given by

$$\xi_{\text{put}} = -sN(-d_1) + ke^{-rt}N(-d_2) \tag{3.17}$$

where

$$d_1 = \frac{\ln\left(\frac{s}{k}\right) + \left(r + \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}} \tag{3.18}$$

and

$$d_2 = d_1 - \sigma\sqrt{t} \tag{3.19}$$

and where $s$ is the current stock price, $k$ the strike price, $r$ the risk-free interest rate, $\sigma$ the volatility, $t$ the time left until expiration, $\xi$ the option value, and $N$ the cumulative normal probability function. These formulae satisfy the put-call parity relationship

$$\xi_{\text{call}} - \xi_{\text{put}} = s - ke^{-rt} \tag{3.20}$$

expressed in the notation of this chapter. In the above equations, $d_1$ and $d_2$ bear no relationship to the transition multipliers discussed earlier.

## Black-Scholes and Forward Expectation

The Black-Scholes pricing formula was originally developed neither by taking the limit as $n \to \infty$ of the binomial model nor by determining the expectation of option price at expiration under the assumption that stock prices are log-normally distributed at that time. Instead, it was developed from a no-arbitrage argument. Nevertheless, the Black-Scholes formula is a direct solution for the simple binomial option price as $n \to \infty$ and for the valuation based on expectation at expiration, given an underlying stock having a log-normal distribution of terminal prices. Although no mathematical proof is offered here, it is instructive to demonstrate these claims numerically. Consider

$$\xi_{0,k}(s_0) = e^{-rt} \int_0^\infty f_{u,\sigma}(x) \xi_{n,k}(x)\, \mathrm{d}x \tag{3.21}$$

where

$$f_{u,\sigma}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left( -\frac{1}{2} \left[ \left( \ln x - \ln u + \frac{1}{2}\sigma^2 \right) \Big/ \sigma \right]^2 \right) \tag{3.22}$$

is the log-normal probability density function that describes the distribution of terminal stock prices, where $\xi_{0,k}$ is the sought-after theoretical option price, $s_0$ the initial stock price, $k$ the strike price, $\xi_{n,k}(x)$ the value of the option at expiration, and $x$ the stock price at expiration.

Equation 3.21 will be recognized as defining expectation for the option at expiration, appropriately discounted for growth in a risk-neutral world. Essentially, the equation is a continuous version of Equation 3.12, in which the sum has been replaced with an integral and the discrete probability density with a continuous density function. In the equation, $u = s_0 e^{rt}$ and $\sigma = v\sqrt{t}$

with $t$ being the time remaining until expiration (measured in years) and $v$ the standard, annualized volatility.

Numerical integration, or *quadrature* as it is sometimes called, can be used to solve for $\xi_{0,k}(s_0)$ in Equation 3.21. This was actually done using the same parameters employed when demonstrating the binomial model. The volatility was set to 85% ($v = 0.85$), the time remaining to 10 days ($t = 10/365.25$), the interest rate to 10% ($r = 0.10$), the strike price to \$50 ($k = 50$), and the initial stock price to \$50 ($s_0 = 50$). The value obtained for $\xi_{0,k}(s_0)$ was \$2.8681983 for the call and \$2.7314929 for the put. For the purpose of comparison, the same parameters were entered into the standard Black-Scholes formula. The option valuations obtained from Black-Scholes agree with those computed by numerical integration of Equation 3.21 to at least the eight significant digits displayed by the Fortran program! Several other sets of parameters were tried with similar outcomes. Likewise, when the binomial model was taken to 3,000 steps, the resultant option prices for a stock with the same parameters agree quite well with the results from Black-Scholes and from numerical integration. Priced on a 3,000-layer binomial tree, the call was \$2.86814 and the put \$2.73144. Monte Carlo data based on the Fortran routine `MCBRW` presented earlier are also in close agreement with option prices obtained from Black-Scholes, numerical integration, and binomial trees. There is no doubt that Black-Scholes is a solution for the forward expectation, discounted for interest, of an option trading on a stock having a log-normal distribution of terminal prices and an expected growth equal to the rate of risk-free interest; and, as $n \to \infty$, so is the Cox-Ross-Rubinstein binomial model.

### Black-Scholes versus Binomial Pricing

Black-Scholes has become the dominant pricing model because it is very easy to calculate both option prices and the Greeks and because it gives acceptably good results. However, Black-Scholes is not flexible and, in its original form, is unable to correctly price American-style options. Although approximations for American-style options have been developed, in many instances there is little practical difference in price between options with early exercise rights and those without, and hence little need to go beyond standard Black-Scholes. Long-term options such as

LEAPS, and options that are deeply in-the-money, do require early exercise to be taken into account if they are to be correctly priced, hence the continued popularity of the binomial model despite its time-consuming calculations. The binomial model is adaptable, handles early exercise with grace, and can be used to price all kinds of exotic options. With suitable adjustments, it can even be used to price options under assumptions of nonconstant volatility, dividends, and interest rates, and of distributions that differ from log-normal. Given today's computing power, the numerical workload is virtually a nonissue. Nevertheless, Black-Scholes continues to be used by many traders and hedgers.

## MEANS, MEDIANS, AND STOCK RETURNS

Before closing, it is worth examining some properties of the log-normal distribution, and, by implication, of other distributions that approach or approximate it. For the log-normal distribution, the median is always less than the mean. This has some significant implications for stock prices and returns. Imagine a stock with an initial price of \$100. If the set of potential terminal prices is described by a log-normal distribution with a median of \$100, then the stock would be expected to finish above \$100 as frequently as it would finish below \$100. That is the definition of the median: the fifty-fifty or midway point. Naively, this might be construed as a break-even situation. However, it is not! With the median at \$100, the stock's mean or expected terminal price will be greater than \$100. This is a positive return, a net profit for the trader when considered over a sufficient number of trades. It is due to the asymmetry of the log-normal distribution, which has a truncated left tail and an elongated right one.

If the expectation for the final stock price is to be \$100 (a true break even situation), then the median must be less than \$100 by some amount. The amount depends on the difference between the median and the mean, which, in turn, depends on the volatility. The relationship between volatility ($\sigma$), the median ($\mu$), and the mean ($E(s_n)$, where $E$ is the expectation operator) for prices ($s_n$) that have a log-normal distribution is defined by

$$E(s_n) = \mu \exp\left(\frac{\sigma^2}{2}\right) \tag{3.23}$$

Therefore, if a mean terminal stock price of \$100 is to be obtained, the parameter $\mu$ of the standard log-normal probability density function must be suitably adjusted. The parameter $\mu$ must be set, not to 100, but to 100 divided by $\exp(\sigma^2/2)$ or, equivalently, 100 multiplied by $\exp(-\sigma^2/2)$. Such an adjustment was incorporated into Equation 3.22 (note the $\sigma^2/2$ term in the exponential) to allow the parameter $\mu$ in $f_{\mu,\sigma}(x)$ to be identified with the expectation or mean, rather than with the median of the log-normal probability density described by the function. In the above example, after adjusting $\mu$, the expectation for the terminal price will be \$100. This is exactly as intended. However, after the adjustment, the stock price will have a slightly greater probability of falling than of rising.

How did the parameter $\mu$ become identified with the median? If $x$ is a random variable with a normal distribution, than $\exp(x)$ is a random variable with a log-normal distribution. Consequently, every log-normal distribution has a corresponding or underlying normal distribution. The logarithm of $\mu$ in the standard formula for the log-normal probability density is identified with the mean of the underlying normal distribution, and $\sigma$ with the standard deviation of that underlying distribution. For any symmetric distribution, including the normal distribution, the median and mean are equal. Hence, the logarithm of $\mu$ is not just the mean of the underlying normal distribution, it is the median as well. Let $x$ be a random variable with a nonzero standard deviation. It can be shown that $M(\exp(x)) = \exp(M(x))$ but that $E(\exp(x)) > \exp(E(x))$ where $M$ is the median operator and $E$ the expectation. The result is that if the logarithm of $\mu$ represents both the median and the mean of the underlying normal distribution, then $\mu$ will represent the median, but not the mean, of the corresponding log-normal distribution.

## Empirical Study of Returns

How do real stocks behave? The trader would probably argue that, in the absence of predictive information, a stock has a roughly equal probability of either rising or falling, at least over a short interval. Stocks that are more volatile would not be expected to be any different in this regard than less volatile ones. Yet, if the expected price in the near future were the same as a stock's current price, then the stock would have to have a higher probability of falling rather than rising in price. Furthermore, the greater the stock's volatility, the more exaggerated the relative likelihood of a downside move. If the trader is right in arguing that the probability of a gain is roughly equal to the probability of a loss, then the expectation for the near-future price should be greater than the current price, i.e., the stock should have a positive expected return, and that return should increase with increasing volatility. Do stock returns have positive expectations that grow with volatility? Would not such returns represent inefficiencies in a supposedly efficient market–profit opportunities that could be exploited by the astute trader?

Table 3–3 contains data relevant to these questions. The first section of the table shows the results from a simple Monte Carlo experiment. This experiment examines the theoretical returns for several volatility levels and holding periods. In the experiment, the distribution of price is assumed to be log-normal and the median of the terminal price is assumed to equal the initial price. The numbers are average price gains measured in percent; they represent the return in dollars from the purchase of a nominal \$100 stock. Consistent with the theoretical analysis, quite substantial gains are generated and these gains increase with annualized volatility and holding period.

The last two sections of Table 3–3 show data similar to the first section, but for real stocks rather than simulated, Monte Carlo stocks. Small, thinly-traded NASD stocks are examined in the first of these two sections, while higher-priced, more widely followed NYSE stocks are examined in the second. For both sets of stocks, returns are positive and tend to increase with annualized volatility and holding period. In other words, there is a *volatility payoff.* Do stock returns have a positive expectation that grows with volatility? The answer is a resounding "Yes."

For very low levels of historical volatility, real stocks deliver higher returns than simulated Monte Carlo stocks with equivalent theoretical volatilities. The explanation is regression to the mean, or mean reversion, a well-known statistical phenomena. When historical volatility is very low, regression to the mean will lead to greater volatility in the period from which the return

**TABLE 3–3**

Expected Percentage Gain in Price as a Function of Holding Period and Volatility

**Log-Normal Geometric Random Walk (Monte Carlo Data)**

| | *Volatility (Annualized Percent)* | | | | | | |
|---|---|---|---|---|---|---|---|
| **Days Held** | **50** | **100** | **150** | **200** | **250** | **300** | **350** |
| 5 | 0.2 | 1.0 | 2.3 | 3.9 | 6.4 | 9.4 | 12.7 |
| 10 | 0.5 | 2.0 | 4.4 | 8.2 | 13.3 | 19.7 | 28.0 |
| 15 | 0.8 | 3.1 | 6.9 | 12.5 | 20.3 | 31.2 | 43.3 |
| 20 | 1.0 | 4.2 | 9.5 | 17.6 | 28.7 | 43.0 | 61.7 |
| 25 | 1.3 | 5.1 | 11.8 | 22.0 | 36.6 | 55.8 | 82.5 |

**All NASD Stocks with Volume < 200,000 Shares and Price < $4**

| | *Volatility (Annualized Percent)* | | | | | | |
|---|---|---|---|---|---|---|---|
| **Days Held** | **50** | **100** | **150** | **200** | **250** | **300** | **350** |
| 5 | 0.7 | 1.3 | 1.9 | 2.8 | 4.3 | 6.1 | 6.6 |
| 10 | 1.4 | 2.5 | 3.4 | 5.2 | 8.1 | 9.1 | 13.7 |
| 15 | 2.1 | 3.5 | 5.0 | 7.9 | 10.9 | 12.8 | 18.4 |
| 20 | 2.8 | 4.7 | 6.5 | 10.1 | 14.2 | 17.6 | 18.6 |
| 25 | 3.6 | 6.0 | 7.9 | 12.5 | 16.8 | 21.9 | 20.9 |

**All NYSE Stocks with Volume > 400,000 Shares and Price > $35**

| | *Volatility (Annualized Percent)* | | | | | | |
|---|---|---|---|---|---|---|---|
| **Days Held** | **50** | **100** | **150** | **200** | **250** | **300** | **350** |
| 5 | 0.3 | 0.9 | 2.1 | 1.5 | 3.1 | 4.8 | 1.9 |
| 10 | 0.7 | 1.7 | 3.9 | 2.7 | 5.4 | 6.2 | 0.7 |
| 15 | 1.0 | 2.5 | 5.6 | 5.1 | 7.8 | 7.4 | 4.5 |
| 20 | 1.4 | 3.4 | 6.8 | 8.3 | 8.7 | 9.8 | 5.0 |
| 25 | 1.7 | 4.3 | 8.3 | 10.0 | 12.0 | 14.0 | 5.3 |

is calculated than in the preceding period that is used to calculate the historical volatility. Mean reversion also partly explains the lower relative returns for real stocks, when compared to Monte Carlo simulacra, in situations where historical volatility is very high. Another factor that contributes to relatively lower returns from real stocks when volatility is high is exploitation of the volatility payoff by traders. Therefore, the answer to the question, “Would not such [excessive] returns represent inefficiencies in the market and profit opportunities that could be exploited by the astute trader?” is also “Yes.”

Although exploitation of inefficiencies acts to attenuate returns, there are still rewards for taking on volatility risk. In addition, when traders enter the market and drive out inefficiencies, they also alter the distribution of returns and the observed volatility. The very fact of excessive volatility indicates that traders who exploit inefficiencies induced by high volatility have not yet entered the fray in sufficient numbers to bring the volatility down and to neutralize the inefficiency. If they had, the volatility would no longer be excessively high. On larger, more widely followed NYSE issues, the mean volatility to which recent historical volatility tends to revert is lower and there are more astute traders around to drive out any inefficiency, volatility-induced or otherwise. This is one possible explanation for the generally smaller returns seen in the NYSE data. No doubt “small-cap” and “ignorance” effects, well-known in the field, also play a role in accounting for the differences in returns between the two groups of stocks compared in Table 3–3.

## SUMMARY

In this chapter, the two most popular option pricing models were dissected and their inner workings explored. The Cox-Ross-Rubinstein binomial model was the first to be examined. The discussion began with the crude Monte Carlo example from Chapter 2. In that model, stock prices were assumed to increase or decrease randomly by $1 each and every day. Several modifications to this model were made: allowing the size of each random movement to be specified, permitting the size of the up moves to differ from the size of the down moves, and making the movements multiplicative or geometric, rather than additive or arithmetic. These changes led directly to the equal probability variant of the Cox-Ross-Rubinstein binomial model. The model was implemented using Monte Carlo, just as the crude model in Chapter 2. The Monte Carlo implementation was used to price a hypothetical option.

The elegant method of binomial trees was then presented. This method was demonstrated to be an extremely flexible and powerful approach to implementing the Cox-Ross-Rubinstein and related models. The same option previously priced with Monte Carlo was priced on a binomial tree with similar results. Pricing American-style options (those with early exercise rights) with binomial trees was also discussed to illustrate the flexibility of the method.

Some of the assumptions underlying the Cox-Ross-Rubinstein model were examined. Expectation at expiration (the mean option price at the time of option expiration) was also considered, given the assumption that the terminal stock price has a distribution that is, or approximates, log-normal. This led directly to an alternative analysis of Black-Scholes that demonstrated that the model is actually a closed form solution for expectation at expiration, one that factors in growth of the underlying stock and discounts the option for the holding interval. The standard Black-Scholes formula was then used to price the same option priced earlier using Cox-Ross-Rubinstein in both its tree-based and Monte Carlo implementations. Finally, using numerical quadrature, the expectation of that same option's price at expiration (with appropriate handling of growth and discounting in a risk-neural world) was determined and shown to be the same as the Black-Scholes price. The price from both the direct numerical integration of the expectation equation and from Black-Scholes agreed in every digit displayed.

In the course of these analyses, it was found that, although each model was arrived at by a different route and has a distinctive surface appearance, they all estimate essentially the same variables, express similar assumptions, and produce virtually identical option prices. The Cox-Ross-Rubinstein model does so as the number of time steps in the binomial tree grows large, whereas the Black-Scholes model does so immediately, without requiring any specification of, or attention to, time steps. Both models essentially calculate the option's average price at expiration, discounted for interest under the assumption that the underlying stock exists in a risk-neutral world and has, or closely approximates, a log-normal distribution of terminal prices (prices at the time of option expiration). By saying the stock and the option exist in a risk-neutral world means that both are assumed to have growth rates equal to the prevailing rate of risk-free interest.

The only significant practical difference between the models, as implemented, was found to be flexibility. The binomial model is extremely flexible and can be used to price all kinds of options, including exotics such as barrier options, as well as options with early exercise rights. The flexibility extends to the point that the method of binomial trees can be modified to handle nonconstant volatility, interest rates, and dividends, including ones that are contingent upon the prior course of events. It can even be adapted to terminal distributions that differ from log-normal. As such, the Cox-Ross-Rubinstein model and the method of binomial trees provide a very general and powerful approach to pricing options. However, in its standard or basic form, it yields basically the same results as Black-Scholes. Black-Scholes has little, if any, flexibility, but is extremely fast and easy to use. It is popular because, for practical purposes, even when dealing with options that have early exercise rights not accounted for in the model, it gives results that are not too far off the mark.

Finally, some characteristics of the log-normal distribution, upon which these models explicitly or implicitly rely, were examined. One such characteristic is that the mean of a log-normal distribution is greater than its median, or 50th percentile. Another characteristic is that the extent to which the mean is greater than the median is directly related to volatility. In other words, if it is assumed that a stock has an equal chance of rising or falling over the holding period, then a positive return that will increase with increasing volatility—a volatility payoff—must also be assumed.

Real data were examined to see if returns were consistent with a log-normal distribution of terminal prices having a median equal to the initial stock price. Actual market data demonstrated that the effect described above was indeed present to some degree: there were volatility-related positive returns. It should be noted that a strictly log-normal distribution is not required for returns to be related to volatility; a volatility payoff may be observed with any distribution that has a longer right-hand tail than a left-hand one and for which increasing volatility or distributional spread results in a greater difference between the mean and the median. The question of whether stock prices really behave as these models assume—i.e., as geometric random walks that yield log-normal distributions of returns—is examined more fully in the next chapter.

## SUGGESTED READING

*Black-Scholes and Beyond* (Chriss, 1997) provides thorough coverage of the popular Black-Scholes and Cox-Ross-Rubinstein pricing models, including the use of binomial trees to value exotics. A handy source for basic numerical algorithms, including algorithms for psuedo-random number generation and for numerical quadrature, is *Numerical Recipes in Fortran 77* (Press et al., 1992). More material on the subject of numerical algorithms, with a focus on ultra-high-precision calculations, can be found in *Precise Numerical Methods Using C++* (Aberth, 1998). *Continuous Univariate Distributions* (Johnson et al., 1994) provides reasonable coverage of the normal and log-normal distributions. Also of possible interest is *Monte Carlo Simulation* (Mooney, 1997).

CHAPTER 4

# Statistical Moments of Stock Returns

There are two good reasons to study the statistical moments of stock returns. The first reason to investigate moments is to check the assumptions made by popular option pricing models. Popular pricing models make some heavy assumptions regarding the statistical nature of price movements. As demonstrated in the previous chapter, Black-Scholes and Cox-Ross-Rubinstein assume a geometric random walk with a log-normal distribution of returns. But do stock prices really follow a random walk? Is each price change or return statistically independent of all others, as the random walk theory would imply? Are the distributions of returns from real stocks log-normal? One way to answer these questions is by examining the statistical moments of actual returns. A statistical analysis involving moments may reveal whether the price behavior of real stocks violates any assumptions made by the standard models and, if so, how.

The second reason to study returns from the perspective of moments is to better characterize the distributions involved and to learn how these distributions are affected by circumstances. If not log-normal, then what are the distributions of returns? How are they affected by holding period, day of week, time of year, and other factors? This breaches the topic of conditional distributions, which are discussed in a later chapter. Option prices are undoubtedly determined by the terminal price distributions,



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

real or imagined, of the underlying stocks. All pricing models make some assumptions regarding such distributions. Given these facts, it is clearly worthwhile to empirically characterize the distributions of stock returns, both conditionally and generally. Such characterizations can help in the quest for better pricing models and for ways to improve the existing ones.

This chapter marks a transition from the primarily theoretical and explanatory stance of earlier chapters to the more empirical and practical viewpoint of later ones. The focus is now on empirical studies designed to answer specific questions. Some of these studies have to do with elucidating the circumstances under which popular models break down and grossly misprice options. Others concern the quest for a deeper statistical understanding of the market and for more accurate option pricing, generally. The ultimate objective is to devise pricing models that have practical value for traders and that are not merely of theoretical interest.

## THE FIRST FOUR MOMENTS

Distributions of returns may be characterized by statistics called *moments*. The first four moments of a distribution are the mean, variance, skew, and kurtosis. The *mean*, or expectation, reflects a distribution's central tendency. In pricing models such as those discussed in Chapter 3, the mean appears as a growth rate related to the rate of risk-free interest. *Variance* is a measure of a distribution's width or range. The square root of variance is the *standard deviation*, known as *volatility* in the world of options. Volatility is a major determinant of option value; as such, it is a statistic required by virtually every pricing model.

While the mean and variance (or standard deviation) describe the location and size of a distribution, skew and kurtosis characterize its shape. *Skew* measures relative asymmetry. Positive skew characterizes a distribution that has, relative to the normal distribution, a shorter left (negative) tail, a longer right (positive) tail, and a peak that appears tilted to the left. Negative skew describes a distribution that has a longer left tail, a shorter right tail, and appears tipped to the right. *Kurtosis* is a statistic reflecting the relative flatness or peakedness of a distribution. A distribution with positive kurtosis has a sharp peak and long tails while one with negative kurtosis is more rectangular, with a flatter top and shorter tails. In stark contrast to growth and volatility, skew and kurtosis are ignored by standard pricing models. However, skew and kurtosis are as readily computed and analyzed as growth and volatility. All four moments provide useful information about the distribution of price movements in stocks and indices and could thus play a role in more advanced option pricing models.

Moments may be calculated from a theoretical distribution or from a sample of data. When calculated from a sample they are known as *sample moments*. Sample moments are statistical estimates of parameters—either parameters governing the underlying or generating distribution, or those that describe the population from which the sample was drawn.

## Calculating Sample Moments

Let $\{x_1, x_2, \ldots x_n\}$ be a sample of $n$ distinct data points. The first moment is the sample mean, which is just the average of the data points in the sample.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.1}$$

The sample variance, skew, and kurtosis are the second, third, and fourth moments about the mean, respectively. They are calculated as

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{4.2}$$

$$\text{skew} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3 \tag{4.3}$$

$$\text{kurtosis} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^4\right] - 3 \tag{4.4}$$

where $s$ is the sample standard deviation (the square root of the sample variance, $s^2$). As apparent from Equations 4.2 to 4.4, *moments about the mean* are averages of powers of the deviations about the mean. For skew and kurtosis, the deviations from the mean are normalized. Normalization is accomplished by dividing each deviation by the standard deviation. Because of such normalization, skew and kurtosis are independent of scale and sensitive only to form. The reason for subtracting 3 in Equation 4.4 is that, when the distribution underlying the sample of data points is normal, kurtosis will approach zero as $n$ grows large.

The above equations require the mean to be calculated prior to any of the other moments. Calculating the mean first can be quite cumbersome for the programmer and can increase processing time due to the need for two passes over the data. An alternative method allows the moments to be calculated directly from the sums of powers of the data points. The method requires only one pass to accumulate the required sums (the $a_k$ in Equation 4.5). Be warned, however, that this method is far more susceptible to round-off error. When writing computer code to calculate moments with the alternative method, do all sensitive arithmetic with double precision! The alternative solution, based on binomial expansions of the terms in Equations 4.1 to 4.4, is as follows:

$$a_k = \sum_{i=1}^{n} x_i^k \tag{4.5}$$

$$\bar{x} = \frac{a_1}{n} \tag{4.6}$$

$$s^2 = \frac{n}{n-1}\left(\frac{a_2}{n} - \bar{x}^2\right) \tag{4.7}$$

$$\text{skew} = \frac{1}{ns^3}(a_3 - 3a_2\bar{x} + 3a_1\bar{x}^2 - \bar{x}^3) \tag{4.8}$$

$$\text{kurtosis} = \frac{1}{ns^4}(a_4 - 4a_3\bar{x} + 6a_2\bar{x}^2 - 4a_1\bar{x}^3 + \bar{x}^4) - 3 \tag{4.9}$$

## Statistical Features of Sample Moments

Suppose that the underlying distribution from which a sample is drawn has a finite mean and variance, and the sample data points are statistically independent. For large $n$, the sample mean will then have a distribution that approximates the normal and, according to the Central Limit Theorem, the approximation will become exact as $n$ approaches infinity. When the underlying distribution is normal, the sample mean will be an unbiased estimate of the *population* mean (the mean of the underlying distribution). The standard deviation of the sample mean will be $\sigma/\sqrt{n}$, where $\sigma$ is the standard deviation of the underlying distribution; it will be well approximated by $s/\sqrt{n}$ for even modest sample sizes. If the underlying distribution is normal, then the sample variance will be an unbiased estimate of the variance of the underlying distribution and will have a chi-square distribution with $n$–1 degrees of freedom. For sufficiently large $n$, the sample standard deviation will have a distribution that approaches normal and a standard deviation that will be close to $s/\sqrt{2n}$.

Statements similar to those above can also be made about sample estimates of skew and kurtosis. Skew and kurtosis should have values that are not significantly different from zero if the underlying distribution is normal. The standard deviation for large $n$ will be approximately $\sqrt{6/n}$ for sample skew and $\sqrt{24/n}$ for sample kurtosis when estimates are made using the sample mean. If the mean is known, and is set independently of the sample, then skew and kurtosis will have standard deviations of approximately $\sqrt{15/n}$ and $\sqrt{96/n}$, respectively. Note the absence of $s$ or $\sigma$ in these approximations: this is because both skew and kurtosis are scale-independent. As with the first two moments, and for similar reasons involving the Central Limit Theorem, the distributions of sample skew and sample kurtosis will tend toward normal as $n$ grows large; however, distributions exist for which skew and kurtosis have infinite standard deviations and thus no meaningful estimates.

How large must $n$ be for the approximations presented above to be of practical use? Given an $n$ of 25 and a normal underlying distribution, the standard deviation of sample skew has a mean of 0.41, determined by a large-scale, high-precision Monte Carlo experiment, compared to 0.49, as found with the approximation described earlier. Observed and approximated standard deviations for skew are much closer for a sample size of 100: 0.23 for Monte Carlo versus 0.24 for the approximation. Rounded to two decimal places, the mean sample skew for all sample sizes tested by Monte Carlo was 0.00—the expected skew for a normally distributed random variable.

Kurtosis requires a larger sample than skew for the approximations provided above to achieve reasonable accuracy. The standard deviation of kurtosis determined by Monte Carlo was 0.67, while the approximation yielded 0.98, for an $n$ of 25. Agreement between Monte Carlo and the approximation is much better with $n$ at 100 (0.44 versus 0.49) and excellent with $n$ at 400 (0.24 versus 0.25). Mean sample kurtosis values estimated by Monte Carlo were –0.45, –0.12, –0.03, and –0.01 for sample sizes of 25, 100, 400, and 2,000, respectively. As $n$ increases, these mean kurtosis figures approach the zero kurtosis expected for a normal distribution.

The statistics for the sample mean and standard deviation are extremely well approximated even for small sample sizes and hence require no additional discussion.

## EMPIRICAL STUDIES OF MOMENTS OF RETURNS

In the studies that follow, actual stock and index returns are examined from the standpoint of moments. The findings from these studies help in answering a variety of questions relevant to the appraisal of options and in understanding the statistical idiosyncrasies of stock price movements. Some observations made in the course of these studies may be of immediate practical value to the stock and options trader.

The first study examines the sample moments of stock returns for several holding periods. Day-of-week effects are analyzed in the second study. The third study investigates how moments of returns are affected by the time of year, i.e., by seasonal factors. Moments are examined in relation to time left before expiration in the fourth study.

Before going into the details of the individual studies, three items that are common across all investigations must be discussed: raw data, analytic software, and Monte Carlo baselines.

### Raw Data

The raw, historical stock data used in the studies were taken from the Worden Brothers TC-2000 end-of-day database. Data were extracted for all optionable stocks. Care was taken to exclude from the extracted data all nonstock securities such as munis, funds, and trusts, many of which trade on the NYSE.

The sampling period, which ran from January 2, 1996 to April 14, 2003, involved a trade-off between sample size and sample representativeness. A longer period would have contained data that was far less representative of the current market, while a shorter one would have been less adequate statistically. The final data sample included both raging bull (January 2, 1996 through March 24, 2000) and devastating bear (March 24, 2000 through April 14, 2003) markets. Data and sampling issues are thoroughly discussed in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000) and in *Computerized Trading* (Jurik, 1999).

Only stocks that were active (alive, not necessarily traded) on the last day of the period were included in the sample. Of course, not all stocks were active over the entire sampling period. An unbroken series of quotes running from January 2, 1996 to April 14, 2003 was available for stocks that were active over the entire period; for those that were not, the unbroken series spanned only a subset of the specified period. Prices for bars or days prior to the active subset were filled with the earliest active price; this was the first nonzero price on a day when the stock had a trading volume greater than zero. Volumes for bars prior to the active subset were set to zero. For bars or days where a stock was active, but during which no actual trading had taken place, the zero volume figure was replaced with a very small positive number. The tiny positive number was intended as a flag to be used by the analytic software to ascertain whether a stock simply had no trading activity on a given day (infinitesimal positive volume), or whether it was dead or not yet issued (zero volume).

All prices in the database were back-adjusted for splits. Because a split factor was included among the extracted data fields, the original, unadjusted prices were recoverable when needed. The final sample was saved in a simple binary database file for speedy access. Each data record in this binary database had the following fields: *Date*, *Open*, *High*, *Low*, *Close*, *Volume* (in 100s), and *Split Factor*. At this stage, there were 2,246 stocks, each with 1,834 bars (not all active), for a total of 4,119,164 data records.

The extracted data for each stock were checked with a utility akin to the one described in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000). The utility was interactive, allowing the user to make on-the-fly corrections to the data. Not unexpectedly, data errors were found. In most instances, the errors were small and easily corrected. An example would be a quote with a high that was just under the close, but where prices were otherwise apparently correct. In a few instances, there were more substantial errors that could not be as easily or rationally corrected. If only one or two prices were involved, but other nearby prices appeared to be correct, an attempt was made to replace the bad prices with more reasonable ones that were estimated from the context of surrounding prices. When several zero (or other highly deviant) prices were followed by a return to normal ones, the stock was typically "trashed," i.e., the data series containing the errors were flagged and deleted. Also trashed was any stock with more than a few errors of any kind. Given the large number of data points in the sample, it was astonishing that only a small number of errors were found and that only five stocks had errors serious enough to justify deletion. This attests to the quality of the Worden stock data. The corrections and deletions should not noticeably impact the results except to make them more reliable by having eliminated extremely deviant, obviously defective data points. Highly divergent data points that echoed real market behavior, such as crashes or other severe but legitimate price swings—e.g., those that might be caused by earnings surprises or bankruptcy announcements—were not "corrected" or removed from the data. Data points had to be clearly erroneous before they were corrected or the stock exhibiting them was trashed.

## Analytic Software

All number-crunching was done using the C programming language. There were several reasons for choosing C over other languages. One reason was that programs written in standard C are portable across a wide variety of machines, operating systems, and compilers. Other reasons were power, flexibility, and speed: C and C++ are powerful, flexible languages for which modern compilers generate highly optimized code that runs quickly. A final and most important reason for making C the language of choice was the ready access to an immense library of routines that were written and accumulated over the years and that have been used and verified in many projects. The use of existing, thoroughly tested routines saves time and gives bugs less opportunity to creep in and invalidate the results. The software used to perform the analyses reported in this book is available from the authors.

## Monte Carlo Baselines

To provide a baseline for comparison, a Monte Carlo simulacrum was constructed for each stock in the binary database. There were two steps involved in creating each simulacrum.

The first step in the construction of each simulated stock was calculation of the historical volatility at each bar in the data series of the corresponding real stock. Historical volatility may be determined from Equations 4.1 and 4.2, given a correct choice of sample. The sample needed to determine an $n$-period historical volatility at a given bar consists of returns from the previous $n$ bars. When computing historical volatility, logarithmic returns are employed. Hence, to calculate the historical volatility at bar $k$, the sample $\{x_1, x_2, \ldots x_n\}$ referred to in Equations 4.1 and 4.2 would be identified with $\{\ln(p_k/p_{k-1}), \ln(p_{k-1}/p_{k-2}), \ldots \ln(p_{k-n+1}/p_{k-n})\}$ where $p_i$ is the price at the close of the $i$-th bar. To obtain the annualized historical volatility, the sample variance ($s^2$ in Equation 4.2) must be multiplied by 252 (the average number of bars or trading days in a year) and the square root then obtained.

The second step in the construction of each Monte Carlo simulacrum was the generation of a series of random prices with a log-normal distribution of returns. Starting with a price equal to the initial price of the real stock, each successive price in the simulated stock was found by multiplying the previous price by a random number having a log-normal distribution. The $\sigma$ parameter for the log-normal pseudo-random number generator was set to the historical volatility. In this case, volatility expressed in daily, rather than annualized, form was required. To obtain the daily (bar-to-bar) volatility, the standard (annualized) historical volatility was divided by the square root of 252. For each bar $k$ of each simulated stock, the historical volatility of the corresponding real stock at bar $k + n/2$ was used in the calculation. The goal was to have the underlying volatility of each simulated stock at any given bar be equal to the observed volatility of the corresponding real stock over a period of $n$ bars centered on that bar. No effort was made to simulate growth due to interest in a risk-neutral world (minimal over the holding periods considered here) or to emulate trends that sometimes seem to appear in real stock prices.

From the above steps, it should be clear that each simulated stock had a local volatility that tracked a centered moving average of the volatility of the corresponding real stock. It should also be apparent that each simulated stock was generated by a known log-normal random process and so is correctly described as a geometric random walk with a log-normal distribution of returns. As such, the theoretical growth, skew, and kurtosis for each simulated stock over any short interval should be close to zero, except for the small effect of a slowly varying level of volatility. By design, the simulacra had volatility levels similar to real stocks; but, possibly unlike real stocks, they also fully satisfied the assumptions made by most popular pricing models, including those discussed in Chapter 3. That, in fact, is what makes these simulated stocks good baselines for gauging the degree to which real stocks violate the standard assumptions.

Returns from both real stocks and simulacra were analyzed side by side in many of the tests reported below. By performing parallel tests, distortions produced by normalization and sampling procedures can be assessed and compensated for in the analysis of the results.

## STUDY 1: MOMENTS AND HOLDING PERIOD

The aim of this study was to determine whether successive returns are statistically independent, whether such returns are log-normal in their distribution, and, if not, how the distribution of returns differs from log-normal. If successive returns are statistically independent, then volatility should scale as the square root of time; e.g., quadrupling the holding period should double the observed volatility. If the distribution of dollar returns is log-normal, then, when expressed in logarithmic form, such returns should have no skew or kurtosis. The presence of notable skew or kurtosis would indicate a deviation from a log-normal distribution of returns.

Two different methods of analysis were employed in this study. The first method was a segmented analysis in which moments were computed for each of a large number of segments and then averaged. The second method involved a more direct analysis in which moments were calculated all at once from the complete set of returns. Regardless of the method, returns from five different holding periods were analyzed. The shortest holding period was one bar; the longer periods were 5, 10, 15, and 20 bars. Multiples of five bars were chosen for the longer periods to minimize weekend and day-of-week effects.

## Segmented Analysis

The idea behind the use of segments was that, within a segment, the underlying characteristics of a stock would be fairly constant and easily estimated. The procedure was as follows: First, a stock was retrieved from the binary database of optionable stocks described earlier. The $n$-period historical volatility was calculated for the stock at each bar; in this instance, $n$ was 100. A simulacrum was then created with a local volatility similar to that of the real stock. Other than a slowly changing volatility, the simulated stock was designed to trace out a geometric random walk that fully satisfied the assumptions, made by standard pricing models, of sequential independence and a log-normal distribution of returns.

Next, a set of segments, each $m$ bars in length, were defined for the stock's data. The last (most recent) segment ended $m/2$ bars prior to the end of the data series. The next-to-last segment ended $m/2$ bars earlier than the last segment. In other words, each segment ended $m/2$ bars before the end of the subsequent segment and an equal number of bars after the end of the preceding one. The first segment began with a bar greater than or equal to $m/2$. Given that each segment was $m$ bars in length, there was a 50% overlap between segments. A segment length, $m$, of 150 bars was used in the analysis. Once all segments were defined, each segment in the set was checked. If the stock was inactive or had a price of less than $2 (uncorrected for splits) at any point within the segment, the segment was discarded. Only complete segments having valid data from start to finish were analyzed further.

At this point, a segment was selected from the set of valid segments for the stock. Logarithmic returns for each of the five holding periods were then calculated from data in the segment. If dollar or percentage-based returns are log-normal in their distribution, then logarithmic returns will be normally distributed. Because of this, volatility is usually calculated from logarithmic returns and their use also makes skew and kurtosis more interpretable. A return for a given holding period was simply the logarithm of the ratio of the price at the end of the holding period to the price at the beginning. Returns were allowed to have overlapping holding periods. For example, to obtain the first 10-bar logarithmic return for a segment, the price at bar 11 was divided by the price at bar 1 and the logarithm was taken. For the second 10-bar return, the price at bar 12 was divided by the price at bar 2 and the logarithm was again determined. Returns were calculated in this manner until the end of the segment was reached. Consequently, a segment of $m$ bars had $m–k$ $k$-bar returns.

The first four moments were then calculated from the logarithmic returns for each of the five holding periods. Calculations were carried out for both returns from the real stock and for returns from the identically processed simulacrum. Each segment, therefore, had eight statistics associated with it for each of the five holding periods: the first four moments for the real stock and the same statistics for the simulacrum.

After all moments were computed, the second moments, in the form of standard deviations or volatilities, were normalized. Normalization was accomplished by dividing the second moment for each of the five holding periods by the second moment for the 1-bar holding period and then multiplying by a bias correction coefficient. This was done separately for both the real and simulated data. The bias correction coefficient was

$\sqrt{m/(m-k+1)}$, where $m$ was the number of returns and $k$ the holding period. Bias correction was necessary to adjust for degrees of freedom lost due to overlap in the holding periods on which the returns were based. Normalization was not required for skew or kurtosis since these statistics are unaffected by scale.

Once a segment was completely analyzed, the resultant statistics were saved to a scratch file and the next segment for the stock was selected for processing. After all segments for a given stock had been processed, the next stock was retrieved and the segmenting and analytic procedures repeated. The process was continued until every valid segment from every stock had been considered. There were 39,333 valid segments taken from 2,241 stocks. The finishing step was to calculate the mean and standard deviation over all segments for each of the eight statistics for each of the five holding periods.

### Results from Segmented Analysis

The means and standard deviations of the segment statistics are presented in Table 4–1. They are broken down by stock type (real, simulated), holding period (1-, 5-, 10-, 15-, 20-bar), and moment (growth, volatility, skew, kurtosis).

***Statistical Independence of Returns*** First consider the question of whether successive returns are statistically independent. If the returns from individual time steps are statistically independent, then the volatility of returns from a stock held for a given period will be proportional to the square root of the number of time steps in that period. Under the assumption that volatility is proportional to the square root of time, and given that all volatilities have been normalized by dividing them by the volatility for the 1-bar holding period (a single time step), the mean volatilities for the 1-, 5-, 10-, 15-, and 20-bar holding periods should theoretically be 1, $\sqrt{5}$ or 2.236, $\sqrt{10}$ or 3.162, $\sqrt{15}$ or 3.873, and $\sqrt{20}$ or 4.472, respectively.

The actual figures for mean volatility presented in Table 4–1 are close to these theoretical values, suggesting that volatility does in fact scale approximately as the square root of holding time. Mean volatility figures for the simulacra lie even closer to the

**TABLE 4–1**

First Four Moments of Normalized Returns Calculated for Five Holding Periods Using the Segmented Method on Both Real and Simulated Stocks

| **Means** | | **1-Bar** | **5-Bar** | **10-Bar** | **15-Bar** | **20-Bar** |
|---|---|---|---|---|---|---|
| Growth | Real | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 |
| | Sim | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Volatility | Real | 1.000 | 2.189 | 3.023 | 3.669 | 4.217 |
| | Sim | 1.000 | 2.225 | 3.128 | 3.809 | 4.378 |
| Skew | Real | –0.028 | –0.035 | –0.052 | –0.062 | –0.051 |
| | Sim | –0.002 | –0.001 | 0.000 | 0.001 | 0.000 |
| Kurtosis | Real | 3.940 | 1.093 | 0.325 | –0.042 | –0.273 |
| | Sim | 0.149 | 0.009 | –0.154 | –0.287 | –0.395 |
| **Standard Deviations** | | **1-Bar** | **5-Bar** | **10-Bar** | **15-Bar** | **20-Bar** |
| Growth | Real | 0.003 | 0.015 | 0.029 | 0.044 | 0.060 |
| | Sim | 0.003 | 0.016 | 0.032 | 0.049 | 0.066 |
| Volatility | Real | 0.000 | 0.251 | 0.483 | 0.698 | 0.920 |
| | Sim | 0.000 | 0.208 | 0.457 | 0.711 | 0.967 |
| Skew | Real | 1.092 | 0.697 | 0.600 | 0.546 | 0.511 |
| | Sim | 0.228 | 0.337 | 0.409 | 0.440 | 0.454 |
| Kurtosis | Real | 7.907 | 1.971 | 1.131 | 0.857 | 0.716 |
| | Sim | 0.589 | 0.626 | 0.634 | 0.612 | 0.590 |

theoretical values, which is not surprising given that these were constructed to have statistically independent returns. If volatility was not proportional to the square root of time, then the assumption that stock returns from successive time steps are independent and that stock prices follow a random walk would have to be rejected. The fact that volatility does scale roughly as the square root of time indicates that returns may be fairly independent of one another. It also implies that it is reasonable to assume an approximate square-root-of-time scaling when pricing options, even if returns do exhibit some complex, nonlinear dependencies that are not reflected in the time-related scaling of volatility.

While the volatility for genuine returns is more or less proportional to the square root of holding time, this relationship is not perfect. For each holding period, the mean volatility for real stocks is slightly lower than that for simulated stocks and the difference grows at least up to a 20-bar holding period. For a 20-bar holding period, real stocks have a mean volatility that is 3.7% lower than that of simulated ones. A possible explanation would be the presence of some mild countertrend activity in the behavior of real stocks over longer holding periods. For most practical purposes, however, a square-root-of-time scaling for volatility may be assumed.

The answer to the question "Are successive returns independent?" is a qualified "yes." This is consonant with the experience of successful traders. Although traders believe that there are small inefficiencies that can be exploited for a profit—otherwise, why trade—most would argue that these inefficiencies are complex and hard to find and that, for the most part, future price movements are unpredictable from (i.e., independent of) the movements that immediately precede them.

***Log-Normality of Returns*** The second question the study intended to address was whether or not returns are log-normal in their distribution. If $x$ is a random variable with a log-normal distribution, then $\ln(x)$ will be a random variable with a normal distribution. Consequently, if dollar or percentage-based returns have a log-normal distribution, as assumed by standard pricing models, then logarithmic returns should be normally distributed. Are they? As mentioned earlier, a normally distributed random variable has neither skew nor kurtosis, i.e., its skew and kurtosis are zero.

As reported in Table 4–1, the mean skew exhibited by returns from genuine stocks clearly differs from zero; it is consistently negative across all holding periods. In contrast, returns from the Monte Carlo simulacra evidence little or no skew. Although the amount of skew seen for real stocks is small, it is statistically meaningful. The conclusion is that returns from real stocks have a distribution that differs somewhat from log-normal: In comparison to what standard pricing models expect, real stocks more frequently take steep dives or make gentle ascents, and less frequently stage powerful rallies—a pattern of behavior readily seen in stock charts.

Positive kurtosis, too, is evidence for deviations from normality that violate the assumptions of popular pricing models. For the 1-bar holding period, the mean kurtosis for real stocks was 3.940, while for simulated ones it was only 0.149, a very substantial difference. This means that the distribution of 1-bar returns from real stocks is leptokurtic. Extreme price changes, either up or down, happen more frequently than allowed for by standard models. Just remember the crash of 1987 or the euphoric bubble of 2000. Such events would happen only once every several thousand years, if market movements truly followed a log-normal random walk. However, such crashes and bubbles occur with alarming regularity. Minimal price changes also occur with greater frequency than one might expect, as anyone who has repeatedly lost money purchasing options on stocks that rest motionless would attest. Conversely, moderate price movements occur somewhat less often than standard models would anticipate.

The results in Table 4–1 indicate that, as the holding period increases, the level of kurtosis comes more in line with what would be anticipated given a log-normal distribution of returns. By the time a 20-bar holding period is reached, the difference between the real and Monte Carlo data is small. On that time frame, the typical stock's price behavior is in closer agreement with the assumptions of most option pricing models. The decline in kurtosis across holding periods could be explained by occasional large shocks to price—sharp movements that span only a few days and that are distinct from the smaller, more normal ones occurring before or after. When 1-day periods are examined, such intense 1- or 2-day thrusts would have a strong impact on the distribution of returns. Their influence would be washed out over longer periods by the larger number of smaller movements taking place at other times. This explanation for the decline in kurtosis with longer holding periods is supported by the experiences of traders. Any trader of small stocks will occasionally see a stock jump 10%, 20%, or even 100% (or decline likewise) in the space of two or three days, and then, after the shock, return to more normal behavior. Such a pattern of price movement frequently occurs in response to rumors and can be very dramatic among smaller, less liquid issues; no doubt it occurs in larger issues as well.

Although small in absolute size, the mean kurtosis figures for the simulacra also differ from zero and become more negative as holding time increases. This is due to the segmenting and normalization procedures employed when computing the moments and is the reason comparisons are made against Monte Carlo baselines rather than against theoretical values.

***Estimating Standard Errors*** The standard deviations in Table 4–1 may be used to estimate standard errors for the corresponding means. A standard error for a sample mean would usually be computed by dividing the standard deviation by the square root of sample size, here the number of segments. The usual method assumes that sample data points are independent; in this instance, however, they are clearly anything but. Because of overlapping segments, and even more so as a result of correlations between stocks, a great deal of statistical dependence exists. Does this mean that a standard error cannot be found? No. A standard error may still be coarsely reckoned by making some subjective adjustments to standard statistical techniques. The trick has to do with degrees of freedom. It is the effective degrees of freedom (or effective sample size) that are reduced by statistical dependence. If one is willing to venture a guess regarding the effective degrees of freedom, given the dependencies on the data, an estimate of the standard error can be obtained. The trick is to divide the sample standard deviation by the square root of the subjectively estimated effective degrees of freedom instead of dividing it by the square root of the actual sample size.

In the present case, there are about 40,000 segments or sample data points. About half of the degrees of freedom implicit in those data points are lost to the 50% overlap between segments. That leaves roughly 20,000 degrees of freedom. When stocks, rather than simulacra, are being considered, these remaining degrees of freedom must be further trimmed. If stocks were perfectly correlated with one another, the 20,000 remaining degrees of freedom would have to be divided by the number of stocks, about 2,000. But, stocks are not perfectly correlated, only partially so. Suppose 90% of the 20,000 degrees of freedom are lost to correlations between stocks, about 2,000 effective degrees of freedom would then remain. Assuming these figures are roughly on target, the standard error can be estimated for any real stock mean in Table 4–1 by dividing the corresponding sample standard deviation by 44.72, the square root of 2,000. For simulated stocks, the standard deviation would be divided by 141.42, the square root of 20,000, since degrees of freedom are lost only to segment overlap and not to correlations between stocks.

As an example, consider the calculation of a standard error for the mean kurtosis of returns from the real stocks over a 5-bar holding period. If there were there no correlations between stocks and no overlap between segments, the standard error of mean kurtosis would be 1.971 divided by the square root of the number of segments or about ±0.010. Since there are correlations and overlaps, 1.971 is divided by 44.72, the square root of the effective degrees of freedom, yielding ±0.044 as the estimated standard error. For the Monte Carlo baseline data, the standard error is 0.626 divided by the square root of 20,000, which works out to be ±0.004. It should be noted that standard errors determined by the method just discussed may be used to compute the statistical significance of differences between means. It is possible, e.g., to roughly reckon the statistical significance of the difference between the kurtosis of simulated and real stocks. An extensive discussion of degrees of freedom, both real and effective, and the use of statistics in difficult circumstances can be found in *The Encycl0opedia of Trading Strategies* (Katz and McCormick, 2000).

### Nonsegmented Analysis

An alternative way to calculate the first four moments across holding periods was tested. In this analysis, the stock data were not segmented. Instead, for every valid bar of every stock, normalized returns were computed. Why were the returns normalized? Because volatility varies dramatically from stock to stock, and even from time to time within the same stock. Without normalization, the returns sampled would differ greatly in magnitude from era to era and stock to stock, and in a way unrelated to the distributions under investigation.

The first step in calculating a normalized return for a given stock was to determine the $m$-period historical volatility at the reference bar. In this instance, $m$ was 100. The historical volatility was multiplied by $\sqrt{k/252}$ to convert it from an annualized measure to an expected volatility for a $k$-bar return. The logarithmic return from the $k$-bar period immediately following the reference bar was then divided by the expected volatility for a $k$-bar return as determined from the historical volatility. These calculations were performed for every valid reference bar, and for both real stocks and simulacra. The result was two sets of normalized returns for each $k$-bar holding period. For each holding period there were a total of 2,473,149 returns from real stocks and an equal number from simulated stocks, certainly sufficient for stable statistics.

Only stocks that were active over the entire period covered in the binary database were analyzed. This limited the total number of stocks to 1,475. The reason for only considering stocks active over the entire historical span was to avoid any bias that might result from the existence of different numbers of stocks in different historical periods contributing to the sample of returns.

Moments were computed across holding periods using the normalized returns for both real stocks and Monte Carlo simulated stocks. In addition to the mean, standard deviation, skew, and kurtosis, the average deviation was calculated. The average deviation is a so-called *robust statistic*. It is robust in the sense that it is less affected by outliers (highly deviant data points) and violations of normality than its nonrobust cousin, the standard deviation. Like the standard deviation, the average deviation measures the spread of a distribution or of a sample of data points.

There were two reasons for performing an alternative, nonsegmented analysis. One reason was the desire to verify results from the segmented method. The other reason was the need to determine whether a simpler, whole-sample method was workable. It was desirable to have in reserve a whole-sample method applicable to situations where insufficient data for the segmented approach would be available, e.g., when studying day-of-month effects.

## Results from Nonsegmented Analysis

The basic pattern of findings from the nonsegmented analysis is similar to that produced by the segmented method, except that the attenuation of skew and kurtosis due to segmentation was avoided. Table 4–2 shows the results from the nonsegmented analysis.

***Volatility and Independence of Returns*** Given how the returns were normalized, the observed volatility should be slightly greater than unity and roughly equal across holding periods if successive price movements are independent and volatility scales as the square root of time. The figures in Table 4–2 show that this is indeed the case. Also implying a market that behaves almost precisely as standard theory would suggest is the fact that for longer holding periods, the volatilities for the

**TABLE 4–2**

Moments of Normalized Returns Calculated Using the Nonsegmented Method for Each of Five Holding Periods and for Both Real and Simulated Stocks

| | | 1-Bar | 5-Bar | 10-Bar | 15-Bar | 20-Bar |
|---|---|---|---|---|---|---|
| Growth | Real | –0.001 | 0.000 | 0.002 | 0.003 | 0.005 |
| | Sim | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 |
| Avg dev | Real | 0.751 | 0.772 | 0.779 | 0.788 | 0.794 |
| | Sim | 0.807 | 0.809 | 0.812 | 0.814 | 0.818 |
| Volatility | Real | 1.073 | 1.072 | 1.067 | 1.068 | 1.070 |
| | Sim | 1.046 | 1.050 | 1.055 | 1.060 | 1.066 |
| Skew | Real | –0.822 | –0.675 | –0.674 | –0.652 | –0.643 |
| | Sim | –0.004 | 0.002 | 0.005 | 0.008 | 0.018 |
| Kurtosis | Real | 22.999 | 9.217 | 7.188 | 5.838 | 5.177 |
| | Sim | 1.034 | 1.073 | 1.114 | 1.181 | 1.435 |

real and simulated stocks are nearly equal; however, the volatility of genuine stocks relative to simulated stocks declines slightly as the holding period increases. A decline in relative volatility with increasing holding period was also observed with the segmented method and might indicate that real stocks exhibit some systematic countertrend activity.

Essentially the same pattern appears when the average deviation is used instead of the standard deviation as a measure of volatility. Theoretically, the average deviation is about 23% smaller than the standard deviation when the underlying distribution is normal. The numbers in Table 4–2 are lower by about 23% for the Monte Carlo simulacra, just right for a sample drawn from a normal distribution, and by around 27% for the real stocks. The greater difference between the standard deviation and average deviation for real stocks is no doubt a consequence of positive kurtosis in real stock returns; a platykurtic distribution (one with positive kurtosis) has longer tails that inflate the standard deviation more than they do the average deviation and thus lead to a greater disparity in the values of these statistics.

***Skew, Kurtosis, and Log-Normality*** As with the segmented method, skew is negative across all holding periods examined. The observed negative skew is greater in the nonsegmented analysis because the attenuation that results when estimates are made from small segments or samples is absent.

Kurtosis is intensely positive across all holding periods. As with skew, the kurtosis figures have not been reduced by segmentation. Consistent with what was found using the segmented analysis, kurtosis declines as the length of the holding period increases, but always remains substantially greater for real than for simulated stocks. The normalization procedure and, to a lesser degree, the slowly changing volatility in the simulated stocks are responsible for the mild kurtosis observed in the Monte Carlo returns.

The mean returns are quite small, as might be expected for short holding periods. The numbers in Table 4–2 are slightly larger in absolute size than in Table 4–1 because the data upon which the means are based are not raw logarithmic returns, but rather normalized logarithmic returns expressed as $z$-scores, i.e., in terms of standard deviations.

### Nonsegmented Analysis of Two Indices

A nonsegmented analysis was carried out on the QQQ and SPY index securities. For those not familiar with these popular, optionable securities, the QQQ tracks the NASD-100 index while the SPY tracks the S&P-500 index. The idea was to see if these index-based securities (and stock indices, generally) behave like individual stocks. Table 4–3 contains the results from the analysis broken down by data type (real or simulated), moment (growth, average deviation, standard deviation, skew, kurtosis), holding period (1-, 5-, 10-, 15-, 20-bar), and index symbol (QQQ or SPY). Because data samples were smaller for indices than for individual stocks, standard errors are larger for the statistics reported in Table 4–3 than for those appearing earlier in Table 4–2.

The mean returns in Table 4–3 are mildly positive for all holding periods. If these numbers appear large, remember that the returns on which they are based were measured in standard deviation units (i.e., as $z$-scores) and not as percentages.

Consistent with theoretical considerations, the volatility of returns from the simulacra is slightly greater than unity for every holding period; except for the 1-bar holding period, the volatility of returns from the real index securities is not: real index returns from longer holding periods have volatility levels considerably below unity, and even further below the Monte Carlo figures. The difference between the real and Monte Carlo volatility is greatest for the 10-bar holding period. Substantial countertrend or mean-reverting activity in both the QQQ and SPY might account for these findings. Unlike returns from individual stocks, successive returns from index securities are not independent and the square-root-of-time scaling assumption may lead to an overestimation of volatility by as much as 10% for certain holding periods.

As with tests involving individual stocks, skew for returns from the index securities is lower than for returns from the simulacra across all holding periods. Negative skew seems to be

**TABLE 4–3**

Moments Calculated Using Nonsegmented Method for Normalized Returns from Two Popular Stock Indices across Five Holding Periods

| SPY | | 1-Bar | 5-Bar | 10-Bar | 15-Bar | 20-Bar |
|---|---|---|---|---|---|---|
| Growth | Real | 0.007 | 0.021 | 0.033 | 0.043 | 0.053 |
| | Sim | –0.006 | –0.013 | –0.019 | –0.025 | –0.031 |
| Avg dev | Real | 0.799 | 0.770 | 0.732 | 0.748 | 0.753 |
| | Sim | 0.818 | 0.827 | 0.838 | 0.828 | 0.804 |
| Volatility | Real | 1.065 | 0.999 | 0.951 | 0.947 | 0.948 |
| | Sim | 1.056 | 1.072 | 1.089 | 1.073 | 1.052 |
| Skew | Real | –0.287 | –0.446 | –0.513 | –0.450 | –0.377 |
| | Sim | –0.005 | 0.376 | 0.370 | 0.293 | 0.253 |
| Kurtosis | Real | 2.518 | 1.478 | 1.302 | 0.725 | 0.354 |
| | Sim | 0.624 | 1.781 | 1.189 | 0.959 | 1.160 |
| **QQQ** | | **1-Bar** | **5-Bar** | **10-Bar** | **15-Bar** | **20-Bar** |
| Growth | Real | 0.024 | 0.059 | 0.086 | 0.106 | 0.126 |
| | Sim | –0.005 | –0.011 | –0.014 | –0.016 | –0.020 |
| Avg dev | Real | 0.814 | 0.778 | 0.761 | 0.774 | 0.784 |
| | Sim | 0.813 | 0.820 | 0.832 | 0.819 | 0.793 |
| Volatility | Real | 1.054 | 0.984 | 0.959 | 0.955 | 0.969 |
| | Sim | 1.046 | 1.056 | 1.071 | 1.049 | 1.027 |
| Skew | Real | –0.085 | –0.277 | –0.105 | –0.040 | 0.067 |
| | Sim | 0.038 | 0.320 | 0.420 | 0.407 | 0.398 |
| Kurtosis | Real | 2.006 | 0.846 | 0.297 | –0.177 | –0.358 |
| | MC | 0.584 | 1.195 | 0.978 | 0.948 | 1.294 |

characteristic of both individual equities and equity indices like the QQQ and SPY. Kurtosis is positive for the shortest holding period, although not so much so as for individual stocks. For longer holding periods, kurtosis becomes effectively negative (below the simulacrum baselines), especially for the QQQ. Like declining relative volatility with increasing holding time, this may indicate the presence of countertrend activity. Negative kurtosis would be anticipated for a security that has prices constrained by resilient barriers of support and resistance, or in cases where prices tend to revert to former levels after severe shocks.

The lower relative volatility and unusual negative kurtosis seen with longer holding periods will affect option prices and cause standard models to produce erroneous results. For instance, when using historical volatility, a model like Black-Scholes will tend to overestimate the value of QQQ and SPY options; it will do so because movements of the underlying securities in the period remaining before expiration are likely to be less than expected, given the historical volatility level. Because negative kurtosis reduces the value of far out-of-the-money options, the overestimation will be worse to the extent that the options being priced are of this kind; standard models tend to overprice far out-of-the-money options when returns have a platykurtic distribution.

## Discussion

When individual stocks are studied, successive returns appear to be nearly independent (at least in terms of any simple, linear relationship) and volatility scales roughly as the square root of time. The assumptions made by popular pricing models regarding the independence of returns and the scaling of volatility seem compatible with observed market behavior. In other words, stock price movements are fairly efficient in the sense of being relatively unpredictable. This is not to say that stock price movements cannot be predicted at all; some predictability that can be captured with appropriate nonlinear prediction models may exist. However, for practical purposes involving individual stocks over the timeframes studied here, the assumption of independent returns appears to be acceptable. Such is not the case for indices. When indices are examined, successive returns are not independent. Index data appear to exhibit dependency, perhaps in the form of mean-reverting or countertrend activity that reduces the relative volatility for certain holding periods.

Whether dependent or independent, returns are clearly not log-normal in their distribution. Returns from both stocks and indices are negatively skewed across all holding periods.

Kurtosis is strongly positive across all holding periods for stocks, with the shortest periods having the most severe kurtosis. For indices, kurtosis is positive for the shortest holding periods, but drops below the Monte Carlo baselines when returns from longer periods are considered.

In plain English, for stocks, and for indices held for short periods, there are more extreme returns, both positive and negative, than expected from a log-normal distribution, as well as returns that reflect little or no change in price. In addition, there is a tendency for the extreme returns to be negative, rather than positive—stocks tend to fall faster than they rise. There are fewer moderate returns in either direction than would be expected. For indices held for longer periods, the pattern is one of a lower-than-expected volatility and an excess of moderate returns in both directions punctuated by the occasional large negative return.

Deviations from independence and log-normality, such as those just discussed, will cause option premiums to differ from what popular pricing models indicate is fair. The difference will be especially pronounced under certain circumstances and for particular options. For example, positive kurtosis gives out-of-the-money options greater value than what might be calculated using a model like Black-Scholes, which assumes a log-normal distribution of returns. The reason is obvious: extreme moves that could result in an out-of-the-money option becoming significantly in-the-money at some future time are more likely than allowed for by the standard models.

Turning this around, if market prices reflect actual value then the implied volatility calculated with a standard model will be greater for out-of-the-money options than for at-the-money options trading on the same security. Higher implied volatility has indeed been observed for many out-of-the-money options. When combined with higher implied volatility for deeply in-the-money options, another effect of positive kurtosis results—the so-called “volatility smile.” The term derives from the smile-like appearance of the curve when implied volatility is plotted against moneyness for a chain of options having the same expiration date.

The reduced volatility in index returns from longer holding periods will cause index options to have less value than anticipated based on standard pricing models. To the extent that serial dependence in returns attenuates volatility, models like Black-Scholes will tend to overprice such options. Negative skew will increase the value of out-of-the-money puts while decreasing the value of out-of-the-money calls. Consequently, standard models will overprice the calls and underprice the puts, and implied volatility for out-of-the-money puts will be greater than for out-of-the-money calls. There is no violation of put-call parity, since options of different terms are involved. In fact, as a consequence of put-call parity, standard models will also overprice in-the-money puts and underprice out-of-the-money calls.

## STUDY 2: MOMENTS AND DAY OF WEEK

This study examines how the moments of returns vary with day of the week. Is Monday a day of sharp corrections? Does Friday often close strong? Which days are least volatile and which are most? Do weekend returns have a greater standard deviation than nonweekend returns? These are some of the issues addressed.

The question of whether weekend returns (those from the close on Friday to the close on Monday) are more volatile than weekday returns is an important one when it comes to pricing options. Consider the fact that the amount of time that remains in the life of an option is an essential variable required by all pricing models. Time may be measured in calendar days, or in bars, or trading days. Pricing models are customarily set up to measure time in calendar days. But is this the most effective way to measure time for the purpose of pricing options? Or should time be measured in trading days? It is all a matter of volatility. If stock price movements over the weekend are statistically like those occurring on other days, then trading days should be used as the measure of time. If prices move more over the weekend—some might argue they do since more price-influencing events can occur in three days than in one—then calendar time is the appropriate measure. Or perhaps weekends should be counted as if they represent some intermediate interval between one and three days. In fact, if each day manifests a different volatility level, it might be argued that every day of the week should be assigned an "effective time" that is proportional to the square of that day's relative volatility. Other issues examined in this study also have relevance to option pricing and are of practical interest to traders of both options and stocks.

## Method

A procedure very much like the nonsegmented analysis from Study 1 was employed. The procedure was applied to both the entire sample of stocks and to the QQQ and SPY index securities. First the historical volatility was computed for every bar of every stock or index. The logarithmic return from each bar was then normalized by dividing it by the historical volatility. For this purpose, historical volatility expressed in daily (rather than annualized) form was used. Only stocks or indices that were active over the entire historical sample were analyzed. So far, the procedure is identical to that employed in Study 1 to compute the 1-bar returns for the nonsegmented method. At this point, however, the returns were sorted into five distinct groups that correspond to the five days of the week. The moments were then separately determined, along with the average deviation, for each group of normalized returns.

## Results

The results appear in Table 4–4 arranged by moment (mean or growth, average deviation, standard deviation or volatility, skew, kurtosis), day-of-week, and market (stocks, indices). Moments for Monday are based on returns from Friday's close to Monday's close, moments for Tuesday are based on returns from Monday's close to Tuesday's, and so on. The number of data points or returns used to calculate each moment is also shown in Table 4–4.

***Growth*** Because growth was computed from logarithmic returns, it does not reflect gains achieved merely as a result of the action of volatility (see Chapter 3). Also keep in mind that the growth figures are based on normalized returns and hence take the form of $z$-scores (ratios to the standard deviations) rather than percentages.

Growth for the average stock is negative on Monday; it is positive on Friday and, to a lesser degree, on Wednesday.

**TABLE 4–4**

Moments of Returns Broken down by Day-of-Week, Real-vs-Simulated and Stocks-vs-Indices

| Stocks | | Mon | Tue | Wed | Thu | Fri |
|---|---|---|---|---|---|---|
| Growth | Real | –0.037 | –0.002 | 0.012 | –0.001 | 0.027 |
| | Sim | 0.000 | –0.003 | 0.000 | 0.002 | 0.001 |
| Avg dev | Real | 0.770 | 0.748 | 0.751 | 0.760 | 0.728 |
| | Sim | 0.809 | 0.811 | 0.809 | 0.807 | 0.809 |
| Volatility | Real | 1.102 | 1.060 | 1.069 | 1.089 | 1.041 |
| | Sim | 1.033 | 1.034 | 1.033 | 1.031 | 1.033 |
| Skew | Real | –1.084 | –0.772 | –0.469 | –0.951 | –0.859 |
| | Sim | 0.007 | –0.010 | –0.003 | –0.001 | –0.014 |
| Kurtosis | Real | 25.632 | 20.577 | 19.297 | 25.501 | 24.610 |
| | Sim | 0.504 | 0.494 | 0.581 | 0.539 | 0.485 |
| Data pts | | 4,75,644 | 5,11,752 | 5,10,410 | 5,01,613 | 5,00,128 |
| **Indices** | | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** |
| Growth | Real | –0.008 | 0.003 | 0.051 | 0.001 | 0.032 |
| | Sim | 0.081 | –0.017 | 0.016 | –0.017 | –0.015 |
| Avg dev | Real | 0.808 | 0.839 | 0.777 | 0.818 | 0.787 |
| | Sim | 0.806 | 0.798 | 0.836 | 0.826 | 0.820 |
| Volatility | Real | 1.114 | 1.087 | 0.999 | 1.048 | 1.049 |
| | Sim | 1.036 | 1.026 | 1.068 | 1.041 | 1.023 |
| Skew | Real | –1.228 | 0.136 | 0.288 | –0.171 | 0.207 |
| | Sim | 0.004 | 0.017 | –0.016 | –0.035 | –0.136 |
| Kurtosis | Real | 5.528 | 1.005 | 1.110 | 0.657 | 1.950 |
| | Sim | 0.348 | 0.511 | 0.244 | 0.157 | –0.092 |
| Data pts | | 658 | 708 | 706 | 694 | 692 |

The same pattern is observed for the combined SPY and QQQ index data except that, on a relative basis, Monday's losses are smaller than those for stocks and Wednesday's gains are greater. For both stocks and indices, growth is negligible on Tuesday and Thursday.

***Volatility and Average Deviation*** For both stocks and indices, volatility is highest on Monday. This might be expected: Monday is traditionally a volatile day. However, the volatility for over-the-weekend returns on Monday is only marginally greater than for weekday returns on Thursday, in the case of stocks, or Tuesday, in the case of indices, and by far lower than would be expected if it scaled with calendar rather than bar or trading time. For any 3-day period (calendar time), the expected normalized volatility is around 1.033 multiplied by 1.732 (the square root of 3), or 1.789; observed Monday volatility is obviously far lower than this: 1.102 for stocks and 1.114 for indices. The number 1.033 comes from the large-sample simulated volatility for normalized returns from a 1-bar period; it is greater than one because the divisor in the normalization, the historical volatility, contributes a little to the total variance of the normalized returns.

Volatility is lowest on Friday for stocks, and on Wednesday for indices. For the indices, Wednesday's volatility is not merely the lowest of the week; it is extremely low, even when considered on its own.

Overall, the average deviation behaves much like the standard deviation, the usual measure of volatility. The average deviation is, however, consistently low relative to the Monte Carlo baselines, whereas the standard deviation is consistently high. This effect is due to the fact that the distribution of logarithmic returns has longer tails than the normal distribution, and thus exhibits kurtosis. Kurtosis inflates the standard deviation and, when returns are normalized as done here, drives down the average deviation, which is much less affected by the longer tails in the distribution.

***Skew*** In the world of individual stocks, skew is decidedly negative for every day of the week. Monday has the most negative skew, whereas Wednesday has the least. For the indices, skew is quite negative for Monday, but mildly positive for Wednesday and Friday.

***Kurtosis*** Strong positive kurtosis exists for stocks across all days of the week. For indices, kurtosis is positive and strong on Monday, moderate on Friday, and weak but still positive on Thursday.

### Discussion

Is Monday the most bearish day of the week? Does the homily "Don't sell stocks on Monday" have any truth? Growth is most negative on Monday for both stocks and indices. Given such negative growth, the answer would have to be "yes" to both questions. Is Monday a day of sharp declines? Monday has the most negative skew and is also the day with the greatest kurtosis. The negative skew together with positive kurtosis suggests that the answer to this question is also "yes." However, the high level of kurtosis indicates that Monday may have not only an excess of severe losses, but an excess of extreme gains as well. Friday and Wednesday, by way of contrast, are more bullish days; this is true for both stocks, where Friday is the stronger of the two days, and for indices, where Wednesday is the stronger day. It is curious that volatility is at its lowest on these two days, the days of strongest growth.

When it comes to volatility, returns over weekends (on Monday) behave much more like returns from other days of the week than like those that might be expected from any three-day period. Although volatility is slightly elevated on Monday, it is nowhere near high enough to justify the use of calendar time when computing theoretical option premiums. Based on the data sample used in the analysis, weekends should be treated like any other days of the week; bar time, not calendar time, should be used when modeling volatility or pricing options.

From the findings of this study, when pricing options, the customary use of calendar time can be expected to result in overpricing on Friday and underpricing on Monday. Such mispricing may be sidestepped by measuring time in bars. To use bar time with any pricing model, simply let $t$ = bars/252, where $t$ is the time remaining in years, rather than $t$ = days/365.25. It should be mentioned that the choice of time measurement is really significant only for very short holding periods.

The higher kurtosis over the weekend, especially for indices, probably reflects the factors normally assumed to justify the use of calendar time when working with volatility. Political events and some types of news that affect the markets may have a greater likelihood of occurring over the weekend because of the three-day span, but such events are probably rare and tend to affect all stocks. Weekday news more often affects specific stocks and its impact is thus smoothed out in the indices. With indices, the countertrend activity generated by arbitrageurs and index traders tends to dominate. Another possible explanation for the high level of kurtosis in Monday returns is that Monday is a day that traders and investors act on decisions made leisurely over the weekend. If these decisions all happen to go the same way—a fairly uncommon event—a large Monday spike, either up or down, may occur and contribute to a higher level of kurtosis.

Tuesday is a typical day for stocks, but a volatile one for indices, especially the QQQ. Why this should be is a puzzle. In fact, for the indices, the average deviation is greatest on Tuesday, not Monday.

## STUDY 3: MOMENTS AND SEASONALITY

One day early in 2003, one of the authors observed that options on Intel (INTC) were trading with implied volatilities of well over 70%. This suggested that a large move was imminent. The very next day, the same options were trading with implied volatilities in the 30% to 40% range and at dramatically lower prices. What happened? An earnings report had been released. The release of the report caused the uncertainty amongst traders to evaporate and led to a collapse of implied volatility. As it happened, the stock moved hardly at all. The implied volatility did not reflect reality; it was excessively high almost exclusively because of the psychological factor of uncertainty. Often, however, large moves do correlate with the dissemination of earnings reports and traders have incorporated this fact into their understanding of the market.

Earnings reports are seasonal; they are released in droves every quarter of every year. Because of index arbitrage, even stocks that are not themselves releasing reports are likely to participate in the movements induced by such releases. And, there may be other seasonal cycles in volatility beyond those induced by earnings reports. For example, October is often considered a highly volatile month, as is the time between late December and early January. In October, volatility seems to have a large downside component—it is the "Month of the

Crash." Conversely, the volatility that characterizes the so-called "January Effect" often involves substantial gains.

Since volatility varies so perceptibly over the year, and since it has a great impact on option value, it seemed appropriate to analyze it for seasonality. The same can also be said for the other moments—growth, skew, and kurtosis. Although seasonality in price (or growth) has been discussed by numerous authors, especially in articles and books on futures trading, seasonality in volatility, skew, and kurtosis has not; yet it is the latter that is most consequential when trading options. For that reason, the following discussion is rather unique and may be of significant practical importance for the options trader.

### Method

First, a stock that was active over the entire historical period was selected. A reference bar was then chosen from that stock's data series. The bar was tested for validity. To be a valid reference, it was necessary that the bar be preceded by enough bars to calculate the historical volatility and that it be followed by a sufficient number of bars to calculate a return. An additional condition for validity was that the bar's date fell between January 1, 1996 and January 1, 2003. If the reference bar was valid, a normalized, logarithmic, 1-bar return was calculated by dividing the closing price of the bar following the reference bar by the closing price of the reference bar, taking the logarithm, and then dividing the result by the daily (unannualized) historical volatility. The historical volatility used to normalize the return was computed over a one-year (252 bar) period to filter out any seasonal variation. Once the return was determined, it was saved along with the date. The next bar from the stock's data series was then chosen as the reference bar. After all valid bars for the currently selected stock had been processed, the next stock was selected and the whole procedure repeated. This sequence of events continued until a normalized return had been computed for every valid bar of every stock.

For each possible day and month, a set of returns was assembled from those saved in the above-described procedure. If a given day and month did not exist (or fell on a weekend) in a particular year, the most recent return that did exist for each stock was used. Hence, for each day and month, returns were taken from every stock and every year in the sample. Because of this, the number of returns assembled for each possible day and month of a year was always equal to 10,325—the number of stocks (1,475) multiplied by the number of years (7).

For each set of returns thus assembled, the first four moments and the average deviation were computed. The result was a table with columns that corresponded to the different statistics and rows that corresponded to the possible days and months of a year. A centered, circular moving average was then computed for each of the columns in the table to smooth out noise. What is a *centered, circular moving average*? It is a moving average that treats data from the end of a series as if it preceded the beginning and data at the beginning as if it followed the end, so that an equal number of data points on each side of the moving reference can always be included in the calculation. A 20-point circular moving average was used for growth, volatility, and average deviation, while a 30-point average served to smooth the noisier skew and kurtosis data. The smoothed columns or data series were then plotted against the day and month to generate Figures 4–1 through 4–4.

## Results

***Growth*** As can be seen in Figure 4–1, the greatest gains in stock prices occur from mid-October to mid-May, with growth peaking in late April, early May, late October, and early December. Losses tend to occur from mid-May to mid-October, with the most negative growth observed in July and especially in September. The pattern of growth in the current data is quite reminiscent of that seen in the *Calendar Effects Chart* (Katz and McCormick, 1990), which was based on S&P-500 index data from 1976 through 1990, a much earlier period. In that data, too, September is a month of generally negative growth and November through the end of June is a period of strong positive growth. The biggest difference between the two sets of results is that the spikes of negative growth historically observed in October occur somewhat earlier, in September, in the more

**FIGURE 4–1**

Growth as Function of Time of Year

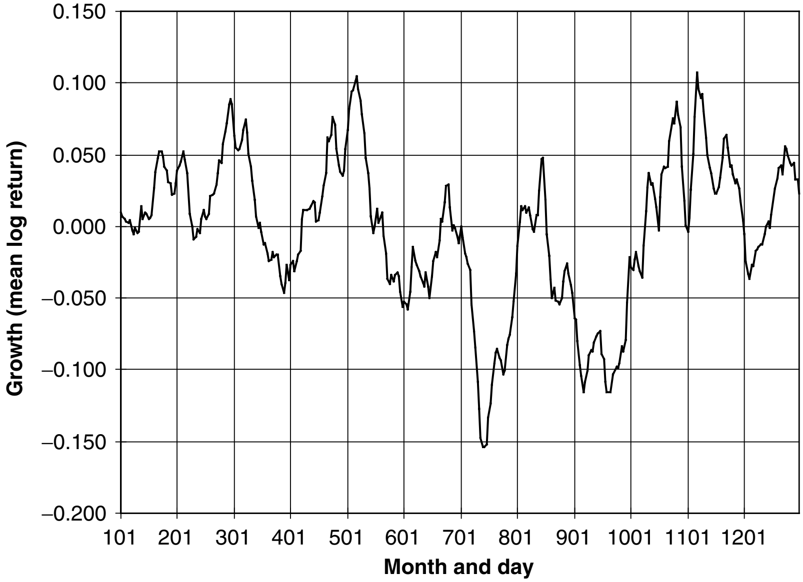
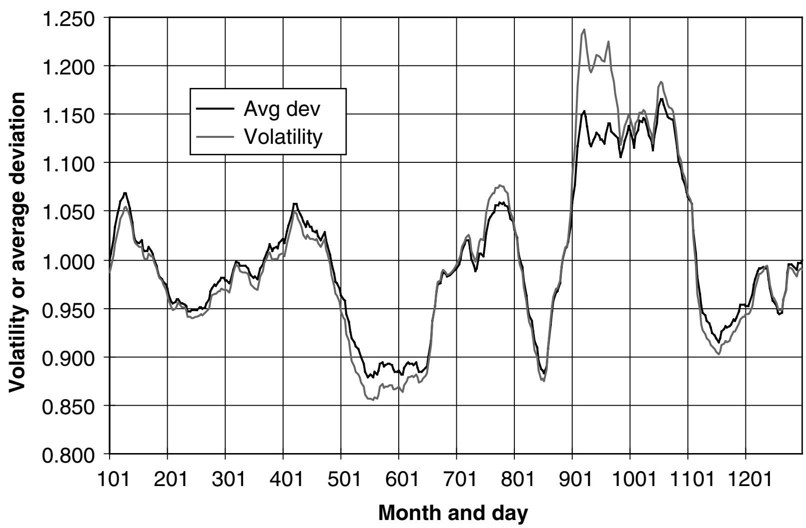



**FIGURE 4–2**

Volatility and Average Deviation Plotted against Time of Year



recent data. This may be due to the fact that several years in the current data sample come from a powerful bear market.

When perusing Figure 4–1, remember that growth was computed from normalized, logarithmic returns. The use of logarithmic returns eliminates, from the growth figures, the small gains that derive from the action of volatility, while the normalization of returns may cause growth to appear larger in absolute magnitude than it is in reality. Normalization also results in a more even contribution by all stocks to estimates of growth; without normalization, more volatile stocks would have a greater impact on estimated growth than do less volatile ones.

***Volatility and Average Deviation*** Figure 4–2 shows that volatility (thin dashed line) peaks in early January, early April, and July, and reaches its highest levels in September and October. Volatility in October is extremely high—about 20% higher than the average volatility over the rest of the year! The high level of volatility in October is significant since the data did not include years having major crashes. A year like 1987 would have driven estimated October volatility levels even higher. Except for September, the peaks in volatility appear in months when earnings reports are typically disseminated. Relatively low levels of volatility are found in February, May, early June, mid-August, and most of November.

The average deviation (thick solid line) tracks volatility very closely. Only in September is there an appreciable divergence between the two measures of distributional spread; the volatility or standard deviation reaches a higher level in September, while the average deviation in September is more in line with its level in October. Reaction to the September 11, 2001 attacks can be ruled out as the cause—unless blame is placed on terrorists attempting to profit from the impending attack—since peak volatility was reached several days earlier.

The observed swings in volatility are not merely of academic interest, but are of sufficient magnitude to be exploited by traders for a profit. From mid-August to early September, volatility rises by more than 20%. It declines an equal amount between late October and mid-November. Several smaller swings of around 10% are also apparent in Figure 4–2.

Note that the volatility plotted in Figure 4–2 was adjusted to have a unit mean. The adjustment was performed prior to generating the chart by dividing every individual data point by the average of all data points in the series, i.e., by the average of the entire year. The same normalization to unit mean was carried out for the average range, also shown in Figure 4–2.

An alternative analysis of volatility was performed to confirm the results of this study. For each bar of each stock, a 15-period historical volatility was computed and then normalized. The historical volatility figures were then averaged by day and month to obtain a chart similar to Figure 4–2. This method of analysis is sensitive only to internal variation (volatility within a given stock at a given time) and not to variation that might occur from year to year or stock to stock. It was suspected that the year-to-year, and possibly the stock-to-stock, variation might be distorting the findings, but this does not appear to be the case. Although not presented here, the results from the alternative analysis were very similar to those obtained from the method described earlier; the only notable difference was that

**FIGURE 4–3**

Seasonality of Skew

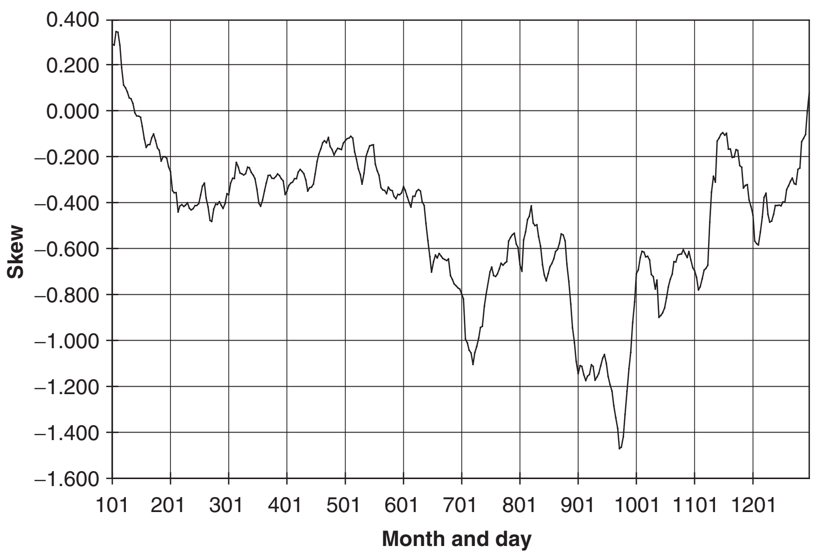
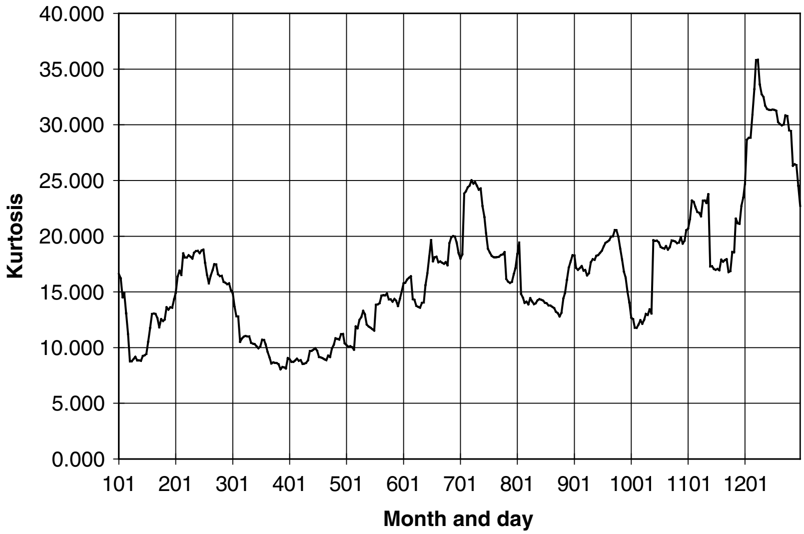


**FIGURE 4–4**

Seasonality of Kurtosis




volatility was somewhat lower in September. The alternative analytic method was not used in this study because, unlike volatility, skew and kurtosis could not be meaningfully estimated from small samples of data taken from individual stocks.

***Skew and Kurtosis*** Figure 4–3 demonstrates that skew is most negative from mid-June to early December, hitting its maximum negative value late in November. Positive skew, something rarely seen with genuine (as opposed to simulated) stock data, is observed in early January.

Figure 4–4 reveals that kurtosis is highest in December. Relatively high values are also seen from late June to November and in February. Lower levels of kurtosis are found in January up to late June.

## Discussion

Is October a month of crashes and wild price swings? Perhaps. Volatility is, in fact, near its highest level in October, skew is strongly negative, and kurtosis is decidedly positive. However, growth happens to be positive, rather than negative, throughout much of the month. Growth and skew are most negative in September, while kurtosis remains strongly positive. Based on a data sample that contains no outright crashes, it appears that while October has good qualifications as a crash month, September has somewhat better ones. The conclusion is that both months have the potential for intense volatility and sharp declines.

What about the January effect? Consistent with the effect, growth is indeed positive in January. Skew is also positive in very late December and in early January—the only time, in fact, that positive skew is observed. Kurtosis, although high in December, has come down to more average levels in January. There are more extreme returns with a below-average bias toward the downside in December and an average number of extreme returns with a bias toward large gains in January. In other words, the January effect is detectable in the data sample used in the analysis.

Is there a quarterly earnings cycle in volatility? Unquestionably! Measured with the average deviation, volatility peaks in January, April, July, and October, precisely the months when large numbers of earnings reports are released. Just as with the average deviation, when measured using the standard deviation, volatility peaks in January, April, and July. Volatility based on the standard deviation also peaks in September, but remains at a high level well into October. Regardless of the measure used, volatility forms bottoms in February, May, August, and November.

There is a curious relationship between growth and volatility that can be seen in Figures 4–1 and 4–2. Volatility tends to base near market highs, when growth is still positive, and to peak just after periods when growth is most negative. This pattern of volatility has been described by several market scholars. According to trading lore, low volatility is often observed proximal to market tops and high volatility near market bottoms. The effect has even been noted with implied volatility: OEX options (including puts) were extremely inexpensive (had very low implied volatilities) as the market topped prior to the 1987 crash; these same puts were very expensive near the bottom that followed.

Seasonal variations, as seen in Figures 4–1 through 4–4, are of real significance to stock and options traders. Consider volatility. The changes in volatility with time of the year are not just theoretical, but are large enough to cause option prices to vary by 10%, 20%, or more. The trader can definitely gain an edge by focusing on positions that are long Vega at times when volatility is seasonally low but about to rise, and positions that are short Vega when volatility is high and about to fall. A glance at Figure 4–1 suggests that, over a period of several years, a trader could profit merely by selling option premium near the end of October and covering in mid-November, or buying in early June and selling late in the month.

Some of the seasonal patterns just discussed have been noted by traders and students of the markets for a long time (e.g., the October jitters) or have solid, logical reasons for their existence (e.g., the quarterly swings in volatility produced by earnings reports). These seasonal patterns are likely to continue to exist in the future and must be seriously considered when trading options or trying to develop a pricing model.

## STUDY 4: MOMENTS AND EXPIRATION

Trading lore says that a lot of volatility occurs around the expiration dates of options and futures. So-called "triple witching days," which occur when both options and index futures simultaneously expire, are notorious for being highly volatile times in the marketplace. Yet the authors' experience in trading options on the QQQ index tells a somewhat different story: index prices often seem to converge to one of the popular strikes as expiration approaches, with declining (rather than rising) volatility—almost as if the market wants the sellers of straddles to achieve maximum profits. Then, again, the crash of 1987 took place on the third Friday in October, the last trading day for the OEX options, and high volatility was recorded for the whole week leading up to the event.

For a variety of reasons, the market might be expected to behave differently at different times vis-à-vis expiration. This study examines the behavior of real stocks for each of 15 trading days prior to the day that options cease to trade. It provides some objective information on how volatility levels and the other moments vary with the approach of expiration.

## Method

The method of analysis is similar to that used to examine moments of returns vis-à-vis days of the week in Study 2. In Study 2, returns for both real and simulated stocks were grouped by the day of the week; in the current study, they are grouped by calendar days remaining until expiration Friday, the last day of trading for expiring options.

In Study 2, the Monte Carlo results were reported for every group or day of the week. In the current study they are not reported for every group or day preceding expiration; instead, the Monte Carlo results are computed for every day, but only their average value and standard deviation are reported. On the basis of the construction of the simulated stocks, moments of their returns should bear little or no relationship to the time remaining before expiration of the nearest option series. Thus it makes sense to summarize the results statistically. The average and standard deviation for each of the moments provide trustworthy baselines against which the data from real stocks may be compared.

## Results

Table 4–5 contains the results. There are seven columns in the table. The first column lists the number of days remaining before expiration. The growth or mean return for each day appears in the second column. The third column contains the average deviation; a robust measure of volatility less affected by skew or kurtosis. For each day, the fourth column contains the standard deviation or volatility of the normalized 1-bar returns. Skew and kurtosis appear in the fifth and sixth columns, respectively. The seventh column lists the number of cases or returns used to compute the statistics in the preceding five columns.

Except for the last two rows, the rows in Table 4–5 correspond to days located at different amounts of time prior to option expiration. The first row, beginning with zero, contains

**TABLE 4–5**

Moments of Stock Returns as a Function of Days Remaining to Option Expiration

| Days | Mean | ADEV | SDEV | Skew | KURT | NCAS |
|---|---|---|---|---|---|---|
| 0 | –0.047 | 0.722 | 1.046 | –1.354 | 38.065 | 117,084 |
| 1 | 0.014 | 0.763 | 1.124 | –1.875 | 54.726 | 118,554 |
| 2 | –0.002 | 0.748 | 1.076 | –0.671 | 20.864 | 118,552 |
| 3 | 0.069 | 0.743 | 1.048 | –0.897 | 23.608 | 118,551 |
| 4 | –0.009 | 0.763 | 1.125 | –2.168 | 48.996 | 109,912 |
| 7 | 0.011 | 0.732 | 1.051 | –1.060 | 32.432 | 114,215 |
| 8 | –0.042 | 0.760 | 1.078 | –0.816 | 20.096 | 118,510 |
| 9 | –0.063 | 0.743 | 1.046 | –0.380 | 10.909 | 118,515 |
| 10 | –0.008 | 0.728 | 1.024 | –0.442 | 10.956 | 118,517 |
| 11 | –0.012 | 0.742 | 1.033 | –0.454 | 10.554 | 114,143 |
| 14 | 0.115 | 0.753 | 1.050 | –0.575 | 14.669 | 114,155 |
| 15 | 0.027 | 0.760 | 1.082 | –0.662 | 13.686 | 115,616 |
| 16 | 0.083 | 0.763 | 1.080 | –0.442 | 25.251 | 115,706 |
| 17 | –0.012 | 0.790 | 1.117 | –0.919 | 24.242 | 117,092 |
| 18 | –0.028 | 0.812 | 1.156 | –0.708 | 16.050 | 108,421 |
| MCAVG | 0.001 | 0.809 | 1.032 | –0.004 | 0.516 | 115,836 |
| MCSDEV | 0.003 | 0.003 | 0.003 | 0.011 | 0.086 | 3,111 |

statistics calculated from the normalized logarithmic returns occurring from the close on Thursday to the close on expiration Friday. The row that begins with four contains statistics for returns accruing from the close on the Friday preceding expiration to the close on Monday of expiration week. There are no rows beginning with 5, 6, 12, or 13, because these days fall on weekends. The last two rows in Table 4–5 contain statistics derived from the Monte Carlo simulations. The row labeled MCAVG contains the averages, over all 15 days for which there were data, of the figures for the simulated stocks. The standard deviations for the same figures appear in the row labeled MCSDEV. These Monte Carlo summary statistics can be used to guide interpretation of the moment statistics for the real stocks that appear in earlier rows of the table.

***Growth*** Negative returns appear on expiration Friday despite the fact that Friday is usually a day of positive returns. Tuesday, typically a neutral day, has positive returns during expiration week. Returns are negative on Wednesday and Thursday in the week prior to expiration week. However, the Friday and Wednesday that fall two weeks before expiration week evidence the strong positive growth normal for those days, and Monday shows negative growth, also normal.

***Volatility and Average Deviation*** Volatility on expiration Friday does not differ much from that on any other Friday, but it is mildly elevated on the previous day, Thursday. Volatility on Monday is also elevated, but well within the range of a normal Monday level. Overall, volatility is lower in the week preceding expiration than in the week before that or in the week of expiration. The average deviation follows a similar pattern. Amazingly, when measured by the average deviation, the lowest volatility of any day occurs on the Friday that the options cease to trade!

***Skew and Kurtosis*** Undoubtedly, skew is overall most negative during expiration week, with Monday, Thursday, and Friday being the most negative days, in that order. In the week prior to expiration, Friday also has highly negative skew. Along with heightened negative skew, expiration week is marked by truly exceptional levels of positive kurtosis. The highest levels of kurtosis appear on Thursday and Monday, with Friday falling just behind. The Friday before expiration also shows elevated kurtosis. Kurtosis is dramatically lower from Thursday in the second week preceding expiration through Wednesday in the week immediately before expiration.

## Discussion

The extremely high level of kurtosis just prior to expiration is the prominent finding of this study. High levels of kurtosis imply that, relative to the normal distribution, there is an excess of extreme returns, and an excess of small or negligible returns, both at the expense of moderate returns. This is consistent with both the authors' observations and with general trading lore. The many instances in which returns are small and the market moves little accounts for the author's observation of a quiet pre-expiration period when trading QQQ options. The infrequent, but extremely large, moves that tend to occur during expiration week explain the general feeling amongst traders that expiration is a time of unexpected market turbulence. The trader who sells premium near expiration will very frequently profit from a quiet market, but once in a while will experience a stunning loss.

The greater negative skew during the week of expiration, taken in combination with the high level of kurtosis, indicates that the infrequent but extreme returns tend to contribute more to declining than to rising prices. Although usually quiet, expiration week can be a week of wicked downturns.

## SUMMARY

In this chapter, moments of returns from different holding periods, days of the week, times of the year, and days remaining before expiration were studied. The findings were that moments of returns were conditionally dependent on such factors.

Successive returns were found to be statistically independent for stocks. This was evidenced by the fact that volatility was roughly proportional to the square root of the holding period. Successive returns were not independent for indices. The variation in returns from certain holding periods was well below what would be expected based on the proportional relationship of volatility to the square root of time.

Although returns were statistically independent, at least for stocks, in no case was their distribution log-normal. For the most part, negative skew and positive kurtosis were noted. Kurtosis was greatest for short holding periods, while skew was more evenly apportioned. Relative to a log-normal process, there was an excess of extreme returns, both negative (losses) and positive (gains), with a bias to the downside. The observation of a negatively skewed, leptokurtic distribution in returns from stocks and indices is nothing new. Over the years, many investigators have characterized the distribution of returns from stocks as having long tails and negative skew.

Moments of returns, including volatility, were found to differ across days of the week; however, returns over the weekend were observed to have only slightly greater volatility than returns from other days. Volatility was not sufficiently greater over weekends to warrant the use of calendar time when pricing options. Bar time, not calendar time, is definitely better for this purpose.

During expiration week, kurtosis was extremely positive and skew was very negative. Levels of skew and kurtosis were by far lower in the week preceding expiration week, although still typical of a negatively skewed, leptokurtic process.

A potent quarterly cycle was observed in volatility. Often 10% to 20% in size, the cyclic swings in volatility are large enough to be important to the trader. The quarterly release of financial reports by active companies is probably responsible for this cycle.

The phenomena observed when studying the moments of returns are sufficiently intense to have practical significance when trading or pricing options. Taking conditional variation in growth, volatility, skew, and kurtosis into account can certainly improve a trader's bottom line. But how do the moments, and differences in their levels under a variety of conditions, affect the theoretical fair value and real-market price of an option? Option value, market price, and influential variables are explored in later chapters, beginning with the next in which one of the moments, volatility, is studied in great depth.

## SUGGESTED READING

A brief discussion regarding the statisitical moments of a sample can be found in *Numerical Recipes in C*, Second Edition (Press et al., 2002). Moments and characteristic functions are discussed from the mathematician's perspective in *Probability* (Lamperti, 1966) and in *Continuous Univariate Distributions* (Johnson et al., 1994).

CHAPTER 5

# Estimating Future Volatility

There is no question that volatility is one of the most critical variables when it comes to determining the worth of an option. However, it is not implied or historical (past) volatility that determines value; it is future volatility or movement that does so. It goes without saying that future volatility cannot be directly measured until the future has become the past. Future volatility must, therefore, be estimated or predicted. Often historical volatility is used as a proxy for future volatility when calculating theoretical option prices. Implied volatility, too, frequently serves as an estimate of volatility in times to come. The best estimator of future volatility, however, is definitely not standard historical or implied volatility, taken as is. Better predictors of future volatility are easily constructed as will be demonstrated in this chapter.

## MEASUREMENT RELIABILITY

One issue to consider in the quest for a better predictor of future volatility is the issue of reliability. The ability to accurately estimate the volatility that lies ahead is limited by the reliability of the measurements used to generate the estimate. How reliable a measure is historical volatility calculated as a standard deviation of returns? How about historical volatility assessed with



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

a more robust estimator, such as the average daily range? In fact, what exactly is reliability? The answer to the last question—and the technology able to answer the other two—can be found in the arcane realm of classical test theory or psychometrics.

Psychometrics is a field where the concept of reliability is well developed. Consider volatility from the perspective of a psychometrician. Volatility was defined in previous chapters as price movement taking place in the context of time; it was measured as the standard deviation of logarithmic returns. However, from the psychometric perspective, volatility is a trait, i.e., a parameter governing an underlying generating process (or distribution) that characterizes an individual stock at a particular point in time. Regarded as a trait, the behavior that volatility affects may be observed, but not volatility itself. What behavior does volatility influence? Price movement. In fact, when volatility is regarded as a trait, each day's price movement may be considered a response to an item in a test. Calculating a 30-day volatility becomes analogous to obtaining a score on a psychological test containing 30 items.

The score on any test may be interpreted as a weighted sum of two components: one, the true level of the underlying trait the test is intended to measure; and two, a random error term. Just as there exists a *true score* for intelligence or depression in an individual at a particular moment, a given stock at a given time has a true level of volatility. And, just as a test will reflect both the psychological trait and the random measurement error, so will a return reflect both volatility and noise. It is the proportion of total variance that can be attributed to the underlying trait—whether depression, intelligence, or volatility—that constitutes reliability.

So how reliable is the standard measure of historical volatility? With the machinery of test theory, a question like this is fairly easy to answer. No problem exists with construct validity since the items, the squared daily returns, clearly reflect the desired underlying trait. However, as measures of the underlying trait, the individual items are fairly unreliable. This is not a serious problem because a set of items provides a more reliable measure than any single item. The 30 returns, which may be used as the basis for a 30-bar historical volatility calculation, yield a more reliable measure of underlying volatility than merely one return. This is because a test's trait-related variance increases directly with the number of items, while its error variance increases only with the square root of the item count.

There are various ways to assess the reliability of a scale or test composed of multiple items. A common approach involves the concept of *internal consistency*. One measure of internal consistency is the *split-half correlation*, from which the reliability of a test may be easily reckoned. To calculate a split-half correlation for a measure of volatility, split the data sample into odd and even bars, determine the volatility from the odd bars, determine the volatility from the even bars, and then correlate the odd and even volatility scores over all samples. The odd and even scores can be correlated using the standard formula for the correlation coefficient:

$$r_{xy} = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \tag{5.1}$$

In this formula, $r_{xy}$ is the *Pearson Product-Moment Correlation*, $x$ and $y$ are the variables for which the correlation is to be found, and $n$ is the number of data points. The *split-half reliability* can readily be determined from the split-half correlation. This correlation is, in fact, the reliability of a test containing half as many items as the full-length measure. Reliability may be extrapolated from a shorter to a longer test using the formula

$$r_k = \frac{kr}{(k-1)r+1} \tag{5.2}$$

where $r_k$ is the extrapolated reliability of the longer test, $r$ is the reliability of the shorter test (or test item), and $k$ is the ratio of the size of the longer test to the size of the shorter one. A split-half reliability for a full test may be computed from the correlation between the two halves by setting $k = 2$ in the above formula. There is no reason not to compute split-half correlations and reliability coefficients for a variety of volatility scales, including the standard and average range measures.

Why the fuss over reliability? Because the reliability of an independent (predictor) variable is directly related to the maximum possible correlation of that variable with the dependent (target) variable. If one variable is used as a predictor, the correlation of that variable with the variable being predicted can never be greater than the square root of the predictor variable's reliability. In other words, reliability limits predictive utility. The more reliable a measure, the better an estimate it can be of some other measure or of an underlying trait or hidden parameter.

Obviously, measures used as independent variables should be as reliable as possible. If a measure of historical volatility lacks reliability, then, even if it measures precisely what is intended, it will still be a poor estimator of future volatility. Sometimes a *biased estimator* can provide better results. In fact, bias in an estimator may be perfectly acceptable if sufficient improvement in the level of reliability can be achieved. The many ways of measuring volatility almost certainly differ in both bias and reliability. For example, a robust measure, like the average daily range, even though biased (in that it may measure something slightly different than intended or have an expectation that deviates somewhat from the correct value) may be sufficiently more reliable to compensate for the slight loss in validity. This is the essential idea behind the use of biased estimators. Volatility measures, including the average daily range, a biased estimator, are analyzed for reliability in several of the studies that follow. One goal *en route* to developing the best possible estimator of future volatility is to determine the best (in the sense of most reliable with fewest data points) measures of historical volatility and other important variables.

## MODEL COMPLEXITY AND OTHER ISSUES

Another issue to consider in the quest for better estimates of future volatility has to do with the complexity of the estimation model. No doubt, a predictor variable should have high reliability. However, to obtain the best possible estimate of future volatility, more than one such predictor may be necessary.

One effective way to combine several predictors is with multiple regression. A model based on multivariate regression is likely to provide better estimates of future volatility than can be achieved with a univariate (single predictor) model. To begin with, a multivariate model can employ historical volatilities for several periods as independent variables. Data of the type adduced in Chapter 4, such as that related to holding period, seasonality, and day of the week, are easily integrated within the context of a multivariate regression model. Even nonlinear relationships amongst the variables can be managed with appropriate techniques. A well-designed multivariate regression is a good way to determine expected future volatility, given a set of relevant independent variables.

Multivariate regressions can be complex, in the sense of having many input variables and model parameters. It is well known that with model complexity comes the danger of curve-fitting. Curve-fitting is of little concern in the investigations presented herein because data points are so numerous, while even the most complex models contemplated involve only a handful of variables.

In addition to examining the use of single predictor approaches, including those based on standard historical volatility and on more reliable measures thereof, models with multiple independent variables and nonlinear relationships will be explored in the quest for improved estimation. The issue of what to use as a dependent variable—the "target," or measure of future volatility—will likewise be considered. One possible choice for the dependent variable is future volatility calculated in the same manner as historical volatility, using the equations from Chapter 4. An alternative and possibly better choice might be the volatility implied by the terminal price of an at-the-money straddle. Both choices with regard to future volatility will be examined as dependent variables in the development of predictive regression models. Finally, the reader should note that more popular methods of forecasting volatility, such as GARCH models, will not be covered, since extensive coverage of these methods already exists in the literature on the subject.

## EMPIRICAL STUDIES OF VOLATILITY

The series of studies to follow concerns the very real problem of estimating future volatility for the purpose of pricing options.

These studies attempt to answer many significant questions including: How good are raw historical and implied volatility as inputs to a pricing model or as predictors of near-term future volatility? How can better estimates of future volatility, which will provide more accurate appraisals of option worth, be calculated? How reliable are various measures of volatility? And how does the statistical phenomenon of regression to the mean affect volatility and option value?

## Software and Data

All analyses were carried out in generic C++ and the same code libraries as those employed in earlier chapters were used. In some instances, final calculations were performed in Microsoft Excel spreadsheets. Excel was also used to prepare charts and to format tables for presentation.

Two binary databases were employed in the studies that follow. The first binary database contained price and volume data for 2,241 stocks. There were 1,834 bars in the period spanned, which ran from January 1, 1996 to April 15, 2003. These data, originally extracted from the Worden Brothers TC-2000 package, were thoroughly checked for errors and found to be quite clean. Details regarding how the data were extracted, examined for errors, and scrubbed cleaned can be found in Chapter 4. The second binary database consisted of implied volatility figures corresponding to every bar of every stock in the first binary database. The figures in the second database were computed using raw end-of-day option quotes obtained from www.stricknet.com.

## Calculation of Implied Volatility

*Implied volatility* is the volatility that, when entered into an option pricing model, yields the observed market price of an option. It is usually calculated for each of several active options trading on a given security at a given time. The individual implied volatilities may then be used by the trader or hedger to compare the relative costs of the options or they may be combined into a general (and more reliable) measure of implied volatility, as was done when constructing the database described below.

Implied volatility numbers in the second binary database were calculated in the following manner: First, one of the stocks in the first binary database was selected. A bar was chosen for that stock. An options chain was then retrieved from a large options database that was assembled using data acquired from www.stricknet.com. The options chain consisted of data for every option trading on the selected stock on the day corresponding to the selected bar. Each option in the chain was examined. If an option had between 3 days and 48 days left to expiration, and if the bid was greater than zero, then an implied volatility was calculated and multiplied by a weight. The multiplier weight was the amount of speculative value in the option divided by the square root of the time remaining before expiration. The weighted implied volatilities were then summed and the result divided by the sum of the weights. This calculation produced a weighted average that was the implied volatility figure. Implied volatility figures were determined separately for calls and for puts. The figures for the selected stock at the chosen bar were then saved in the second binary database. At this point, processing moved to the next bar of the currently selected stock. When all bars of the selected stock were processed, the next stock was selected and the sequence of steps outlined above was repeated. Repetition continued until an attempt had been made to compute implied volatility figures for every bar of every stock. It should be noted that the second binary database contains only about one year of implied volatility figures due to options data availability.

The intent behind the weighting scheme used when calculating implied volatility was to give greater emphasis to at-the-money or near-the-money options with substantial speculative (time) value than to options that were either far out-of-the-money or deeply in-the-money. At-the-money options have the greatest amount of time value and are often the most liquid; thus, their prices most reliably reflect the volatility-dependent speculative component of option value. Options with no bids were not analyzed since their true market price is indeterminate. Options with less than three days of life were ignored because they have little time value and may have more erratic market prices (especially if strikes sufficiently close to the stock price are not available), while those with more than 48 days were disregarded because they are typically much less active and also fall outside the short-term trader's timeframe that is the focus of this book.

## STUDY 1: UNIVARIATE HISTORICAL VOLATILITY

Historical volatility is often used when calculating a theoretical option price. Yet the value of an option ultimately depends not on past volatility, but on future volatility. How good is raw historical volatility as a proxy for future volatility? Are there conditions under which the use of historical volatility will lead to serious pricing errors? It is possible to correct historical volatility figures to obtain better estimates of future volatility and, therefore, of option value? These are some of the questions that this study attempts to answer. Additionally, this study compares several measures of both historical and future volatility. Comparisons are made between standard and average range volatility as historical measures, and between standard, average range, and implied terminal straddle price volatility as measures of future price movement.

### Method

Data analysis was performed in three computational blocks. The steps carried out in the first computational block were as follows: First, a stock was selected. A bar from that stock was then chosen as a reference bar.

Several criteria were used when choosing reference bars. Data points from the first $m$ bars of each stock's data series were avoided as reference bars because they were retained for the calculation of historical volatility. Likewise, the last $n$ bars of each stock's data series were reserved for the calculation of future volatility. In addition, the stock had to be actively trading during all $m$ bars in the period preceding any reference bar and, at no point during this period, could the price fall below $2.

These criteria ensured that the data analyzed were from periods during which the stocks were alive and trading above penny stock levels. Given 1,834 bars (days) per stock and 2,241 stocks, the number of reference bars was 4,109,994 less those eliminated from consideration by the aforementioned criteria.

Once a stock was selected and a reference bar chosen, historical volatility (the independent variable) and future volatility (the dependent variable) were calculated. Historical volatility was calculated over a 30-bar period ($m = 30$) ending at the reference bar. Future volatility was determined from a 10-bar period ($n = 10$) immediately following the reference bar. The historical and future volatility figures were saved to a scratch file.

At this point, the next reference bar was chosen and the historical and future volatilities were again calculated and saved. This process was repeated until the last reference bar of the currently selected stock was reached, and then another stock was selected. The sequence continued until historical and future volatility had been calculated for every valid reference bar of every stock in the database.

In the second computational block, historical volatility figures (derived from the above procedure) were sorted into 50 equally spaced levels or bins. The lowest bin was centered at 0.05 and the highest bin was centered at 2.10, which covered most of the range of historical volatility observed. For each level of historical volatility, the mean and standard deviation of future volatility was ascertained. For example, in the analysis employing standard historical and future volatility, bin 10 contained 211,087 data points with a historical volatility centered at 0.468, and was associated with a future volatility having a mean of 0.479 and a standard deviation of 0.252. Bin 45 was centered at 1.933 and had only 1,375 historical volatility numbers falling in its range; for those cases where historical volatility fell in this bin, the future volatility had a mean of 1.153 and a standard deviation of 0.619. For every level of historical volatility, this bin-by-bin analysis provided a reasonably stable estimate of the mean and standard deviation of the corresponding future volatility.

The analysis performed in the third computational block was that of a second-order polynomial regression. Data points were not binned. Coefficients $a$, $b$, and $c$ were computed to make $\hat{y} = cx^2 + bx + a$ a least-squares estimator of $y$, where $y$ was the observed future volatility and $x$ the historical volatility. The multiple correlation, which measures the goodness of fit, was also determined.

A total of six analyses were performed as described above. Each analysis examined the relationship between a particular kind of historical volatility and a specific measure of future volatility. Two measures of historical volatility were analyzed: one was the standard historical volatility, based on the standard deviation of logarithmic returns; the other measure was average range volatility, which is based on daily price ranges. The standard historical volatility was computed using Equation 4.2 in Chapter 4; every options trader is familiar with this kind of historical volatility. Historical volatility based on the average range was computed as

$$v = 0.627 \frac{1}{m} \sum_{i=0}^{m-1} \ln \frac{H_{k-i}}{L_{k-i}} \tag{5.3}$$

where $m$ represents the lookback period, $k$ the index of the reference bar, $H_{k-i}$ the high of bar $k$–$i$, and $L_{k-i}$ the low of bar $k$–$i$.

Three measures of future volatility were tested. The first two measures were the same ones used for historical volatility, but calculated from the $n$ bars directly following the reference bar rather than from the $m$ bars immediately preceding it. The third measure of future volatility was based on the terminal price of an at-the-money straddle.

The third measure of future volatility was, in fact, the implied volatility of an at-the-money straddle with $n$ days of life remaining. The implied volatility was calculated from the terminal price, not the market price, of the straddle. In performing the calculation, strike prices for the options comprising the straddle were set equal to the stock price at the reference bar. The implied volatility was the volatility that, when entered into Black-Scholes, produced a theoretical price equal to the terminal price of the straddle $n$ bars after the reference bar. Because volatility is linearly related to the price of an at-the-money straddle, the expectation of the volatility implied by the terminal price is equal to the volatility implied by the expected terminal price. Given that the mean or expected terminal price is the theoretical value of an option, the implied volatility calculated from the terminal price of an at-the-money straddle is precisely what must be estimated in order to obtain accurate theoretical option prices.

Analyzed psychometrically, the only problem with the straddle-based measure of future volatility is that it is based on only one test item and, therefore, has low reliability. Reliability, however, is of minor concern in the present instance, when the measure is used as a dependent variable. When dealing with measures used as dependent variables, bias and construct validity are of much greater concern. Considered from the perspective of bias and construct validity, the straddle-based volatility measure wins hands down.

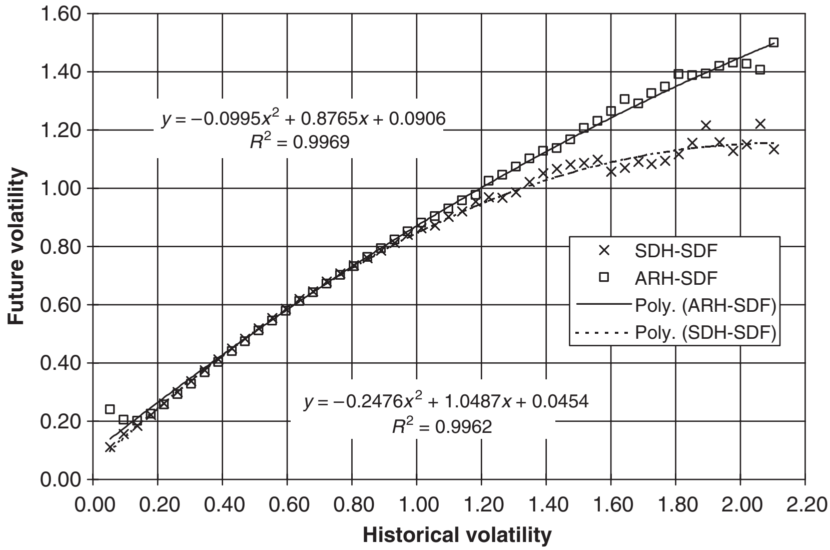
## Results

Figure 5–1 shows the expected level of future volatility ($y$-axis) for each bin or level of historical volatility ($x$-axis). Two data series are presented in the figure. One data series illustrates the relationship between standard historical volatility (SDH) and standard future volatility (SDF), represented by data points tagged by small X's. The second data series, marked by small rectangles, depicts the relationship between average range historical volatility (ARH) and standard future volatility (SDF). The smooth curves drawn through the data points are second-order polynomial regression lines. The dotted curve corresponds to the data points marked by Xs and expresses the statistical relationship between standard future and historical volatility; the solid curve is the regression between the more robust average range historical volatility and standard future volatility. The regression equations, which express future volatility ($y$) as a function of historical volatility ($x$), appear in Figure 5–1 along with the squared multiple correlations ($R^2$).

***Regression to the Mean*** Three observations are readily made when examining Figure 5–1. The first observation involves *regression to the mean*, a phenomenon well known to statisticians. Volatility clearly reverts to the mean. The phenomenon can

**FIGURE 5–1**

Standard Future Volatility (SDF) as a Function of Standard (SDH) and Average Range (ARH) Historical Volatility




be seen in the fact that low levels of historical volatility are followed by higher levels of future volatility, while high levels of historical volatility are followed by lower future levels. For instance, with a standard historical volatility of 0.20, the expected future volatility is around 0.25; a historical volatility of 0.80 decays to a future volatility close to 0.72. Future volatility tends to revert to its mean value, which is approximately 0.50 or 50%. Regression to the mean is more severe with standard historical volatility than with average range historical volatility because the standard measure is less reliable.

***Quadratic (Nonlinear) Relationship*** The second observation derived from Figure 5–1 is that the relationship between future and historical volatility is nonlinear; in fact, it is quadratic. A quadratic regression yielded an exceptionally high $R^2$ for both data series: 0.9969 for the average range (ARH-SDF) and 0.9962 for the standard (SDH-SDF). When tested, higher order polynomials produced little or no improvement.

The nonlinear relationship between historical and future measurements suggests that mean reversion becomes exaggerated beyond normal when volatility is very high. There are several possible explanations. One possible explanation for the exaggerated mean reversion is that the occasional high volatility event (such as might be attributed to an earnings report) is usually followed by a rapid collapse back to normal or baseline levels. From the trader's perspective, stocks usually exhibit mildly varying baseline volatility and only occasionally become extremely unstable. Another explanation for the nonlinear capping effect, seen in Figure 5–1, is the exploitation of high levels of volatility by traders. Extreme volatility is probably not sustainable because it represents an inefficiency that may be exploited for a profit and eliminated or, at the very least, attenuated. Finally, however measured, volatility has an asymmetric distribution; it can go up without limit, but down only to zero.

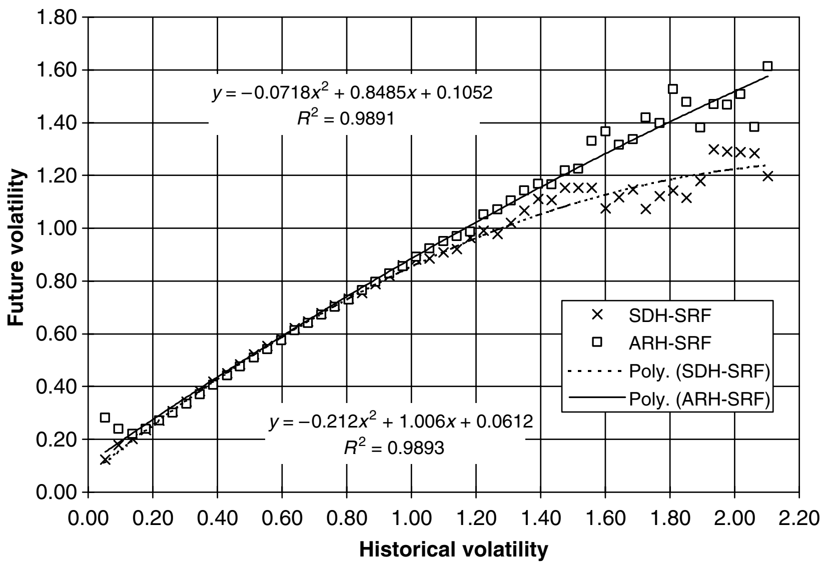
***Changing Relationship with Changing Volatility*** The third observation based on Figure 5–1 involves the change in the relationship between historical and future volatility at different levels of historical volatility. Between about 15% (0.15) and 100% (1.00), both measures of historical volatility provide almost identical results with overlapping data points. In this range, either measure may be used to estimate future volatility. As the level of historical volatility exceeds 100%, the curves diverge; the curve for the standard measure displays greater droop than does the curve for the average range. The divergence might be due to average range volatility being less sensitive to short bursts of activity. To achieve a high level of average range volatility, there must be sustained price movement. Such movement probably tends to linger longer into the future than the short-lived spikes that can easily inflate the standard historical volatility measure. At the left end of Figure 5–1, the standard measure can be seen to provide a better prediction of future volatility than the average range. However, this is only true for levels of historical volatility below 15% or 20%. The reason may be that even a single quiver in stock price can drive up the standard measure but will have little effect on the average range volatility. A stock that exhibits even one such quiver is more likely to have wider future price swings than one that is steadfastly quiet.

### *Straddle-Based versus Standard Future Volatility*

Figure 5–2 again depicts the relationship between future and historical volatility. The only difference in the data underlying Figure 5–2 and Figure 5–1 is that, in the former, future volatility is based on what is implied by the terminal price of an at-the-money straddle that has 10 bars of life left until expiration. It is easily seen that the data series in these figures have the same overall contour and the regression equations that define the polynomial smoothings are similar. The regressions in Figure 5–2 have a somewhat lower $R^2$ due to a greater dispersion of data points about the regression curves. This is a consequence of the lower reliability of the straddle-based measure—straddle-based volatility is a one-item measure, in the psychometric sense. The mean future straddle-based volatility was 0.507, very close

**FIGURE 5–2**

Straddle-Implied Future Volatility (SRF) as a Function of Standard (SDH) and Average Range (ARH) Historical Volatility



to that of the standard future volatility. Lastly, the three observations made for Figure 5–1 also apply to Figure 5–2.

It should be noted that another chart (not shown) was generated with standard and average range historical volatility as the independent ($x$-axis) variables, and average range future volatility as the dependent ($y$-axis) variable. Except for higher $R^2$ levels in the regressions—a consequence of the greater reliability of the average range measure—the data in the chart appear very much like those shown in Figures 5–1 and 5–2. Likewise, the conclusions and observations are essentially the same and are, therefore, not presented.

***Longer-Term Historical Volatility*** The analysis shown in Figure 5–1 was repeated using 100-bar (rather than 30-bar) measures of historical volatility. Although the correlation between historical and future volatility was expected to be larger with 100 bars, this was not the case; however, regression to the mean was slightly attenuated for historical volatility levels below 80%. The reason for expecting a higher correlation and less mean reversion involves measurement reliability. Imagine that, over relatively short periods, a stock has some "true" underlying volatility that is fairly constant. A trader assessing it over a 100-day period is comparable to a psychometrician measuring a personality trait using a questionnaire containing 100 items. In other words, in the analysis, there were 100 somewhat independent measurements of the hypothetical true volatility. The number obtained from these 100 one-day measurements should be more reliable and, if the assumption of a relatively constant underlying volatility is correct, more valid than volatility calculated from a 30-bar sample. In the market, however, validity tends to be lower with longer periods because longer samples are less representative of current market conditions. Reliability, however, is still likely to show a gain with an increase in the number of bars or test items. In the analysis of the 100-bar measures, the decrease in validity, due to changing levels of volatility over time, roughly neutralizes the increase in reliability. The greater reliability of the longer-term volatility measurement, however, is sufficient to somewhat attenuate the reversion to the mean. If the regression lines in Figure 5–1 are extrapolated to the $y$-axis, the zero intersection is not reached; instead, the lines intersect the axis at a future volatility between 5% and 10%. The equivalent regression lines for the 100-bar analysis fall closer to zero when similarly extrapolated, thus providing further evidence of reduced mean reversion. However, the decay of very high volatility in the future is not much different and, overall, the results are quite similar to those found with the 30-bar historical volatilities that were presented in Figure 5–1.

***Raw Data Regressions*** The regressions in Figures 5–1 and 5–2 were calculated (in Excel) for the binned data points actually depicted in the figures. Regressions were also computed directly (in the third computational block) from the raw data points. The equations generated from the raw data were quite similar to those for the binned data, indicating that the same curvilinear relationships were being described, although the $R^2$ was much lower for the former. A lower $R^2$ for the raw data was anticipated since each raw data point represents an individual measurement, while the data points in the figures represent averages of future volatility for all cases in a given bin. All regressions had a large positive linear component, a small negative squared component, and a small positive intercept. The multiple correlation (not squared) was highest ($R = 0.808$) for the regression relating average range historical to average range future volatility. Given that the average range measure is the most reliable, this was not surprising. The regression equation for estimating the most construct-valid measure of 10-bar future volatility (based on the implied volatility of an at-the-money straddle) from the 30-bar average range historical volatility is $\hat{y} = 0.084 + 0.851x - 0.060x^2$. This regression, calculated over 3,395,614 data points, yielded a multiple correlation of 0.395. Given the large number of data points, the regression should be quite stable.

## Discussion

For almost the entire spectrum of historical volatility, the average range provides a better estimate of future volatility than does the standard measure. The factor responsible is the greater reliability and robustness of the average range as a measure of volatility. The average range is more reliable because the average deviation holds up better than the standard deviation when the underlying distribution differs from normal, and because each item is itself a complete test of volatility that takes into account more data—the entire range of a day—than does a single close-to-close return.

Although commonly used when pricing options, raw historical volatility is a poor predictor of future volatility. This is true regardless of how it is measured. The situation worsens as historical volatility moves away from around 50% and becomes especially bad when it exceeds 100%. When historical volatility is high, future volatility tends to fall below the historical level. Consequently, options tend to be worth less than a standard pricing model using raw historical volatility might indicate. Likewise, when historical volatility is low, future volatility is likely to be higher and option value greater than might be expected. The use of raw (uncorrected) historical volatility as an input to a standard pricing model is a bad practice that may lead to erroneous theoretical option premiums.

Given the findings, the standard measure should be used to price options when historical volatility is extremely low (below 15%). For all other levels of historical volatility, the average range measure is more appropriate. Regardless of the measure used, historical volatility must be corrected before it is employed as an estimate of, or proxy for, future volatility. Much better estimates of expected future volatility and, in turn, theoretical option price can be obtained by correcting historical volatility using the relationships described in the regressions and depicted in Figures 5–1 and 5–2.

How is historical volatility corrected for use in pricing an option? Assume that the observed standard historical volatility for a stock is 150%—not an uncommon figure in the technology sector. Should this number be directly entered into an option pricing model? No! It is easy to determine from the curve in Figure 5–1 that a standard historical volatility of 150% corresponds to an expected future volatility of around 110%. Alternatively, the regression equation shown in the figure could have been used to calculate the expected future volatility. It is the expected future volatility, not the historical volatility that must be entered into a model like Black-Scholes in order to establish the theoretical price for an option. How great would the error be if Black-Scholes were calculated using raw historical volatility, as it often is, rather than with expected future volatility? Would it be significant enough to matter to a trader? Suppose the stock was trading at $50, and a $55 call with one month of remaining life was under consideration. With a volatility of 150%—the raw historical volatility in the imagined scenario—Black-Scholes prices the option at $6.36. A more accurate theoretical price of $4.16 is obtained when the 110% expected future volatility is used in the calculation. Should the trader be concerned with a 53% pricing error? Is it worthwhile estimating expected future volatility? You be the judge.

An implication of the above is that, when historical volatility is very high, selling premium might not be a bad idea, especially if market prices are tracking standard models using uncorrected historical volatility. Likewise, in low historical volatility situations, if option prices are tracking raw historical volatility, a trader might wish to buy premium in the expectation that future volatility will revert to its higher mean.

It should be noted that, for the sake of simplicity, a 10-bar period was settled upon for future volatility in the studies described above. Obviously, a trader using an option model is not going to price only options having 10 bars of life remaining. Analyses like those presented, however, may easily be carried out for options whose remaining life spans any desired period. Tests have demonstrated similar results when other potential holding periods were studied. It is not that difficult to prepare a set of analyses covering a range of time frames that can provide all the information necessary for pricing options with different amounts of life remaining. Given the task at hand, however, it may be better to construct a more unified nonlinear regression or neural network estimator that takes as one of its inputs the period over which future volatility is to be reckoned.

## STUDY 2: BIVARIATE HISTORICAL VOLATILITY

The previous study involved a univariate model in which a single, short-term measure of historical volatility was used to estimate expected future volatility. Can the addition of a longer-term measure of historical volatility improve the quality of estimation? The primary objective of the current study is to discover whether a multivariate model that employs two measures of historical volatility—one a short-term measure, the other a long-term one—can yield a better estimate of future volatility than a univariate model that employs only one measure.

A secondary objective of this study is to further examine mean reversion. In Study 1, future volatility appeared to revert to the mean of the entire market. Perhaps, even more than it reverts to the mean of the whole market, short-term historical volatility regresses to the mean of a given security's long-term volatility. It seems reasonable to suppose that securities have characteristic volatilities that are structurally and fundamentally determined, and that persist over extended periods. If this is the case, it gives further weight to the supposition that a model that employs both a long-term and a short-term historical measure can yield better predictions of future volatility than a model that employs either measure alone.

## Method

The analysis was performed in two computational blocks. The first computational block began with the selection of a stock and a valid reference bar. A reference bar was valid if there were at least $m1 + m2$ bars preceding it and $n$ bars following it, if no price was less than \$2 in the $m1+m2$ preceding bars, and if the stock was active during the entire period. Once a valid reference bar was selected, the short-term historical volatility was calculated from the $m1$ bars immediately preceding the reference bar. A long-term historical volatility was computed from the $m2$ bars that preceded the first bar used to calculate the short-term measure. Future volatility was determined from the $n$ bars that immediately followed the reference bar. In this study, $m1$ was 30, $m2$ was 70, and $n$ was 10. Standard volatility, based on the second moment of logarithmic returns, was used for all measurements. The three volatility measurements (two historical and one future) were saved to a scratch file. At this point, the next reference bar was selected. When all valid reference bars for a given stock had been processed, the next stock was chosen. The sequence of steps was repeated until all three volatility figures had been calculated for every valid reference bar of every stock.

The second computational block involved the binning procedure. Because there were two independent variables, the bins formed a two-dimensional grid, rather like a spreadsheet. The bins in each row corresponded to a distinct level of short-term volatility, while those in each column corresponded to a specific level of long-term volatility. The bin in the upper left corner was centered on both a short- and long-term historical volatility of 0.05. The center of the bin at the lower right corner fell at 2.05. Since the bins formed a two-dimensional grid, there were far more of them than in Study 1. This meant that fewer data points would be assigned to any one bin and that many bins would likely have no data points at all assigned to them. To avoid a lot of empty or inadequately filled bins, the range of volatility was subdivided into only 20 levels, rather than the 50 used in Study 1; in this way, the total number of bins was reduced to 400 ($20 \times 20$) and the number of data points falling in each bin was increased. There were two numbers associated with each bin: one was the number of instances in which the two historical volatilities fell within the two ranges associated with the bin; the other was the average (expected) future volatility for those data points.

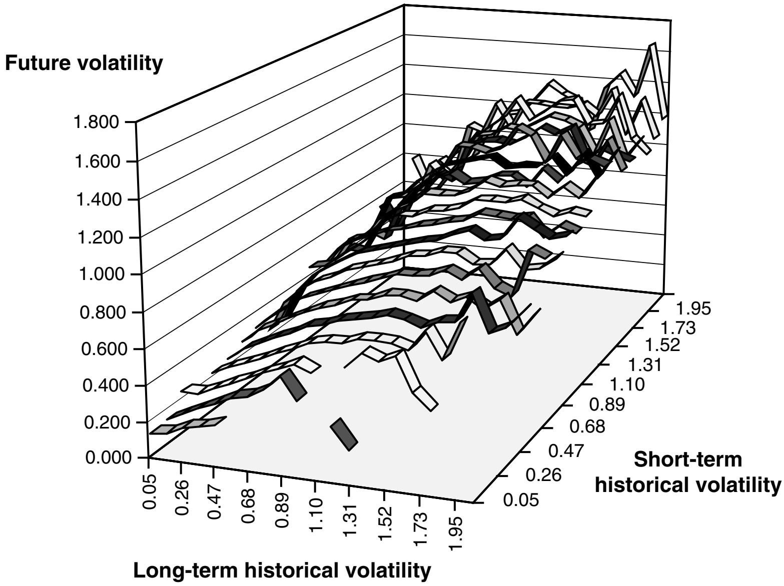
## Results

Figure 5–3 depicts short- and long-term historical volatility as predictors of future volatility. Data from each bin is individually represented in this chart. The $x$-, $y$-, and $z$-axes represent the short-term historical, long-term historical, and expected future volatilities, respectively. Each ribbon in the three-dimensional space of the chart illustrates the relationship between short-term historical and expected future volatility for one level of long-term historical volatility. Collectively, the ribbons define a volatility surface in three-dimensional space.

Figure 5–4 shows a subset of the data from Figure 5–3. The $x$-axis represents short-term historical volatility, the $y$-axis

**FIGURE 5–3**

Standard Long- and Short-Term Historical Volatility as Predictors of Standard Future Volatility



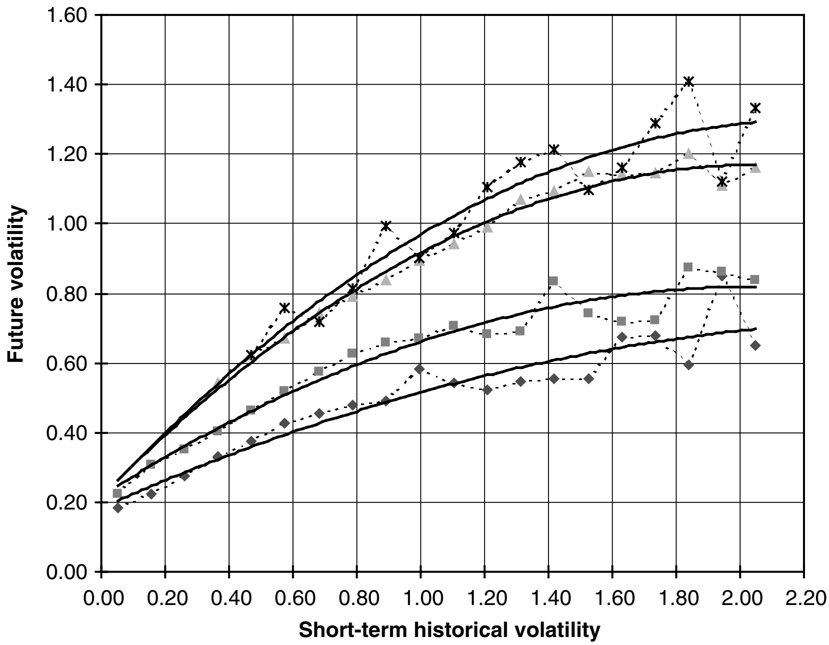

represents expected future volatility, and the four data series (dotted lines and markers) correspond to four selected levels of long-term historical volatility. From bottom to top, the data series are for long-term historical volatilities falling into bins centered at 0.261, 0.471, 0.992, and 1.839. As in Study 1, simple polynomial regressions were used to smooth the data series; they appear as solid lines.

The raw data points in Figure 5–4, which represent averages taken from the two-dimensional bins, have a much greater variation about their polynomial regressions than was the case with the data points in Figure 5–1 or 5–2. The reason is that each two-dimensional bin in the current study contained fewer cases over which to average future volatility than did the one-dimensional bins in the previous study.

***Independent Contributions*** Figure 5–3 shows that each measure of historical volatility can independently account for

**FIGURE 5–4**

Standard Future Volatility as a Function of Long- and Short-Term Standard Historical Volatility




variation in expected future volatility and, therefore, should contribute to a model's ability to generate accurate predictions. For any level of short-term historical volatility, future volatility tends to increase with increasing long-term historical volatility. Likewise, with few exceptions, for any level of long-term historical volatility, future volatility increases with increasing short-term volatility. The exceptions are for very low levels of long-term volatility, where future volatility rises with short-term volatility up to a point and then declines. Future volatility is greatest when both short- and long-term historical volatility levels are high, and least when both historical volatilities are low.

The unique contribution of each independent variable to the level of the dependent variable is also clearly evident in Figure 5–4. For any of the four levels of long-term volatility, future volatility generally rises with rising short-term historical volatility, i.e., as the chart is scanned from left to right. Likewise, when short-term historical volatility is fixed (when that chart is scanned from bottom to top), higher levels of long-term volatility are associated with higher levels of expected future volatility. In other words, when one independent variable is held constant, the other continues to contribute useful predictive information. This observation is quite significant, since two independent variables can provide better prediction than one variable only if each can demonstrate a correlation with the dependent variable that persists when the other is held constant.

***Reversion to Long-Term Mean*** Figure 5–4 shows how the level of long-term historical volatility influences the relationship between short-term and future volatility. When long-term historical volatility is around 26% (represented by the bottom curve), future volatility varies only between 20% for very low short-term levels, to about 70% for extremely high short-term levels. Short-term levels greater than about 30% result in lower future levels, while those less than 30% are followed by higher levels of expected future volatility. In other words, future volatility appears to revert to a mean near 30% when long-term volatility is close to 26%. When long-term volatility is near 184% (as represented by the top curve in Figure 5–4), future volatility has a greater range and higher average level, running from about 40%, when short-term volatility is 5%, all the way up to 120%, when short-term volatility is near its maximum. For the curve representing 184% long-term volatility, the "fixed point," where short-term and expected future volatility meet, appears to be around 100%. Clearly, future volatility seems to revert to a level somewhere between the mean of the market and a stock's own mean as represented by its long-term historical volatility. Furthermore, low long-term levels seem to be more enduring than high ones. This makes sense in that certain stocks (e.g., utility stocks) are unlikely to change fundamentally into the kind of stocks that have high volatility; but high volatility stocks (e.g., new technology issues) can and do change into more stable stocks over time.

### Discussion

The results suggest that stocks, especially utility stocks and other relatively placid issues, have characteristic levels of volatility that tend to endure over long periods. Future volatility tends to revert to these characteristic levels as well as to the mean volatility of the market as a whole.

The statistics also demonstrate that each measure of historical volatility can explain variation in future volatility even when the other is held fixed. A model that employs the two historical volatility measures used in the study will, therefore, predict future volatility more accurately than a model that employs only one. Given the levels of short- and long-term historical volatility, the expected future volatility may be ascertained from a chart (like the one in Figure 5–4) or from a regression (as in Study 1). To make this work in the current instance, however, charts or regressions are required for each of many levels of long-term historical volatility. One solution to the problem of a family of charts or regressions is investigated later in this chapter, in Study 4. Despite the complexities, the findings are encouraging in that they indicate progress in the search for a better estimator of expected future volatility.

## STUDY 3: RELIABILITY AND STABILITY

Several different volatility measures appear in the studies presented in this chapter. All are presumed to reflect roughly the same thing—the amount of price movement occurring per unit of time—with varying degrees of reliability. Exactly how reliable are these different measures? In one instance, the better estimation of future volatility provided by the average range was explained in terms of that measure's higher reliability. Is such an explanation justified? How similar are the different measures in what they actually assess? Is the assumption that they all measure essentially the same hidden variable a reasonable one? Finally, how stable is volatility when considered as a trait? These kinds of questions are the focus of this study.

The current study has three specific objectives. The first objective is to determine reliability coefficients for each of the three price-based volatility scales used in the studies. For the standard and average range scales, reliability can be calculated with the split-half technique discussed earlier. The split-half method cannot be used for the straddle-based measure because straddle-based volatility is a one-item test; in that case, reliability must be inferred by more complex methods. The second objective is to estimate how similar the underlying traits that the different volatility scales measure are to one another; this involves calculating correlations between perfectly reliable versions of the three measures of interest. The problem is that only real, and thus less than perfectly reliable, measures exist. However, it is possible to estimate theoretical correlations between underlying traits measured by imperfect (real) scales from the observed correlations between these scales together with each scale's reliability. Such theoretical correlations are referred to as *attenuation-corrected*, or as having been *corrected for attenuation*. The third objective in this study is to calculate stability coefficients for the standard and average range volatility measures. Again, the problem is one of estimating a theoretical correlation: that which exists between two perfectly reliable measurements of volatility made at two points separated in time.

## Method

First, a stock was selected. Then, a valid reference bar was chosen. A reference bar was considered valid if there were at least $m$ bars preceding it, if no price in those $m$ bars fell below $2, and if the stock was active throughout that $m$-bar period. Volatility was first calculated from the odd-numbered bars in the $m$-bar period immediately preceding the reference bar, and then from the even-numbered bars. The calculations were done for both the average range and the standard measures. The result was four volatility scores: the odd bar standard, the even bar standard, the odd bar average range, and the even bar average range. These four numbers were saved to a scratch file. At this point, the next valid reference bar was selected. When all reference bars for a given stock had been processed, the next stock was drawn and the sequence was repeated. The result was an array, saved in a file, with four columns that corresponded to the four volatility scores and with over 3,000,000 rows (the exact number dependent on $m$) that corresponded to the stocks and reference bars.

Correlations were calculated for the saved volatility scores using Equation 5.1. Volatility scores (columns) were correlated over all cases (rows) to produce a correlation matrix. The matrix was written to a simple text file. Since the entire procedure just outlined was carried out for three values of $m$ (5, 10, and 30), a file containing three correlation matrices was produced. Recall that $m$ specified the period of, or the number of bars in, each of the volatility measures. Once computed and written to the file, the correlation matrices were loaded into an Excel spreadsheet. Also copied into the spreadsheet were correlations between historical and future volatility that had been determined in previous studies in the course of performing raw-data regressions.

Equation 5.2 was used to determine split-half reliability coefficients for the standard and average range volatility measures from the correlations between the odd and even subscales. This resulted in three sets of reliability coefficients. Each set corresponded to a level of $m$, while each coefficient within a set corresponded to a particular measure of volatility. The 10- and 30-bar reliability figures derived from the calculations are relevant to the 10- and 30-bar volatility measures used in Studies 1 and 2. Reliability data may be extrapolated to the 70-bar measures in Study 2 by means of Equation 5.2. In addition to reliability coefficients, the attenuation-corrected correlations between the two historical volatility measures were ascertained. There were three such correlations, one for each $m$ or period of measurement. An attenuation-corrected correlation may be computed as

$$r_{\dot{x}\dot{y}} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}} \tag{5.4}$$

where $r_{\dot{x}\dot{y}}$ represents the theoretical correlation between the true-score components of measures $x$ and $y$, where $r_{xy}$ is the observed (raw or uncorrected) correlation between $x$ and $y$, and where $r_{xx}$ and $r_{yy}$ are the reliability coefficients for $x$ and $y$, respectively.

Attenuation-corrected correlations between the historical 30-bar standard and average range measures and the equivalent 10-bar measures of future volatility were computed from the split-half reliability coefficients determined above together with the raw correlations copied from earlier studies. These attenuation-corrected correlations, of which there were four, were averaged to obtain a coefficient of stability.

The straddle-based measure's reliability was then estimated by solving for $r_{yy}$ in Equation 5.4. To solve the equation, $r_{\dot{x}\dot{y}}$ was set to the stability coefficient. Since reliability could not be directly determined for the straddle-based measure, the stability coefficient could not be determined. However, stability estimates could be and were calculated (see above) from the other volatility measures. These estimates were quite near to one another, and so the average was taken as an approximation of the stability coefficient that might have been obtained with the straddle-based measure, if it had been possible to directly calculate that measure's reliability. The observed correlation between historical and future volatility needed for $r_{xy}$ was taken from Study 1, while the reliability of the historical measure required for $r_{xx}$ was computed in an earlier step. Two estimates of straddle-based reliability were made with the procedure just described: one estimate was based on the standard historical and future measures and the other involved the average range measures.

Finally, to verify the results, a third estimate of reliability was made for straddle-based volatility using a very different procedure. The third estimate was computed with Equation 5.2 and entailed extrapolating the known reliability of the 5-bar standard measure to the reliability of a 1-bar measure. In addition, to verify the theory and calculations, reliability coefficients for the standard and average range measures were extrapolated from the 5- and 10-bar analyses to a 30-bar scale and the figures compared with those actually obtained when 30-bar volatility measures were directly analyzed.

## Results

Reliability coefficients for the standard 5-, 10-, and 30-bar measures are 0.58, 0.69, and 0.83, respectively; the coefficients are 0.89, 0.93, and 0.97 for the average range measure. Using Equation 5.2 to extrapolate from the 5- and 10-bar scales to the

30-bar measure yields reliability coefficients of 0.89 and 0.87 for standard volatility and coefficients of 0.98 and 0.97 for average range volatility. For the average range measure, the reliability extrapolations are virtually perfect, which affirms that the calculations and theory are valid in the current context. The extrapolations for the standard measure are satisfactory, although not perfect. The extrapolations fall short of perfection because the standard measure involves the square root of a sum of squares, whereas the extrapolation logic implicit in Equation 5.2 assumes that test items are simply summed, as is the case for the average range measure.

Straddle-based volatility has a reliability coefficient of 0.25 when inferred from Equation 5.4 either from the standard or from the average range data. When determined by extrapolation from the 5-bar standard data to a 1-bar test using Equation 5.2, a reliability of 0.22 is found for the straddle-based measure. The two figures agree acceptably well, given the great disparity in the methods by which they were arrived at, thus confirming the validity of the analysis. With a reliability near 0.25, the largest correlation that any predictor can have with the future volatility implied by the terminal price of an at-the-money straddle is around 0.50.

The average attenuation-corrected correlation between 30-bar historical and 10-bar future volatility is 0.80. This is the correlation that theoretically exists between perfectly reliable measures of volatility separated in time; it is the correlation between the underlying trait of volatility in the 30-bar period immediately preceding the reference bar and the same underlying characteristic in the 10-bar period that follows the reference bar. The aforementioned correlation is thus a measure of persistence in the underlying trait or parameter; it is a stability coefficient.

A correlation of 0.98, corrected for attenuation, exists between the standard and average range historical measures. This very high correlation reflects the extent to which the two distinct volatility scales measure the same underlying variable after adjustments are made for their less-than-perfect reliability.

## Discussion

There is no doubt that the average range is by far the most reliable measure of the three examined, especially when volatility must be appraised over short intervals. The straddle-based measure, being a one-item test, has the lowest reliability. Standard volatility is a much more reliable measure than the one derived from an at-the-money straddle, but does not attain the performance of the average range. How good is the performance of the average range? To achieve the reliability of a 10-bar average range, a standard measure taken over more than 30 bars would be required.

While straddle-based volatility may be excellent as a dependent or target variable that represents future market movement when determining a regression equation, the average range is probably the best practical choice for assessing historical volatility. This is especially true when historical volatility must be determined from a restricted number of bars. Any great concern over differences in what the average range and standard scales reflect—differences that are not merely hypothetical, but that have practical significance—should be assuaged by the remarkably high attenuation-corrected correlation that exists between the two volatility measures.

Volatility appears to be a relatively enduring and predictable characteristic, as evidenced by a rather high stability coefficient. At least when adjusted for nonlinearity, regression to the mean, and other statistical phenomena, this finding justifies the use of historical measures, especially the more reliable ones, in models designed to estimate future volatility.

Finally, the current study demonstrates the merit of applying theory and practice developed in one field to problems in another. For example, the work originally pioneered by psychologists studying mental tests is pertinent, not just to psychometrics, but to any arena—including the study of volatility—where accurate measurements depend on samples of noisy, error-laden test items or mini-measurements.

## STUDY 4: MULTIVARIATE PREDICTION

In an effort to build a better estimation model for future volatility and, in turn, option price, Study 4 uses what was learned from previous studies. Although the estimation model developed in the current study is not the last word, it is the first to use both historical volatility measures and exogenous variables in a sophisticated multivariate regression. As such, it represents a major step on the road to better volatility predictions and option valuations.

As in Study 2, the model examined here employs both long- and short-term historical volatility. However, the historical measures used in the current study are the more reliable average range measures, rather than the standard volatility measures employed in the previous study. Linear, second order, and interaction or cross-product terms are included amongst the independent variables in the regression. By incorporating both volatility measures in one integrated model, the need to have a family of regressions involving short-term volatility, one regression for each level of long-term volatility, is avoided. In addition to independent variables derived from the historical volatility measures, the first three harmonics of a three-month cycle are factored into the model in order to capture the quarterly seasonal rhythm in volatility discussed in Chapter 4. Analysis in terms of only the first three harmonics conserves degrees of freedom, requiring only six rather than the 252 that would be lost if the raw seasonality data from Chapter 4 were used. Analysis in terms of harmonics, therefore, has a stronger theoretical foundation and less potential for producing curve-fit results. Harmonic decomposition also efficiently handles the issue of cycle phase. Phase is critical since forward estimates must be made; in other words, the phase of the observed quarterly cycle must be advanced sufficiently to achieve correct alignment with the volatility look-ahead period.

Because 10-bar future volatility is the dependent variable, day-of-week effects are irrelevant and ignored. Given the purpose of this chapter, and for reasons of simplicity and space, results for holding intervals (periods of future volatility) of other than 10 bars are not presented. However, by changing a single parameter in the program that performs the calculations, the approach used in this study can be effectively applied to the estimation of future volatility over any interval desired.

## Method

As usual, a stock was drawn from the database and a valid reference bar was selected. A reference bar was valid if there were at least $m1 + m2$ preceding bars during which the stock was active and had no price less than \$2, and if there were at least $n$

following bars. The straddle-based measure of future volatility, the dependent variable, was determined from the $n$ bars that followed the reference bar. Short-term average range volatility was calculated over the $m1$ bars immediately preceding the reference bar, while long-term volatility was determined from $m2$ consecutive bars ending at the first bar used to compute the short-term volatility. Cross-product and second-order terms were also evaluated. In these calculations, $m1$ was 30, $m2$ was 70, and $n$ was 10. After all volatility-related figures had been calculated, the time of the year at the reference bar was converted to radians. January 1 was mapped to zero radians (or zero degrees), while December 31 corresponded to a number just short of $2\pi$ radians (or 360 degrees). The time of the year, measured in radians, was used to compute the first three harmonics of a quarterly seasonal cycle. All in all, the variables computed were as follows:

```
Y   = Future straddle-based volatility
X1  = Short-term average range historical volatility
X2  = Long-term average range historical volatility
X3  = X1 * X2
X4  = X1 * X1
X5  = X2 * X2
X6  = X1 * SIN (4 * T)
X7  = X1 * COS (4 * T)
X8  = X1 * SIN (8 * T)
X9  = X1 * COS (8 * T)
X10= X1 * SIN (12 * T)
X11= X1 * COS (12 * T)
```

In the above Fortran-like notation (used for consistency with program and spreadsheet code), `Y` is the dependent variable, `X1` through `X11` are the independent variables, and `T` is the time of the year in radians. Once computed, `X1` through `X11` were written to a scratch file followed by `Y`. The next reference bar was then selected and, when all reference bars for the current stock had been processed, the next stock was drawn from the database. This sequence was repeated until every reference bar of every stock had been analyzed, and the associated dependent and independent variables written. A scratch file containing a 12 column by 3,213,002 row array was the result.

By using data from the scratch file, a multivariate regression analysis was performed. The regression equation produced by the analysis was then employed to compute the expected future volatility for every data point. The resultant figures were sorted into 50 bins and the mean future volatility corresponding to each bin, or level of predicted volatility, was determined. In addition, the cyclic element was calculated for each day of the year using only the sine and cosine terms in the regression equation. When performing this latter calculation, the sines and cosines were not premultiplied by the short-term historical volatility, only by the regression weights.

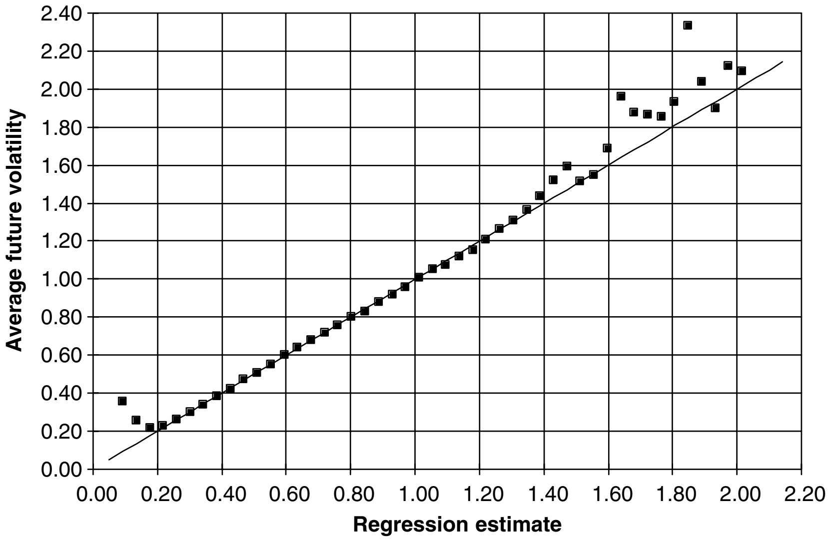
### Results

Each data point in Figure 5–5 shows the average observed future volatility (on the $y$-axis) corresponding to a specific bin or level of expected or predicted volatility (on the $x$-axis). In this case, the expected volatility is that predicted by the multivariate regression model. A solid, diagonal line that represents a perfect one-to-one match between predicted and observed volatility levels also appears in the figure for the sake of comparison.

For predicted levels between 18% and 140%, the relationship between average future volatility and regression-based expectation is almost perfect: the data points fall right on the diagonal line. When predicted levels are extremely low, actual levels of future volatility are greater than the regression estimates and lie above the line that represents a perfect relationship. The excess future volatility appears as a hook near the left end of the chart. Such a hook was also observed in previous studies, when the average range was used as the historical volatility measure. For high levels of expected future volatility, the data points are more scattered and lie slightly above the diagonal line. Increased scatter is a consequence of fewer raw data points falling in the relevant bins. Slightly higher levels of observed future volatility with high levels of predicted volatility imply the presence of some mild nonlinearity that is unaccounted for in the regression by the second-order terms. Perhaps future volatility has a nonlinear relationship with the cycle harmonics, for which only linear terms were entered into the model.

**FIGURE 5–5**

Average Future Volatility Levels versus Multivariate Regression Estimates

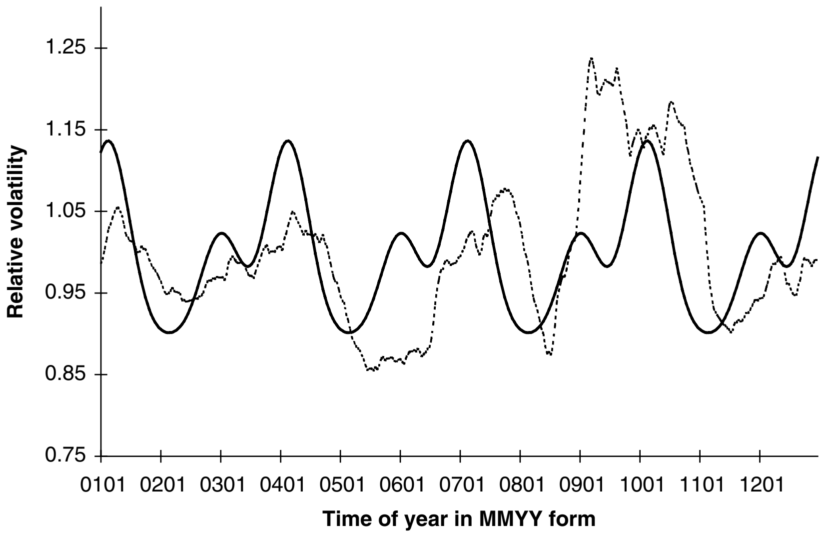



It is very encouraging that, in the present study, the relationship between average and predicted future volatility remains relatively linear—even if not perfectly so—up to far higher levels of volatility than in previous studies. In previous studies, the curves that represented relationships between various independent variables and observed future volatility flattened out at a level between 120% and 140%, never reaching any higher. In other words, regardless of the values taken on by the independent variables, an average future volatility of greater than around 140% could not be achieved due to what was referred to as a "capping effect." In the current study, the capping effect is gone; predictions for volatility levels as high as 200% can be made with reasonable confidence.

The cyclic component of the multivariate model is plotted in Figure 5–6 along with the raw seasonality data from Chapter 4. In the figure, the smoother, solid line represents a weighted sum of the six cycle-pertinent regression terms (those containing sines and cosines). The more ragged, dotted line is the raw seasonal

**FIGURE 5–6**

Relative Volatility versus Time of Year




volatility observed in the market. Time of the year is represented on the $x$-axis, while relative volatility is measured along the $y$-axis.

The cyclic component of the multivariate regression leads the actual market cycle, just as it should in a model designed to predict volatility down the road. Cycle highs appear in January, April, July, and October, between the 4th and 5th of the month. Given that a 10-bar holding period is equivalent to 14 calendar days, the center of the look-ahead interval should be seven calendar days, thus placing the predicted cyclic volatility high around the 11th or 12th of the month. Cycle lows occur almost exactly one month after the highs. The amplitude of the cycle-induced swings is not inconsequential. From the valley at 90% relative volatility to the peak at 114%, there is a 24% spread, certainly sufficient to be of interest to options traders.

Because the cyclic component is computed only from a three-month cycle and its second and third harmonics, each of the four cycles that occur in a year are identical; this is not the case for real market volatility cycles, which vary somewhat from quarter to quarter. A model using a one-year cycle and its

12 harmonics would capture these variations, but would require the calculation of many more regression coefficients and, therefore, would consume many more degrees of freedom. As degrees of freedom are consumed, the risk of curve fitting (in the bad sense) increases. Given a data sample extending back by only seven years, it seemed prudent to limit the number of regressors and conserve degrees of freedom.

The regression produced a multiple correlation of 0.40 between predicted and observed future volatility. When corrected for attenuation due to the less than perfect reliability of the straddle-based measure of future volatility, the multiple correlation was just over 0.80 and almost equal to the stability coefficient. A correlation this high would require a perfect historical measure to achieve, were historical volatility the only input to the model. Of course, there were other inputs that contributed sufficiently to compensate for the lack of perfect reliability in the historical measures. The regression statistics, including the weights, appear in Table 5–1.

In Table 5–1, CASES represents the number of data points, RMUL the multiple correlation, RADJ the same correlation adjusted for shrinkage (given the large number of data points, shrinkage is negligible), F the standard variance ratio statistic for the model, PROB the statistical significance (the probability that a multiple correlation as high as that obtained could be due to chance), RIDGE the ridge coefficient (nonzero only for ridge regression), YMEAN the mean of the dependent variable, YSDEV the standard deviation of the dependent variable, XMEAN the mean of the independent variable, XSDEV the standard deviation of the independent variable, WEIGHT the regression weight, STUD-T the Student's *t*-statistic used to assess the significance of a regression weight, and RXY the Pearson Product-Moment Correlation between the independent variable and the dependent variable.

## Discussion

When compared to the models explored in Studies 1 and 2, the model constructed in this study yields predictions of mean (expected) future volatility that achieve greater accuracy over a much wider range. Future volatility levels of up to a whopping 200% can

**TABLE 5–1**

**Regression Statistics for a Multivariate Model Designed to Predict Near-Term Future Volatility from Average Range Historical Measures and Seasonal Cycle Harmonics**

| | | | | | |
|---|---|---|---|---|---|
| CASES | 3,213,002 | F | 56,369.98 | YMEAN | 0.507 |
| RMUL | 0.4022 | PROB | 0 | YSDEV | 0.557 |
| RADJ | 0.4022 | RIDGE | 0 | | |

| VARNO | Description | XMEAN | XSDEV | WEIGHT | STUD-T | PROB | RXY |
|---|---|---|---|---|---|---|---|
| 0 | INTERCEPT | 1 | 0 | 0.0774 | 67.33 | 0 | 0 |
| 1 | X1 | 0.522 | 0.287 | 0.6673 | 139.03 | 0.0000 | 0.391 |
| 2 | X2 | 0.521 | 0.274 | 0.1806 | 34.74 | 0.0000 | 0.354 |
| 3 | X1 * X2 | 0.337 | 0.368 | –0.1014 | –10.10 | 0.0000 | 0.372 |
| 4 | X1 * X1 | 0.355 | 0.434 | 0.0224 | 4.61 | 0.0000 | 0.365 |
| 5 | X2 * X2 | 0.347 | 0.392 | 0.0430 | 7.25 | 0.0000 | 0.330 |
| 6 | X1 * SIN 4T | 0.015 | 0.423 | –0.0150 | –21.85 | 0.0000 | 0.021 |
| 7 | X1 * COS 4T | –0.009 | 0.419 | 0.0791 | 115.96 | 0.0000 | 0.047 |
| 8 | X1 * SIN 8T | –0.012 | 0.421 | 0.0508 | 74.94 | 0.0000 | 0.036 |
| 9 | X1 * COS 8T | –0.005 | 0.421 | 0.0263 | 38.77 | 0.0000 | 0.016 |
| 10 | X1 * SIN 12T | 0.005 | 0.426 | 0.0121 | 18.09 | 0.0000 | 0.013 |
| 11 | X1 * COS 12T | –0.013 | 0.416 | 0.0151 | 22.01 | 0.0000 | 0.003 |

X1 = Short-term average range historical volatility, 30 bar.
X2 = Long-term average range historical volatility, 70 bar.
T = Time of year in radians.

now be anticipated, levels that were beyond the reach of earlier models. Furthermore, almost perfect accuracy is achieved when predicting future expectation for the range of volatility (between 20% and 140%) most often encountered by traders. To combine cycle harmonics with the most reliable measures of historical volatility in one multivariate model was a smart move. But what about implied volatility? Could it, too, contribute to a model designed to predict future volatility? So far, only historical volatility and seasonal cycles have been examined. Implied volatility has been ignored. Study 5 corrects this ignorance: it takes a look at implied volatility and its ability (or lack thereof) to forecast future price movement.

## STUDY 5: IMPLIED VOLATILITY

Because implied volatility is often used to price options, it is examined here just as historical measures were examined earlier. The current study investigates how implied volatility compares with historical volatility when used as a predictor of future market behavior. It also attempts to determine whether implied volatility is subject to the capping effects, other nonlinear response patterns, and regression to the mean that were found with the historical measures. The approach taken in the present study is the same as in Study 1, except that implied, rather than historical, volatility serves as the independent variable.

### Method

First, a stock was chosen. Then, a valid reference bar was selected. A reference bar was considered valid if (1) data was available for $m$ consecutive bars ending at the reference bar, (2) the stock was active over all $m$ bars, (3) the stock's raw (not split-corrected) price was greater than \$2 for every one of those bars, (4) the average of the put and call implied volatility figures at the reference bar was greater than 0.05 and less than 2.10, (5) the absolute difference between the put and call implied volatility figures was less than 0.20, (6) data were available for at least $n$ consecutive bars following the reference, and (7) the last bar fell on expiration Friday, if trading occurred on that day, or on the most recent day during which trading did take place. These conditions guaranteed that valid data were available for further analysis. Condition (5) was necessary to prevent the use of implied volatility figures based on stock prices that had changed significantly in the 15 minutes between closing of the options exchanges and closing of the stock exchanges.

Once a valid reference bar was located, the independent and dependent variables were determined. The first independent variable was the implied volatility. It was computed by averaging the separate put and call figures taken from the database, which, in turn, were calculated as described earlier. The second independent variable was just the square of the first; it was intended to represent the second-order term in the regression.

By including a second-order term, quadratic relationships could be fitted. Future volatility was computed over the *n* bars that followed the reference and served as the dependent variable. All three variables were written to a scratch file.

The next reference was then selected and, once all reference bars for the current stock had been dealt with, the next stock was drawn from the database. The process continued until every reference bar and stock had figured in the calculations. A scratch file that contained an array with three columns (variables) and 18,633 rows (data points or reference bars) was the result.

By using data from the scratch file, two analyses were executed. The first analysis involved sorting the implied volatility figures into 20 bins or levels and calculating for each level of implied volatility, the mean future volatility. Implied volatility was sorted into only 20 bins—rather than the 50 bins into which historical volatility was classified in Study 1—because there were far fewer valid data points. The second analysis was a multiple regression with a linear and second-order term, i.e., a second-order polynomial regression.

The entire procedure delineated above was carried out for each of the three types of future volatility: standard, average range, and straddle-implied. In every analysis, *m* was 30 and *n* was 10.

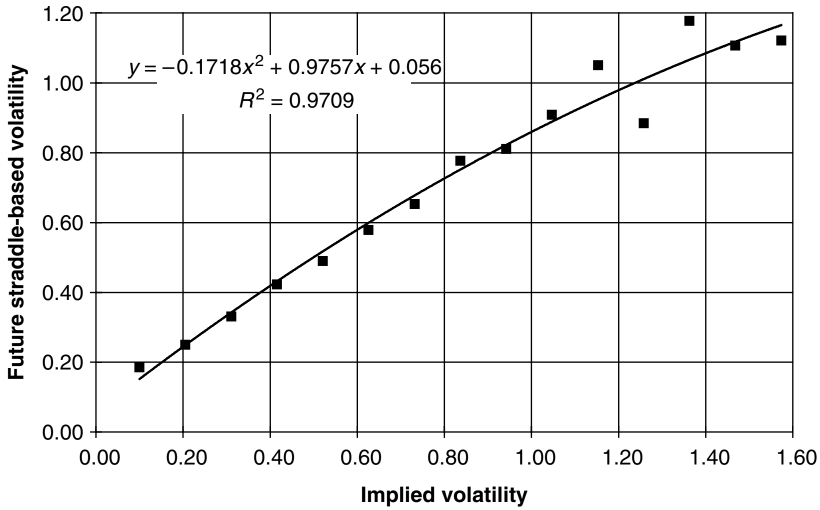
### Results

Each data point in Figure 5–7 shows the expected future straddle-based volatility associated with a given bin or level of implied volatility. A second-order polynomial smoothing is represented by the solid curve drawn through the data points. The smoothing equation, determined from the 20 data points shown in the chart, appears in the upper left corner. Charts for the analyses involving the standard and average range measures of future volatility have been omitted.

Except for a greater dispersion of data points about the polynomial curve and a smaller range of volatility on the *x*-axis, both of which are consequences of the smaller number of data points, the results from Study 5 shown in Figure 5–7 appear quite cognate to those that appear in Figure 5–1 from Study 1.

**FIGURE 5–7**

Implied Volatility as a Predictor of Future Volatility




In both studies, there is regression to the mean and the relationship between the independent and dependent variable flattens as volatility rises—the so-called "capping effect." It is no surprise that the smoothing polynomials that characterize the relationship are quite similar across both studies, with a modest negative coefficient for the second-order term and a positive coefficient that is slightly greater than unity for the linear term.

The regression performed on the 18,633 raw data points yielded a predicted, straddle-based volatility of $\hat{y} = -0.199x^2 + 1.034x + 0.021$, where $x$ is the implied volatility. A multiple correlation of 0.370 for predicting straddle-based volatility from implied volatility was achieved by the model. This may be compared to the analogous correlation in Study 1 of 0.373, when the standard historical measure was used to predict future straddle-based volatility, and of 0.395, when the average range was employed. When corrected for attenuation due to lack of reliability in the straddle-based dependent variable, the multiple correlations were 0.74, 0.75, and 0.79, respectively. Regression equations computed to predict standard and average range (rather than straddle-based) volatility had very similar weight coefficients. The raw multiple correlations were much greater when predicting the more reliable measures of future volatility but, after correction for attenuation, the numbers were in the same ballpark as those just presented.

### Discussion

These results demonstrate that implied volatility behaves very much like its historical counterpart when it comes to the estimation of future market action. For example, both historical and implied measures evidence significant regression to the mean, as well as the capping effect. This means that both historical and implied volatility require adjustments before they are entered into an option pricing model.

Given these results, it is easy to see that there is no great edge in the use of implied volatility to predict future volatility or to price options; in fact, there are some disadvantages, like the inability to determine an option's nonrelative worth. This is not to suggest that implied volatility has no virtues. Since implied volatility performs almost as well as 30-bar historical volatility, it must be a reasonably reliable measure. Implied volatility has another virtue: it is current. Unlike historical volatility, which must be computed from a sample of bars extending well into the past, implied volatility can be determined from a single, recent bar (e.g., the reference bar). Of course, measures like the average range require fewer bars to achieve good reliability, and so can also yield more current volatility estimates than can the standard historical measure. The intriguing possibility, however, is that implied volatility might contribute some independent information to a model that includes historical measures, the result being better overall performance. This possibility is examined in the next study.

## STUDY 6: HISTORICAL AND IMPLIED VOLATILITY COMBINED

Can implied volatility add predictive information to a model that already includes an historical measure? How strong is the relationship between implied and historical volatility? Are they measuring essentially the same variable? Is implied volatility more influenced by volatility in the future than in the past? These are some of the questions addressed in the current study.

Study 6 employs several analytical methods, ranging from the familiar correlation and multiple regression techniques to the more obscure (at least in this subject area) path analysis. The extent to which each kind of volatility contributes to a predictive model may be ascertained from the regression weights and associated $t$-statistics computed in the course of a multiple regression analysis. Correlation coefficients, perhaps corrected for attenuation, can answer questions regarding the strength of various relationships. When subjected to analysis by the method of path coefficients (Wright, 1934), correlations may also shed light on causal or directional influence.

## Method

As in previous studies, a stock was drawn from the database and a valid reference bar was selected. The criteria for the selection of a valid reference bar were identical to those used in Study 5. Once a valid reference bar was located, the independent and dependent variables were calculated as follows:

`Y` = Future $n$-bar volatility
`X1` = Implied volatility (average of put and call measures)
`X2` = Average range $m$-bar historical volatility
`X3` = `X1 * X2` Interaction (cross-product) term
`X4` = `X1 * X1` Second-order term for implied volatility
`X5` = `X2 * X2` Second-order term for historical volatility

These variables were then written to a scratch file and the next reference bar was selected. When all reference bars for the current stock had been processed, the next stock was drawn from the database. The above steps were repeated until all valid reference bars had been processed for all stocks. A scratch file containing an array with six columns (one for each variable) and 18,633 rows (one for each valid data point or reference bar) was produced.

Using data from the scratch file, three procedures were performed; the first was multiple regression. Regression weights for the independent variables `X1` through `X5` were determined along with all associated statistics. The second procedure involved sorting the predicted values of future volatility (based on the regression equation derived in the previous step) into 20 bins. For predictions falling into each bin, the mean future volatility was computed. The third procedure was calculation of a correlation matrix for all variables listed above, both independent and dependent. Data generated by each procedure was written to a standard text file for import into an Excel spreadsheet, where it could be further manipulated and prepared for presentation. The entire series of steps described in this section were repeated for each of the three measures of future volatility: the standard measure, the average range measure, and the straddle-based measure.

## Results

Table 5–2 presents regression weights and related statistics for two models designed to predict future volatility (`Y`) from variables based on the historical and implied measures (`X1` through `X5`). One of the two models is for prediction of straddle-based volatility; the other is for prediction of average range volatility. Similar statistics for the standard measure are not presented. The captions used in Table 5–2 are identical to those which appear in Table 5–1, discussed earlier.

***Regression Results*** As readily seen from the regression weights and *t*-statistics, both implied and historical volatility make independent and statistically significant contributions to the predicted future volatility level. Historical volatility has a greater impact on predictions when the dependent variable is the average range, while implied volatility contributes more to the model when the dependent variable is the standard measure. Because the individual second-order terms possess negative weights in both models, as either implied or historical volatility rises, its relationship with future volatility levels off. This is consistent with earlier observations involving either implied or historical volatility on its own. However, there is a synergy between historical and implied volatility. In both of the regressions presented in Table 5–2, the cross product or interaction term has a substantial positive weight. The synergistic interaction that occurs when both implied and historical volatility are high tends

**TABLE 5–2**

**Multiple Regression Statistics for Models Designed to Predict Future Volatility from Historical and Implied Measures**

Regression for Straddle-Based Future Volatility

| | | | | | |
|---|---|---|---|---|---|
| CASES | 18,633 | F | 593.36 | YMEAN | 0.517 |
| RMUL | 0.3707 | PROB | 0 | YSDEV | 0.493 |
| RADJ | 0.3704 | RIDGE | 0 | Y = | SRF-10 |

| VARNO | Description | XMEAN | XSDEV | WEIGHT | STUD-T | PROB | RXY |
|---|---|---|---|---|---|---|---|
| 0 | INTERCEPT | 1.000 | 0.000 | 0.0160 | 0.89 | 0.3733 | 0.000 |
| 1 | X1 = IMPLIED | 0.548 | 0.237 | 0.7610 | 9.35 | 0.0000 | 0.367 |
| 2 | X2 = HISTORICAL | 0.513 | 0.240 | 0.3093 | 3.87 | 0.0001 | 0.329 |
| 3 | X3 = X1 * X2 | 0.330 | 0.301 | 0.4713 | 2.50 | 0.0125 | 0.341 |
| 4 | X4 = X1 * X1 | 0.356 | 0.333 | –0.2823 | –2.86 | 0.0043 | 0.345 |
| 5 | X5 = X2 * X2 | 0.320 | 0.320 | –0.4024 | –3.85 | 0.0001 | 0.301 |

Regression for Average Range Future Volatility

| | | | | | |
|---|---|---|---|---|---|
| CASES | 18,633 | F | 7,599.27 | YMEAN | 0.522 |
| RMUL | 0.8192 | PROB | 0 | YSDEV | 0.262 |
| RADJ | 0.8191 | RIDGE | 0 | Y = | ARF-10 |

| VARNO | Description | XMEAN | XSDEV | WEIGHT | STUD-T | PROB | RXY |
|---|---|---|---|---|---|---|---|
| 0 | INTERCEPT | 1.000 | 0.000 | 0.0104 | 1.76 | 0.0791 | 0.000 |
| 1 | X1 = IMPLIED | 0.548 | 0.237 | 0.2981 | 11.16 | 0.0000 | 0.775 |
| 2 | X2 = HISTORICAL | 0.513 | 0.240 | 0.7198 | 27.40 | 0.0000 | 0.799 |
| 3 | X3 = X1 * X2 | 0.330 | 0.301 | 0.3209 | 5.18 | 0.0000 | 0.788 |
| 4 | X4 = X1 * X1 | 0.356 | 0.333 | –0.0994 | –3.06 | 0.0022 | 0.733 |
| 5 | X5 = X2 * X2 | 0.320 | 0.320 | –0.2840 | –8.28 | 0.0000 | 0.756 |

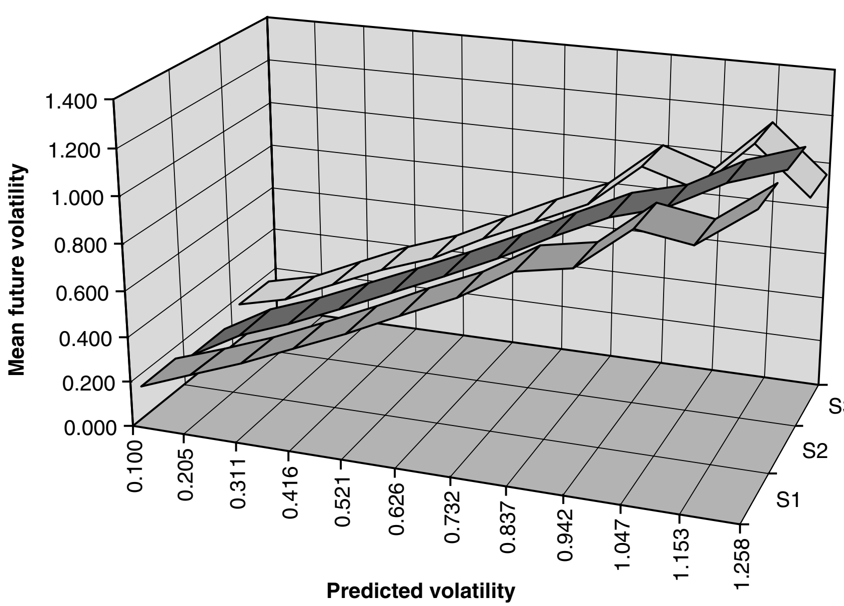
to cancel the downward pull of the individual second-order terms in the model.

Multiple correlations of 0.819 and 0.371 are achieved when predicting the average range and standard measures, respectively. The correlations are 0.851 and 0.741 when corrected for attenuation caused by lack of reliability in the future or dependent measure. These corrected correlations may be compared to similar ones of 0.840 and 0.789, which were computed with data taken from Study 1, and in which only historical volatility was employed as a predictor. The correlations from the current study are not larger than those obtained in the earlier study (where historical volatility was the only predictor); this suggests that the addition of implied volatility contributes little or nothing to a model's predictive ability, which runs counter to the findings from the multiple regression. Perhaps the high correlation that exists between implied and historical volatility—a correlation that suggests one can serve as a surrogate for the other—explains the apparent absence of benefit from the addition of the implied measure. When examining the numbers, also keep in mind that the earlier analysis was based on a larger sample of data that spanned additional years, embraced a greater variety of market conditions, and had a somewhat wider range of volatility.

The association between model-generated predictions and future volatility is shown in Figure 5–8. This figure was generated from the binned data. The $x$-axis, which runs from left to right represents the level of predicted volatility. Ticks on this axis

**FIGURE 5–8**

Mean Future Volatility for Three Measures versus Predicted Future Volatility



correspond to the centers of the bins into which the predictions were sorted. The $y$-axis, which runs from front to back, reflects the type of future volatility that is predicted. In order, from front to back, are the standard, average range, and straddle-based measures. Mean future volatility is plotted on the $z$-axis, which runs from bottom to top.

Series of data points are represented in the chart as ribbons. The middle ribbon follows the straightest path. It depicts a near-perfect relationship between predicted value and future expectation when future volatility is determined from the average range. The ribbon furthest back illustrates the greater dispersion of mean future volatility when predicted levels are high and the straddle-based measure is used as the dependent variable. Such dispersion is explained by the lower reliability of the straddle-based measure and by the smaller number of predictions falling into the higher bins. The ribbon closest to the front of the chart represents the relationship between predicted and mean future volatility for the standard future measure. It ranks somewhere between the other two ribbons in terms of linearity and absence of dispersion.

***Correlational Analysis*** A strong correlation of 0.857 (0.875 when corrected for attenuation) exists between historical and implied volatility. The correlation of historical volatility is 0.799 (0.842) with the future average range and 0.329 (0.667) with the straddle-based measure. For implied volatility, the correlation is 0.775 (0.811) with the average range and 0.367 (0.739) with the straddle-based measure.

Raw correlations with the future average range are by far higher than raw correlations with the straddle-based measure for both historical and implied volatility. When corrected for attenuation, correlations with the average range still trump those with the straddle-based measure, but to a much lesser extent. The implication is that the reliability estimate for the straddle-based measure is a little high and that, consequently, the correlations are undercorrected.

***Path Analysis*** A path analysis was performed to gain deeper insight into the causal structure implied by the observed correlations. Consider a theoretical model in which historical volatility influences both implied and future volatility, and in which implied volatility influences only future volatility. Given the causal assumptions of the model, figures that represent the intensities of the influences may be determined such that the observed correlations among the variables are fully accounted for or, barring that, maximally explained in a least-squares sense. Such figures are known as *path coefficients*. With future volatility measured by the more reliable average range and correlations corrected for attenuation, the path coefficients for the influence of historical on implied, historical on future, and implied on future volatility are 0.875, 0.565, and 0.317, respectively. These path coefficients demonstrate that implied volatility is highly controlled by historical volatility and that both independently affect future volatility, with historical volatility having a somewhat greater influence.

An alternative causal model assumes that future volatility is affected by historical volatility and that both historical and future volatility influence implied volatility. How can this be? According to McMillan (1996), activity in equity options often reflects insider knowledge regarding corporate plans, impending bankruptcy filings, takeover bids, and yet-to-be-released earnings reports. It takes no stretch of the imagination to see how implied volatility may, therefore, be influenced by impending events, known only to insiders, that will generate future volatility in the underlying stock. For this causal model, the path coefficients are 0.842 for the effect of historical on future volatility, 0.662 for historical on implied, and 0.254 for future on implied.

## Discussion

Can implied and historical volatility work together to improve the prediction of future volatility? According to the regression calculated in this study the answer is “Yes”—regression coefficients for both variables were significantly different from zero. Path analysis also clearly demonstrates that, even though implied volatility is itself highly determined by historical volatility, each variable has a direct influence on future volatility. Even when the causal model was flipped around in the second path analysis (in which future volatility was hypothesized to be an influence on implied volatility), a positive path coefficient appeared; however, the path going the other way (from implied to future volatility) in the first causal analysis had a much larger coefficient of determination.

The fact that the regression models developed in the current study did not produce higher multiple correlations than those observed in Study 1 seems to indicate that no benefit derived from the addition of implied volatility. As discussed earlier, this finding was most likely an artifact of differences in data sample and other related factors. It can be concluded, however, that the contribution of implied volatility to a model that already employs historical volatility is relatively modest. No doubt, this is a consequence of the fact that implied volatility is mostly determined by the actions of traders responding to pricing models that employ historical volatility measures.

## STUDY 7: RELIABILITY OF IMPLIED VOLATILITY

The goal in this study is to examine the reliability of implied volatility just as the reliability of the other measures were investigated in Study 3.

### Method

The method in Study 7 closely follows those found in the earlier studies. A stock was drawn from the first database and a reference bar was selected. The reference bar was checked for validity exactly as in the previous two studies. Given a valid reference bar, put and call implied volatilities were extracted from the second database for four consecutive bars, ending at the reference. The put and call figures were averaged for each bar, yielding four numbers that were saved to a scratch file. At this point, the next reference bar was selected and, when all reference bars had been processed, the next stock was drawn. The sequence continued until all stocks and reference bars had been analyzed. The result was a scratch file containing four columns, each corresponding to an implied volatility, and 18,633 rows, each corresponding to a valid reference bar.

A four-by-four correlation matrix was then computed for the data in the scratch file. The figures in the correlation matrix were used to compute estimates of reliability, as well as to examine the stability of implied volatility over short periods of time. For curiosity's sake, the first centroid factor was extracted from the correlations.

## Results

The correlation coefficients computed for implied volatility appear together with the centroid factor loadings and other data, in Table 5–3.

As might be expected, the correlation matrix has a so-called *Toeplitz* form: all correlations that fall just below the diagonal have one value, those that fall just below the preceding correlations have another value, and so on. Obviously, the correlations that fall immediately below the diagonal all represent the relationship between implied volatility on one bar with implied volatility on an immediately adjacent bar. All correlations that fall

**TABLE 5–3**

Correlations, Factors, and Other Data for Implied Volatility Figures

| | *Reliabilities Appear in Diagonal Implied Volatility Correlations* | | | | *Factor Analysis* | | |
|---|---|---|---|---|---|---|---|
| **VARNO** | **0** | **1** | **2** | **3** | **Loading** | **H**2** | **U**2** |
| 0 | 0.986 | 0.986 | 0.976 | 0.966 | 0.988 | 0.976 | 0.024 |
| 1 | 0.986 | 0.986 | 0.986 | 0.977 | 0.993 | 0.986 | 0.014 |
| 2 | 0.976 | 0.986 | 0.986 | 0.986 | 0.993 | 0.986 | 0.014 |
| 3 | 0.966 | 0.977 | 0.986 | 0.986 | 0.988 | 0.976 | 0.024 |
| MEANS | 0.535 | 0.536 | 0.532 | 0.531 | TOTVAR | 15.698 | |
| SDEVS | 0.223 | 0.222 | 0.219 | 0.218 | | | |
| NCAS | 15,692 | | | | | | |
| NVAR | 4 | | | | | | |

another step away from the diagonal represent the relationship between implied volatility on one bar with that measured two bars away in either direction. For practical purposes, the correlations just off the diagonal may be treated as reliability coefficients, suggesting a reliability of 0.986 for implied volatility. Moving away from the diagonal, the correlations become smaller as a result of changes in implied volatility over time, and hence more like the stability coefficient computed in Study 3; in Study 3, however, the constancy under examination was for a much longer period than just two or three days (roughly 28 days separated the center of the 30-bar historical volatility from the center of the 10-bar future measure).

The factor loadings also provide estimates of reliability. In this case, the squares of the loadings (found in Table 5–3 in the column labeled H**2) for variables 2 and 3 may be taken as approximations to reliability. These approximations are extremely close in value to the off-diagonal elements of the correlation matrix.

## Discussion

Implied volatility is extremely reliable in that repeated measurements produce nearly identical results. And, this high reliability is achieved with only a single measurement instance or test item; this is in stark contrast to standard historical volatility, which requires an abundant sample of bars to achieve good reliability. The ability to obtain a highly reliable measurement from a single bar that lies as close as possible in time to the period over which future volatility must be estimated can be a great advantage. Implied volatility thus has the virtue of being current. If the reliability and contemporaneousness were matched by predictive worth, implied volatility would be the ultimate estimator of future market action. As demonstrated by both regression and path analysis, implied volatility has definite value when it comes to predicting future volatility; however, implied volatility is mostly determined by historical volatility, which can provide equally good, if not better, predictions. Implied volatility is, therefore, an important measure—possibly worthy of inclusion as an independent variable in a model designed to forecast volatility—but not an awe-inspiring one that is superior to most others.

It should be noted that the reliability coefficients computed in this study were used to correct the correlations that involved implied volatility in Study 6 for attenuation.

## SUMMARY

Potentially useful independent variables were examined and various models designed to estimate future volatility were tested. Much was learned as a result of the effort. Some of the findings are summarized here.

Consider measurement reliability and predictive worth. All else being equal, the more reliable an independent variable, the more it can contribute to a prediction or estimation model. For a given number of bars, it is unquestionable that implied volatility is by far the most reliable of the independent variables tested. Next comes the average range, which achieves respectable reliability with a fairly small measurement sample. Less reliable is the standard measure. At the bottom of the list is straddle-implied volatility. When used as an independent variable, the average range has the best overall ability to forecast volatility; it represents a good combination of high reliability and high validity. Standard volatility performs acceptably, as does implied volatility. On its own, however, implied volatility can only be used for relative and not absolute valuation when pricing options. Straddle-based volatility, though perhaps the most construct-valid, was far too unreliable for use as an independent or predictor variable.

Regardless of the kind of volatility examined, a so-called "capping effect" was observed. As the level of the independent variable (whether historical or implied volatility) increased, the dependent variable (future volatility) at first rose rapidly, but then reached a limiting plateau. In all cases, the relationship was fairly well described by an elementary quadratic equation with a positive linear coefficient that was slightly above unity and a negative second-order coefficient about one-fifth the size. The capping effect was least with the average range historical measure. Even though the capping effect may be an artifact of the mathematical nature of volatility, it must be accounted for in any model intended to estimate future market action.

In the course of the investigations, it was found that better estimation of future volatility could be achieved quite readily by the use of more than one independent variable. For example, a model that employed historical volatility from two periods, as well as the first three harmonics of the quarterly earnings cycle, provided distinctly better estimations of future volatility than any of the simpler models. In addition, this model could successfully anticipate much higher levels of future volatility, shattering the cap. Another model that demonstrated improved prediction and reduced capping was the one based on a combination of both historical and implied measures. Implied volatility was not incorporated into the model that employed cycles and two historical volatility measurements; examining the effect of bringing implied volatility into this more complex model is left as an exercise for the reader.

Regardless of the number of variables involved in a model, it is clearly necessary to take into account reversion to the mean and lack of linearity when attempts are made to forecast volatility for the purpose of pricing options. The direct use of unadorned historical or implied measures as estimates of future market behavior will lead to biased predictions, especially for high or low volatility levels. Hence, the standard practice of using raw historical or implied volatility as an input to Black-Scholes is ill-advised. With minor adjustments, more accurate future volatility estimates can be obtained. Even the simple procedure of consulting a graph (such as one of those sprinkled throughout this chapter) to correct historical volatility for nonlinear response and regression to the mean before plugging the numbers into Black-Scholes or some other popular model can dramatically improve the accuracy of the theoretical option premiums thus obtained.

Chapter 6 moves away from the subject of volatility and examines the use of conditional distributions to price options. At first it might appear more appropriate for the subject of conditional distributions to follow Chapter 3, another chapter concerned with distributions and option prices; however, estimates of volatility are required for much of what follows and therefore, some serious coverage of volatility seemed to be a warranted prelude to the study of conditional distributions.

## SUGGESTED READING

Good coverage of measurement reliability, construct validity, and regression to the mean can be found in *Foundations of the Theory of Prediction* (Rozeboom, 1966). Consult *Psychometric Theory* (Nunnally, 1978) for a general overview of classical test theory and a less technical presentation on reliability and validity. Another basic source on the subject is *Essentials of Psychological Testing* (Cronbach, 1970). These books are classics in their field. The *Handbook of Psychological Testing* (Kline, 2000) and *Introduction to Classical and Modern Test Theory* (Crocker and Algina, 1986) are newer works on the same subject.

Linear regression, in both its univariate and multivariate forms, is well covered in *Introduction to Linear Regression Analysis* (Montgomery and Peck, 1982) and in *Classical and Modern Regression with Applications* (Myers, 1986). These books also tackle the statistical phenomenon of regression to the mean.

For the mathematically literate reader wishing to learn more about factor analysis, *Modern Factor Analysis* (Harmon, 1976) is the accepted bible of the field. The clearest presentation on causal inference from correlational data is the original paper on the subject, "The Method of Path Coefficients" (Wright, 1934). *Causal Analysis* (Heise, 1975) and *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations, and Causal Inference* (Shipley, 2002) may also be recommended. For the basics on modeling time series with GARCH, a reasonable exposition can be found in "Generalized Autoregressive Conditional Heteroskedasticity" (Bollerslev, 1986).

CHAPTER 6

# Pricing Options with Conditional Distributions

Theoretical option premiums are easily calculated from distributions of returns. This fact was clearly demonstrated in Chapter 3, where it was shown how price distributions determine option value and where some of the specific theoretical distributions commonly assumed to govern market behavior were examined. Furthermore, it was established in that chapter that popular models like Black-Scholes essentially calculate future expectation or theoretical fair value under the assumption that price movements or returns have a log-normal distribution.

An examination of moments in Chapter 4, however, revealed that the distribution of real-market returns is not log-normal; rather, the empirical distribution of returns is leptokurtic and negatively skewed. Consequently, theoretical option premiums calculated with models like Black-Scholes must be more or less inaccurate. Given the observed distribution of returns, out-of-the-money options can be expected to have greater value than popular models suggest, for example, and out-of-the-money puts are likely to be worth more than the corresponding calls. Fortunately, a log-normal distribution of returns is not a necessary presupposition when pricing options. Distributions that better reflect actual market behavior may be employed. The use of more fitting distributions is likely to result in more trustworthy assessments of option value.



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

The focus in the present chapter is on pricing options with empirically derived conditional distributions. What is a *conditional distribution*? It is a distribution that depends upon one or more conditioning factors. By way of illustration, it is not very difficult to establish the distribution of returns associated with a given combination of initial stock price, volatility, and skew. The resultant distribution may be considered conditional, since it is clearly dependent on the particular levels specified for stock price, volatility, and skew—the conditioning variables. The use of a conditional distribution methodology allows many different factors to be taken into account in an empirical pricing model.

## DEGREES OF FREEDOM

When developing an option pricing model that relies upon empirically defined conditional distributions, it is absolutely critical to conserve degrees of freedom. The reader may recall that a greater dispersion was observed in the expected future volatility when data were binned for two historical volatility measures, rather than for one. The reason was that the two-dimensional solution involved a vastly greater number of bins than did the one-dimensional solution and, consequently, fewer data points fell in any single bin in the two-dimensional solution than in the one-dimensional one. Analysis of empirical distributions involves a binning procedure analogous to that used when studying volatility, as well as a comparable potential for some of the bins to contain insufficient data for statistically stable results. Several of the studies that appear below use a technique first illustrated in Chapter 3, where a bin-based analysis was performed to demonstrate how an option may be priced using a distribution; however, that analysis was conducted as part of a Monte Carlo experiment where as many data points as desired could readily be generated. The luxury of a virtually unlimited supply of data points does not exist when dealing with live markets. Hence, the number of bins or classes created by the conditioning variables must be kept as small as possible so as to conserve degrees of freedom. Fewer bins means more data points in each bin and, therefore, more stable and statistically meaningful results.

The fact is that any kind of theoretical simplification that can reduce the number of conditioning variables without necessarily undermining the model will make the results more reliable and useful. For example, if all stocks can somehow be adjusted so that they start out at a nominal price of $100, it may be possible to eliminate initial stock price as a conditioning variable, thus conserving degrees of freedom by reducing the total number of bins or categories involved in the analysis.

### Use of Rescaling to Conserve Degrees of Freedom

Is it reasonable to adjust or rescale the initial price of a stock to a nominal $100? If certain assumptions regarding the nature of stock price movements are met, the answer is a credible "yes." The assumptions are quite reasonable for a broad range of stock prices. The principal assumption is that stock prices manifest proportional movement, i.e., a $100 stock gaining $10 is, in essence, the same as a $10 stock rising $1 or a $5 stock ascending $0.50. This is an assumption made by almost every pricing model, either explicitly or implicitly, and one that seems quite acceptable. It is easy to verify that an elementary scaling procedure can bring any stock to the desired price. If the option's strike is likewise rescaled, an option price may be computed that will be equal to the similarly rescaled price of an option computed from the original (not rescaled) data. Since the ability to rescale in this manner is critical to a number of studies in this chapter, the process is illustrated here.

Imagine a stock that is trading at $20 and has a volatility of 70%. Further, suppose a call with a $22.50 strike and 10 days left to expiration is to be priced, and the risk-free interest rate is 10%. Black-Scholes asserts that the theoretical premium for this option is $0.207. Multiply the stock and strike prices by a scaling factor of 5. The rescaled prices are now $100 for the stock and $112.50 for the strike. If these numbers are entered into Black-Scholes, a theoretical option premium of $1.035 is obtained. This premium is exactly equal to the original option premium multiplied by 5, the scaling factor. In other words, the option premium computed for the nominal $100 stock may be divided by the scaling factor to obtain the premium appropriate for the $22.50 strike option on the $20 stock. Rescaling may thus be used to bring all stocks to the same price, thereby eliminating stock price as a necessary conditioning variable.

## GENERAL METHODOLOGY

To continue with the example, imagine that numerous stocks have been subjected to rescaling and are all now at the nominal $100 price. Further, imagine that the historical volatility is known for each stock. At this point, consider stocks with historical volatilities between 80% and 85%. Check their terminal prices 10 days later and sort them into a set of bins that span the range of possible price outcomes, just as was done for the Monte Carlo data in Chapter 3. Perhaps there are 7,500 instances of a price between $100 and $100.50, 7,200 instances of a price between $100.50 and $101, and so forth. The result is a histogram showing the frequency of observation plotted against the level of terminal stock price. This histogram, in fact, represents a *discrete frequency distribution* for a nominal $100 stock conditional upon a volatility that lies between 80% and 85%. Such a histogram or frequency distribution may be used to generate option premiums for a variety of strikes for both puts and calls that expire in 10 days. The entire procedure can be repeated for stocks with volatilities between 85% and 90% or between 90% and 95%. In other words, a family of histograms may be computed such that each histogram corresponds to a particular level of historical volatility. These histograms would contain all the information necessary to price both puts and calls with different strikes for stocks that have different levels of volatility. Given such a family of histograms, a low-order polynomial could be fitted in order to regularize the data for increased accuracy and to allow options at intermediate volatility levels to be priced.

In the real world, a price of $100 is rarely found when a stock and a reference bar are selected. However, any stock may be rescaled to a nominal $100 by multiplying its price by a suitable constant. The corresponding strike must also be multiplied by the same constant. An option premium can then be computed based on distributions established for a stock with a nominal initial price of $100 and that premium can then be divided by the scaling constant to obtain the desired option premium. Follow this logic and a rudimentary option pricing model ensues. To make the model realistic, prices may be adjusted for the rate of risk-free interest and for other factors.

The studies in this chapter use precisely this methodology to investigate option valuation with distributions that are conditional upon volatility and many other factors of interest. Later studies consider more complex questions, such as whether distributions of returns combine over time through a process of convolution.

### Data and Software

All the following studies exploit the same two binary databases used in the previous chapter when investigating volatility. The first database contains end-of-day stock quotes (consult Chapter 4 for details regarding its construction). The second database consists of implied volatility figures determined as described in the previous chapter. In addition, several studies require raw option prices taken from a third binary database that was assembled from data supplied by www.stricknet.com. The third database contains, for every stock and every bar, a complete option chain that consists of all options trading on that equity on that day. The fields in the options database include the *option symbol*, *underlying* (stock) *symbol*, *date*, *strike*, *bid*, *offer*, *volume*, *open interest*, *expiration year*, *expiration month*, and *option type* (put or call).

The software libraries and toolsets are the same as those employed in the previous chapters. Most of the calculations were performed in generic, ISO-standard C and C++, with some final analysis occasionally being conducted in Microsoft Excel. Excel was also used to prepare tables and charts for presentation.

## STUDY 1: RAW HISTORICAL VOLATILITY

This study is intended to demonstrate how options may be priced using conditional distributions. To keep things simple, the analysis is limited to one basic conditioning variable. Given the conditioning variable, a family of distributions is calculated. The distributions are then used to determine theoretical option premiums.

The conditioning variable in this study is standard historical volatility and options are assumed to have a holding period of 10 days. Interest rates are incorporated into the model. This is necessary to compensate for the stock price gain that would be expected in a risk-neutral world due to interest accruing over the 10-day holding period. Likewise, theoretical option premiums, computed on the basis of future price expectation, are discounted by the same interest rate. Options, like any underlying securities, are expected to grow at the current rate of risk-free interest; therefore, if the expected option price at some point in the future is known or estimated, that price needs to be discounted by the interest to obtain the current premium. Dividends are ignored, as are individual levels of growth, skew, and kurtosis. Because empirical distributions are employed, the growth, skew, and kurtosis that tend to characterize all stocks in the database are not ignored, but are implicitly incorporated into the model. Once the distributions are determined, they are used to price some options. Finally, the option prices generated from the distributions are compared to those computed with Black-Scholes.

## Method

The approach was to determine the distribution of terminal prices for a nominal $100 stock for each possible level of raw standard historical volatility, and to then compute (from the distributions) the theoretical option premiums for several strikes and option types.

The steps were as follows: First, a stock was selected from the 2,246 stocks in the database. The $m$-bar standard historical volatility was then calculated for each bar in the entire series of $n$ bars using a fast, vectorized subroutine. In this study, $m$ was 30 (a popular period over which to calculate historical volatility) and $n$ was 1,834. Next, a bar (the current or reference bar) was chosen. When choosing a bar, only bars from $m + 2$ through $n - 12$ were considered. The objective was to leave sufficient room at the beginning of the data series for computing the $m$-bar standard historical volatility and, at the end, for determining the terminal price 10 bars in the future. In selecting the current bar, only bars for which the stock was trading on the previous $m$ bars and subsequent 10 bars were considered. Also, the stock's raw (not corrected for splits) price on all $m$ bars preceding the current bar was required to be greater than $2. These are the same kinds of criteria used to select valid reference bars in studies discussed in earlier chapters. Once a valid reference bar was found, the raw historical volatility, the stock price at the current or reference bar, and the price 10 bars beyond the current bar (10 days in the future) were assembled.

The two assembled stock prices were used to calculate the terminal price of a nominal $100 stock. To calculate this, the future price of the stock was divided by the current stock price (at the reference bar) and the result was multiplied by 100. To wit, the stock's current price was rescaled to a nominal $100 and a future price consistent with the new scale was computed. For example, a stock with a current price of $20 and terminal price of $22 would be rescaled so as to have a current price of $100 and a future price of $110. Once determined, the future stock price was discounted for interest at the current rate over the holding period. The intent was to remove the implicit effect of interest on the distributions about to be calculated. Compensation for growth due to interest in a risk-neutral world was accomplished by multiplying the terminal stock price by `exp(-r*t)`, where `r` was the rate of interest at the reference bar, and `t` was the holding or look-ahead period measured in years (`t` = 10 bars/252 bars per year). The result was the adjusted terminal stock price.

From `p`, the adjusted terminal price, an array index was computed (using C language notation) as

```
ip = (int)(0.5 + (nlvlp-1)*(p-bmnp)/(bmxp-bmnp))
```

where `ip` represents the array column corresponding to a specified level of terminal stock price, `nlvlp` the number of price levels, `bmnp` the center of the first bin (corresponding to the lowest price level), and `bmxp` the center of the last bin (corresponding to the highest price level). In this study there were 100 price levels or bins, each corresponding to a column in the array, with the lowest level or bin centered at a price of $20 and the highest centered at a price of $250.

The raw historical volatility, `v`, was also used to calculate an index into an array. The formula for this array index (again in C notation) was

```
iv = (int)(0.5 + (nlvlv-1)*(v-bmnv)/(bmxv-bmnv))
```

where `iv` represents the array row associated with a specified level of historical volatility, `nlvlv` the number of volatility levels, `bmnv` the center of the first bin (corresponding to the lowest volatility level), and `bmxv` the center of the last bin (corresponding to the highest volatility level). There were 20 volatility levels, each corresponding to a row in the array, with the lowest level or bin centered at a volatility of 10% (0.10) and the highest centered at a volatility of 200% (2.00).

The contents of the array element addressed by the two index variables (`iv` and `ip`) was then incremented by one, allowing frequency statistics to be accumulated. At this point, the next valid reference bar was chosen and, when all valid bars for the currently selected stock had been processed, the next stock was selected. The sequence continued until all stocks and all valid reference bars had been included in the analysis. These calculations produced a set of frequency distributions residing in the rows of an array. Each row in the array corresponded to a level of volatility and each column was associated with a level of terminal stock price. Each array element thus comprised a frequency associated with a specified two-dimensional bin. There were 2,000 bins or elements in the array. The data in each row represented a frequency histogram for a given level of standard historical volatility.

By using the frequency histograms residing in the array rows, option prices for several strikes were calculated in a manner identical to that described in Chapter 3. A strike price and volatility level were selected for performing this calculation. The volatility level provided an index into the desired row in the array, the one that contained the appropriate distribution.

Working across the row, column by column, each frequency was multiplied by the easily calculated terminal (expiration) price of an option having the selected strike. This was done for calls, puts, and straddles. For each volatility, strike, and option type, the sum of the frequency-by-price products was divided by the sum of the frequencies. The resultant option premiums, based on the actual price distributions observed in the real market, were saved in another array.

Additionally, for purposes of comparison, Black-Scholes prices were calculated for the same set of volatility levels, strikes, and option types. As the reader may remember from Chapter 3, these are the prices that would be found if the histograms in the array rows defined appropriate log-normal distributions.

## Results

Table 6–1 presents the cumulative probability distributions for terminal stock prices for a nominal $100 stock. Cumulative probabilities derived from the empirical frequency distributions (EMP) are shown side-by-side with those derived from log-normal distributions (LN) for each of three levels of standard historical volatility. Parameters for the log-normal distributions were determined from the corresponding empirical distributions. The total frequencies (CASES) and the root-mean-square differences (RMSD) between the empirical and log-normal probabilities are also recorded in Table 6–1. The cumulative probability for a specified stock price level is just the sum of the frequencies for that and all lower price levels divided by the total frequency. For reasons of space and clarity, probabilities are presented in the table only for every third terminal price level.

Consistent differences between the empirical and log-normal distributions are easily discerned from the cumulative probabilities in Table 6–1. For terminal stock prices sufficiently below the nominal $100, the cumulative empirical probability is much greater than the cumulative log-normal probability; and, for prices sufficiently above $100, the empirical probability is much further below 100% than the corresponding log-normal

**TABLE 6–1**

Empirical and Log-Normal Cumulative Probabilities for 3 Levels of Standard Historical Volatility and 25 Levels of Terminal Stock Price

| Stock Price | *30% Volatility* | | *60% Volatility* | | *90% Volatility* | |
|---|---|---|---|---|---|---|
| | EMP | LN | EMP | LN | EMP | LN |
| 40.91 | 0.00006 | 0.00000 | 0.00047 | 0.00000 | 0.00089 | 0.00000 |
| 47.88 | 0.00018 | 0.00000 | 0.00090 | 0.00000 | 0.00203 | 0.00001 |
| 54.85 | 0.00042 | 0.00000 | 0.00197 | 0.00000 | 0.00479 | 0.00025 |
| 61.82 | 0.00094 | 0.00000 | 0.00415 | 0.00010 | 0.01042 | 0.00282 |
| 68.79 | 0.00212 | 0.00000 | 0.00993 | 0.00207 | 0.02336 | 0.01657 |
| 75.76 | 0.00567 | 0.00016 | 0.02520 | 0.01751 | 0.05370 | 0.06001 |
| 82.73 | 0.01859 | 0.00747 | 0.06787 | 0.07847 | 0.12047 | 0.15160 |
| 89.70 | 0.08194 | 0.08667 | 0.18242 | 0.21807 | 0.25036 | 0.29218 |
| 96.67 | 0.35892 | 0.35590 | 0.41085 | 0.42476 | 0.43959 | 0.45994 |
| 103.64 | 0.78605 | 0.71005 | 0.67432 | 0.63993 | 0.63526 | 0.62353 |
| 110.61 | 0.95245 | 0.92167 | 0.84625 | 0.80801 | 0.77961 | 0.75899 |
| 117.58 | 0.98892 | 0.98701 | 0.93238 | 0.91175 | 0.87473 | 0.85716 |
| 124.55 | 0.99699 | 0.99861 | 0.97000 | 0.96445 | 0.92970 | 0.92091 |
| 131.52 | 0.99906 | 0.99990 | 0.98591 | 0.98724 | 0.96036 | 0.95872 |
| 138.48 | 0.99964 | 0.99999 | 0.99317 | 0.99585 | 0.97811 | 0.97951 |
| 145.45 | 0.99980 | 1.00000 | 0.99639 | 0.99876 | 0.98750 | 0.99027 |
| 152.42 | 0.99987 | 1.00000 | 0.99805 | 0.99966 | 0.99244 | 0.99555 |
| 159.39 | 0.99992 | 1.00000 | 0.99888 | 0.99991 | 0.99542 | 0.99803 |
| 166.36 | 0.99994 | 1.00000 | 0.99931 | 0.99998 | 0.99715 | 0.99915 |
| 173.33 | 0.99996 | 1.00000 | 0.99957 | 0.99999 | 0.99811 | 0.99964 |
| 180.30 | 0.99997 | 1.00000 | 0.99969 | 1.00000 | 0.99865 | 0.99985 |
| 187.27 | 0.99998 | 1.00000 | 0.99983 | 1.00000 | 0.99914 | 0.99994 |
| 194.24 | 0.99999 | 1.00000 | 0.99988 | 1.00000 | 0.99942 | 0.99998 |
| 201.21 | 0.99999 | 1.00000 | 0.99990 | 1.00000 | 0.99957 | 0.99999 |
| CASES | 617,672 | | 363,755 | | 132,195 | |
| RMSD | 0.0296 | | 0.0246 | | 0.0231 | |

probability. This is true for all three volatility levels. As an example, the probability that a real stock with a 30% historical volatility will decline from an initial price of \$100 to a final price of \$75.76 or below is 0.567% (about 1 in 200), while the log-normal distribution suggests a diminutive 0.016% (about 1 in 6,000) probability of occurrence. At the other extreme, the actual probability that the same stock will rise to $131.52 or above is 0.094% (1 – 0.99906), a rise in price that the log-normal distribution indicates should occur only 0.01% (1 – 0.99990) of the time. In other words, extreme price changes—either up or down—are far more common with real stocks than would be anticipated from a log-normal random walk.

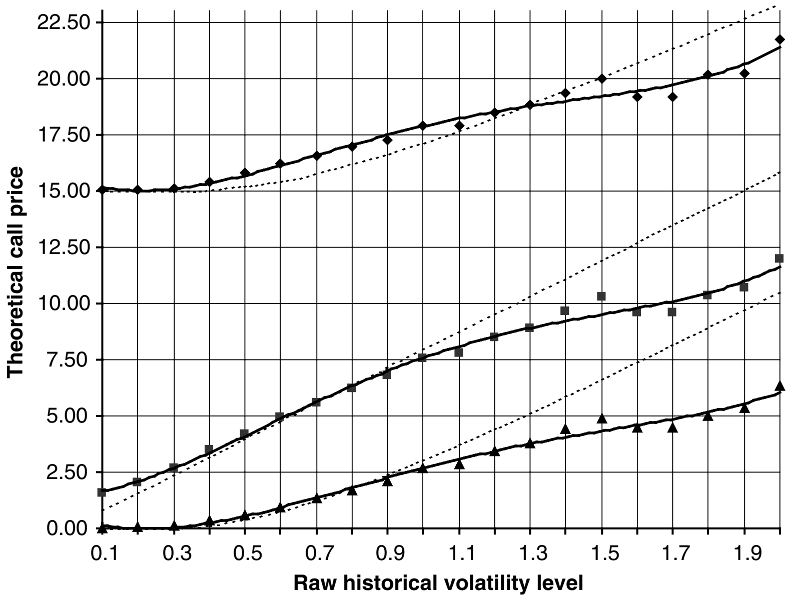
The situation is reversed for terminal stock prices that are only moderately above or below the $100 level. For prices moderately below the $100 level, the cumulative probabilities given by the log-normal distributions are greater than those obtained from the empirical ones. The cumulative probabilities obtained from the log-normal distributions are less than those derived from the empirical distributions when prices are moderately above the $100 level. This reflects a lower relative frequency of moderate price changes in real stocks when compared to what a model like Black-Scholes is predicated upon. Although not easily discerned from the figures in Table 6–1, the relative frequency of unchanged (or minimally changed) prices is high. Overall, real stocks exhibit a distribution of returns with longer tails and a sharper peak than should be observed if stock prices were tracing out log-normal random walks. In other words, real stocks display a leptokurtic distribution of terminal prices.

Figure 6–1 shows theoretical call premiums as a function of standard historical volatility. The solid curves represent premiums computed from the empirical or real-world distributions, while the dotted ones depict those taken from Black-Scholes (which assumes a log-normal distribution). Premiums computed with Black-Scholes and those derived from the empirical distributions are displayed for three strikes. The topmost curves are for a call struck at $85, and hence well in-the-money. Curves for an at-the-money call with a strike of $100 appear in the middle of the chart. The lowermost curves are for an out-of-the-money call that has a strike of $115.

Clearly evident in Figure 6–1 is the fact that Black-Scholes flagrantly overprices options when the level of historical volatility is high. Black-Scholes tends to underprice options when historical volatility is low and time value exists in the options being

**FIGURE 6–1**

Theoretical Call Prices from Black-Scholes and from Empirical Distributions for Three Different Strikes




appraised. It is, naturally, the time value that is poorly estimated by Black-Scholes in situations characterized by either high or low historical volatility. For each pair of curves, there is a crossover point where both the conditional distribution methodology (presumably accurate, at least for the data sample used in the analysis) and Black-Scholes yield the same theoretical fair value. The crossover point depends on the moneyness of the option: in-the-money calls cross at about 130% historical volatility; at-the-money and out-of-the-money calls cross at an historical volatility between 60% and 80%.

One force unmistakably at work in producing the deviations from Black-Scholes, seen in Figure 6–1, is regression to the mean. Comparatively high or low historical volatility levels are likely to be followed by more moderate levels of volatility in the future. This was convincingly demonstrated in Chapter 5, where future volatility was the focus of prediction efforts. When options are priced using conditional distributions of empirical origin, regression to the mean and other prediction issues are implicitly managed; the distributions conditional upon historical volatility (or any other variable) are of future, not of current, stock prices. Black-Scholes, on the other hand, does not implicitly correct for differences between historical and future volatility. Consequently, unless volatility is explicitly corrected, Black-Scholes assigns excessive value to options when historical volatility is high and insufficient value when it is low. The conditional distribution methodology, taking into account regression to the mean, is not subject to this kind of systematic pricing error.

A bullish bias in the data sample may explain why the curves for the in-the-money call cross at a higher level of historical volatility than do the curves for the other calls. Another viable explanation is the volatility payoff. Either way, deeply in-the-money calls are rich in Delta and gain in value directly with increasing stock price. The conditional distributions derived from the data have not been detrended and, therefore, reflect any tendency for in-the-money calls to show price strength. Any such price strength can push the crossover point to the right—i.e., towards higher levels of historical volatility.

## Discussion

This study established that options may be readily priced using conditional distributions derived from actual market data. It also confirmed some earlier findings, such as the fact that the distribution of stock price movements is leptokurtic and that Black-Scholes systematically misprices options due to mean reversion when raw historical volatility is used as a model input and is either relatively high or relatively low. Finally, the study demonstrated one virtue of a pricing model based on conditional distributions: the implicit management of prediction issues. There was no need to explicitly correct historical volatility for regression to the mean (or for nonlinearities like the capping effect discussed in Chapter 5) before generating theoretical option premiums. Other virtues, not demonstrated, include the fact that conditional distributions may be detrended and otherwise modified, that they statistically summarize actual market behavior (whether log-normal or not), and that they may be determined for a wide range of conditioning factors. Some of these virtues were demonstrated in a two-part article written for *Futures* (Katz and McCormick, 2001b,c) in which distributions that are conditional on factors other than just volatility—e.g., trading venue (NASD versus NYSE), overall market mien (bullish versus bearish), and skew (positive versus negative)—were adduced and used to price options.

## STUDY 2: REGRESSION-ESTIMATED VOLATILITY

The purpose of this study is to examine the effect of using an improved volatility estimate when computing the conditional distributions. Except for the fact that the raw historical volatility used in the previous study has been replaced with a dramatically better estimator of future market behavior in the current one, both studies are fundamentally the same. The current study uses a regression estimator that is based on two historical volatility measures and three seasonal harmonics. This estimator forecasts near-future stock price volatility fairly well and fully corrects for regression to the mean. As was the case in the previous study, skew and kurtosis assessments for individual stocks are not included among the conditioning variables; nevertheless, the skew and kurtosis commonly observed in stock price movements are implicitly accounted for in the model.

Why use a multivariate regression estimator for expected volatility rather than just include the relevant inputs directly as conditioning variables? If there were an infinite (or incredibly large) number of data points, it would make perfect sense to simply include the two historical volatility measures and the three cycle harmonics used in the multiple regression as conditioning variables. In the real world, however, the inclusion of all these variables in the model would consume so many degrees of freedom that none would be left from which to obtain stable distribution statistics or option prices. How many degrees of freedom would be used? Equivalently, how many bins would need to be filled with data to make things work? Suppose each of the twelve variables entered into the multiple regression (Chapter 5, Study 4) was partitioned into only 10 levels. In that case, a staggering 1,000,000,000,000 12-dimensional bins would need to be addressed. Because the conditional distribution methodology can directly model nonlinear relationships, things are not quite as bad as they might seem; the second-order and cross-product terms used in the regression could be dropped as inputs with no loss of accuracy, leaving just eight variables. Eight conditioning variables, each having 10 levels, would require a mere 100,000,000 bins or array cells—a definite improvement. Because multivariate regression is constrained to simple linear relationships between its inputs and output, far fewer degrees of freedom (in the present instance, only about 12) are consumed. The fewer the degrees of freedom consumed, the more meaningful the results.

### Method

The analysis took place in two computational blocks. In the first block, the conditional frequency distributions were accumulated in the rows of an array. Option premiums for a variety of option types, strikes, and volatility levels were calculated in the second block.

***First Computational Block*** Processing in the first block began with the selection of a stock and a reference bar. The reference bar was then checked for validity. A reference bar was considered valid if the following conditions held: (1) there were at least $m1 + m2$ bars preceding the reference; (2) the stock was active and had no raw price less than \$2 in the interval defined by those bars; and (3) there were at least $n$ bars following the reference. In this study, $m1$ was 30, $m2$ was 70, and $n$ was 10. Note that, in this instance, $n$ refers to the holding or look-ahead period, not to the total number of bars in each data series, which remains at 1,834, as in the previous study.

Once a valid reference bar was located, long-term and short-term historical volatility measures were calculated. Short-term volatility was determined from the $m1$ bars immediately preceding the reference. The $m2$ bars that came just before those employed for the short-term measure were used to compute long-term volatility. Volatility was measured by the average range method. Three harmonics of the quarterly cycle were also determined. The cycle harmonics, as a set of sines and cosines, were computed from the time of the year, expressed in radians. An initial estimate (or forecast) of expected volatility over the $n$ bars following the reference (the holding period) was then calculated as

$$\begin{aligned} x = {} & 0.077 + 0.667h_s + 0.181h_l - 0.101h_s h_l + 0.022h_s^2 \\ & + 0.045h_l^2 + h_s(-0.015 \sin 4\theta + 0.079 \cos 4\theta + 0.051 \sin 8\theta \\ & + 0.026 \cos 8\theta + 0.012 \sin 12\theta + 0.015 \cos 12\theta) \end{aligned} \tag{6.1}$$

where $h_s$ represents the short-term historical volatility, $h_l$ the long-term volatility, and $\theta$ the time of the year measured in radians ($2\pi$ radians = 360 degrees = 1 full year). Equation 6.1 may be recognized as the multivariate regression estimator developed in the previous chapter. To further improve the estimate of expected future volatility, some additional linearization was performed using the third-order polynomials presented immediately below:

$$\begin{aligned} v &= 0.1093x^3 - 0.1821x^2 + 1.0743x - 0.0071 \quad \text{for } x > 0.25 \\ &= -49.47x^3 + 40.76x^2 - 10.19x + 1.021 \quad \text{for } x \le 0.25 \end{aligned} \tag{6.2}$$

In Equation 6.2, $x$ represents the volatility estimate obtained from the regression model specified by Equation 6.1, while $v$ represents the fully linearized estimate of volatility that was intended for use in the determination of the conditional distributions. The first polynomial in Equation 6.2 corrects for small nonlinearities that are not completely accounted for by the regression model. The second polynomial (which has substantial higher-order coefficients) corrects for a more significant nonlinearity—the hook, appearing at the left end of Figure 5–5 in Chapter 5—that is often observed when average range volatility measures are used.

In addition to the expected future volatility ($v$), the terminal price of a nominal \$100 stock ($p$) was calculated. Calculation of the terminal price involved dividing the stock price $n$ bars after the reference by its price at the reference bar and multiplying the result by 100, exactly as in the previous study.

To discount for growth due to interest in a risk-neutral world, the resultant terminal price was multiplied by $\exp(-rt)$, where $r$ was the interest rate at the reference bar and $t$ the holding period measured in years ($t = n$ bars/252 bars per year). The intent was to eliminate the effect of interest on terminal price. Elimination of the interest effect avoided the need to include this fairly well-understood influence in the model as a conditioning variable. Direct inclusion of interest as a conditioning variable would have led to less stable results due to the degrees of freedom that would have been consumed—degrees of freedom better saved for more complex and less well-understood variables that have greater short-term impact on option prices. Anyhow, interest can be more effectively handled in a theoretical manner, one that does not consume any extra degrees of freedom, once the conditional distributions are determined.

The estimated future volatility and interest-discounted terminal stock price were massaged into row and column indices, respectively. Indices were calculated with the same formulae and parameter values employed in the previous study. Once the indices were determined, the array element addressed by them was incremented by one; that is how the bin counts, that form the basis for the conditional frequency distributions, were accumulated.

Finally, another valid reference bar was chosen and, after all valid reference bars for the selected stock were processed, the next stock was selected. The sequence was repeated until all reference bars for all stocks had been analyzed. The result was an array in which each row contained a frequency distribution for the terminal price of a nominal \$100 stock, residing in an interest-free universe, and having a specified level of expected future volatility.

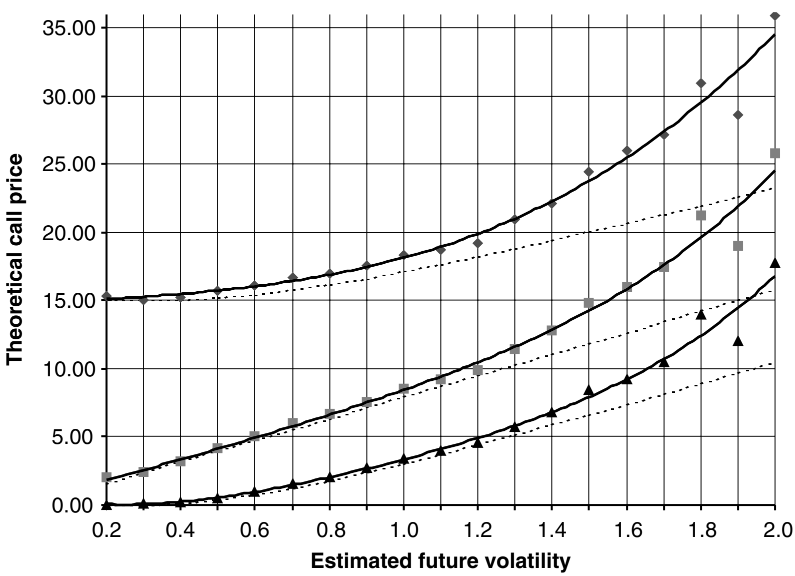
***Second Computational Block*** In the second block, theoretical option premiums for several strikes and option types were calculated from the distributions adduced in the first block. The calculations began with the selection of a volatility level and strike price. The volatility level provided an index into the array row containing the required frequency distribution. Working across the row, column by column, each individual frequency was multiplied by the easily calculated terminal (expiration) price of an option having the selected strike. The calculations were carried out for calls, puts, and straddles. For each volatility, strike, and option type, the sum of the products of price and frequency was divided by the sum of the frequencies. The resultant forward expectations—theoretical option premiums—based on price distributions characteristic of genuine stocks, were placed in another array. Black-Scholes premiums were also calculated for the same set of volatility levels, strikes, and option types in order to provide a baseline for comparison.

## Results

Figure 6–2 shows the theoretical call premiums computed from the empirical distributions and from Black-Scholes. The Black-Scholes premiums are depicted in the figure by dotted lines. Premiums derived from the empirical distributions appear as small black markers. The solid lines drawn roughly through the markers

**FIGURE 6–2**

Theroretical Call Prices from Black-Scholes and from Empirical Distributions When Good Volatility Estimates Are Used



represent smoothing polynomials that were fitted to the empirically derived premiums. Data are presented in Figure 6–2 for three strikes: the topmost curves are for in-the-money calls struck at $85, the middle curves are for at-the-money calls struck at $100, and the lowermost curves are for out-of-the-money calls struck at $115.

As clearly seen in Figure 6–2, theoretical fair call premiums derived from the empirical distributions are generally larger than those calculated with Black-Scholes. The extent to which the empirical distributions yield premiums that exceed those produced by Black-Scholes depends on option strike and level of volatility. Premiums derived from both approaches are quite close to one another for the out-of-the-money call shown in the figure until volatility reaches about 130%; at that point, premiums obtained from the conditional distributions begin to take off relative to those computed using Black-Scholes. Although the volatility thresholds differ, the same pattern appears for the at-the-money call and for the in-the-money call. The threshold at which call premiums derived from distributions of real returns begin to accelerate their growth (relative to those computed with Black-Scholes), declines with increasing moneyness of the options. Regardless of moneyness, for high volatility levels, the empirical distributions yield call premiums that greatly exceed those obtained from Black-Scholes.

Although not shown, theoretical premiums for straddles and puts were also examined. At low volatility levels (less than about 80%), both in-the-money and out-of-the-money puts were found to have higher premiums when analyzed with conditional distributions than when analyzed with Black-Scholes. Empirically derived premiums, however, were less than Black-Scholes premiums for at-the-money and in-the-money puts, especially at high levels of volatility (greater than about 100%). Both Black-Scholes and conditional distribution calculations produced nearly identical theoretical premiums at all levels of volatility for at-the-money straddles.

## Discussion

Premiums based on the conditional distributions deviate from Black-Scholes for both calls and puts. At all levels of volatility, calls have higher relative premiums; at high levels of volatility, they have much higher relative premiums. Both in-the-money and at-the-money puts have lower relative premiums at high levels of volatility. Higher relative premiums are found at low volatility levels for puts that are either in-the-money or out-of-the-money. Premiums computed from conditional distributions for at-the-money straddles, however, deviate hardly at all from those computed using standard Black-Scholes.

***Deviant Call Premiums*** Why are theoretical call premiums for high-volatility stocks so much larger when derived from the conditional distributions than when computed with Black-Scholes? One possible explanation for the larger call premiums associated with the empirical distributions is the volatility payoff. The volatility payoff, a positive return that accrues merely from the action of volatility in a universe of asymmetrically distributed stock price movements, was thoroughly examined in Chapter 3 under the header *Means, Medians, and Stock Returns.* Another explanation might be the association of volatility with market bottoms and imminent gains. Such an association might be generic, or it could be specific to volatility estimates that take seasonality into account, such as the regression estimate used in the current study. Finally, an explanation for the larger call premiums that were found with the conditional distributions might lie with the data; perhaps the data sample used in the analysis was dominated by bullish trends that tend to be magnified by high-Delta options.

The last explanation, that a persistent bullish trend in the data sample is somehow responsible for the greater empirically derived fair premiums, makes the least sense of the several proffered. Although it spans a bullish period, the data sample employed in the study also spans one of the most bearish periods in recent market history. Additionally, returns in the sample were characterized by a near-zero mean and consistently negative skew (see Chapter 4).

That volatility marks bottoms and leads rising stock prices is a more tenable hypothesis. It is a hypothesis that is, at least, consistent with the experiences of traders and technical analysts. The idea that a correlation may exist between the seasonal element in the volatility estimate and bullish trends in the marketplace also has some support. Some of the most volatile periods of the year are also the most bullish ones.

The best explanation proffered as to why theoretical call premiums for high-volatility stocks are so much larger when derived from the conditional distributions than when calculated with Black-Scholes is the one involving a volatility payoff. Option pricing models generally assume that a stock's forward expectation should reflect only growth due to interest in a risk-neutral world. Growth due to volatility in a world of proportionally distributed returns is something that popular models do not assume. As a matter of fact, standard option pricing theory argues that no risk or volatility payoff exists. But tests clearly demonstrate that stocks do yield positive returns that are related to volatility. These returns may be somewhat less than anticipated, given stock prices hypothesized to have a log-normal distribution and an equal probability of either rising or falling over the holding period, but they are assuredly great enough to be significant to traders. Perhaps the risk-neutral assumption made by popular option models is inconsistent with the reality of the marketplace; all risks may not be efficiently hedged or eliminated, even in the world of options. If stocks exhibit volatility-related returns, then it is probable that such returns impact option premiums, especially those of in-the-money options with abundant Delta. In other words, at high levels of volatility, the fair premiums estimated from the empirical distributions may be responding to volatility-induced gains in stock price that are not reflected in models like Black-Scholes that pay homage to the doctrine of risk-neutrality.

A little experiment was performed to test the hypothesis that a volatility payoff in the underlying stock is responsible for the deviant empirical premiums. Theoretical call premiums were calculated using a modified Black-Scholes model. This model was identical to standard Black-Scholes in every respect, except that it did not factor out growth in the underlying stock that could be attributed to volatility. In the standard model, the mean of the underlying stock's log-normal distribution of returns is assumed to be zero when there is no interest; any growth that can be attributed to volatility is factored out by the $\sigma^2/2$ term that appears in Equations 3.18 and 3.22 (see Chapter 3). This expresses the idea that there is no growth unrelated to interest in a risk-neutral universe. In the modified Black-Scholes, the median (rather than the mean) of the underlying log-normal process was assumed to be zero, thus implying a positive volatility-related growth in the underlying stock. The premiums computed with the modified Black-Scholes, which had been revised to include the theoretical volatility payoff, were astonishingly similar to those derived from the conditional distributions, but very different from the premiums calculated using standard Black-Scholes.

How similar were the modified Black-Scholes premiums to those computed from the empirical distributions? Here are some figures: for a predicted volatility of 150%, the theoretical fair premium for the at-the-money call was $14.20 when estimated from the empirical distributions, and $14.80 when computed with the modified Black-Scholes. Standard Black-Scholes priced the same option at $12.05, a figure much lower than the other two. The conditional distribution methodology yielded a fair premium of $23.70 for the in-the-money option, which was close to premium of $23.50 produced by the modified Black-Scholes; standard Black-Scholes yielded a much lower figure of $20.30. With the volatility increased to 180%, the figures for an at-the-money call were $18.50, $18.50, and $14.40 for the conditional distribution, modified, and standard models, respectively. As can be seen, there is a good fit—perfect in the last example—between theory (modified Black-Scholes) and observation (conditional distributions). These findings give strong support to the hypothesis that there exists in stock returns a volatility payoff that has a direct effect on option value, particularly when volatility and option Delta are high.

It should be observed that the underlying volatility payoff implied by the results appears larger than it did for the real stocks analyzed in Table 3–3 (see Chapter 3), but almost perfectly in line with the corresponding theoretical figures that appear in that table. The reason is that the analysis on which the real stock figures in Table 3–3 was based employed standard historical volatility and so these figures are affected by mean reversion. Of course, the theoretical figures in Table 3–3 are not affected by mean reversion. Mean reversion is not a factor in the present study since a much better regression estimator of future volatility was employed.

Finally, a test was run to discover whether the presence of low-priced, illiquid stocks in the data sample could be responsible for the higher relative call premiums derived from the conditional distributions under high volatility conditions. The test entailed recalculation of the entire analysis with a $10 price minimum for a valid reference bar. In the original analysis, the price minimum was set at $2. The results from the calculations with a $10 minimum were very much like those obtained using the $2 minimum. Low-priced stocks with large ask-bid spreads were obviously not responsible for the deviant call premiums observed at high levels of volatility.

***Other Deviant Premiums*** For both in-the-money and out-of-the-money puts on low-volatility stocks, the excess over Black-Scholes in the premiums derived from the conditional distributions is most likely an effect of positive kurtosis and negative skew. Premiums computed using distributions of returns from real stocks are correct in the presence of skew and kurtosis; those computed with Black-Scholes are incorrect.

The lower relative premiums computed from the empirical distributions for at-the-money and in-the-money puts when volatility levels are high almost certainly result from the so-called "volatility payoff," the effect of which is the opposite on puts to what it is on calls. In-the-money puts, like in-the-money calls, possess high Delta and apparently respond strongly to trends, including those that reflect the growth associated with volatility.

***Nondeviant Premiums*** Because implied volatility based on the terminal price of an at-the-money straddle served as the dependent variable when constructing the regression estimator for volatility, it is no surprise that both the standard Black-Scholes and the conditional distribution approaches yield the same premiums. Think of it this way: the volatility that must be entered into Black-Scholes to obtain the terminal straddle price was precisely that volatility to be estimated by the multivariate regression. The regression-based volatility estimate was, in fact, specifically designed to force Black-Scholes to yield an empirically correct valuation in the specific case of an at-the-money straddle.

## STUDY 3: RE-ANALYSIS WITH DETRENDED DISTRIBUTIONS

Positive returns or price trends, perhaps deriving from the volatility payoff, were thought responsible for the higher call and lower put premiums that were observed with the empirical model when it was compared to Black-Scholes in the previous study. In that study, trends were left in; they were not selectively removed from the observed distributions. Were positive trends or returns—linked, for whatever reason, to high levels of volatility—at the root of the differences between the models in the previous study? Standard models like Black-Scholes assume that no trends other than those due to interest-related growth in a risk-neutral world exist in the marketplace. How close to those obtained from Black-Scholes would theoretical option premiums derived from conditional distributions be for stocks that are trendless, and thus satisfy the risk-neutral assumption, but that are otherwise identical to real stocks? Equivalently, what would be found were the conditional distributions of the previous study *detrended*—i.e., were their first moments somehow adjusted to eliminate trend? Finally, how can trend be adjusted in, added to (as might be necessary to account for growth due to interest or other factors), or removed from a pricing model based on conditional distributions? These are the questions addressed in this study.

### Method

As in Study 2, the analysis was carried out in two computational blocks. The first of the two blocks was unchanged in the present study. Discussion of the first block is, therefore, not repeated (see Study 2 for details). The few differences that existed between the analyses appeared in the second computational block where, in the previous study, fair option premium was computed from each conditional distribution as

$$\xi = \frac{\sum_i f_i \mathrm{P}(s_i)}{\sum_i f_i} \tag{6.3}$$

where $\xi$ represented the forward expectation, $f_i$ the frequency associated with the $i$-th terminal stock price level for the specified conditional distribution (in the language of the previous study, the $i$-th element of the array row associated with the specified level of predicted volatility), and $s_i$, the corresponding terminal stock price. Given the terminal stock price, the function P simply returned the terminal premium (the premium at expiration) for an option of the specified strike and type.

To remove the effect of trend from the underlying distribution of terminal stock price and, in turn, from the theoretical option premium, $s_i$ was replaced with $s_i - u + 100$ in the current study, leading to

$$\xi = \frac{\sum_i f_i \mathrm{P}(s_i - u + 100)}{\sum_i f_i} \tag{6.4}$$

where $u$ was set to $(\sum f_i s_i)/(\sum f_i)$ the first moment of the frequency distribution defined by $f$, i.e., the expectation for the terminal stock price. The logic involved is simple: The stock price associated with each frequency is shifted up or down by some fixed amount. Shifting the stock prices shifts the first moment of the stock price distribution, its expectation or mean. The amount of shift is chosen to make the expectation for the terminal stock price whatever is desired—in this case an even \$100, the initial price of the stock, implying the absence of any growth or trend. The terminal option premium is then computed for each level of shifted stock price and the option's forward expectation or fair current premium determined in the usual manner.

The second computational block was the same as in Study 2 (see that study for procedural details) except for the change to the premium calculation that was just discussed.

### Results

Table 6–2 presents theoretical call premiums for volatility levels (VLTY) ranging from 20% to 200% in increments of 10%. Data are shown in the table for options having three strikes or levels of moneyness: an $85 in-the-money strike; a $100 at-the-money strike; and a $115 out-of-the-money strike. Fair premiums derived from the detrended empirical distributions (EMP) are listed side-by-side with those calculated using standard Black-Scholes (BS). The numerical data in Table 6–2 are

**TABLE 6–2**

Theoretical Call Premiums for Three Strikes and a Range of Volatilities Computed from Detrended Conditional Distributions and with Black-Scholes

| | *Strike = $85* | | *Strike = $100* | | *Strike = $115* | |
|---|---|---|---|---|---|---|
| **VLTY** | **EMP** | **BS** | **EMP** | **BS** | **EMP** | **BS** |
| 0.20 | 15.04 | 15.00 | 1.87 | 1.59 | 0.03 | 0.00 |
| 0.30 | 15.09 | 15.01 | 2.40 | 2.38 | 0.08 | 0.02 |
| 0.40 | 15.20 | 15.06 | 3.16 | 3.18 | 0.22 | 0.14 |
| 0.50 | 15.40 | 15.20 | 3.98 | 3.97 | 0.50 | 0.39 |
| 0.60 | 15.69 | 15.44 | 4.82 | 4.77 | 0.93 | 0.76 |
| 0.70 | 15.98 | 15.77 | 5.59 | 5.56 | 1.43 | 1.24 |
| 0.80 | 16.33 | 16.18 | 6.32 | 6.35 | 1.93 | 1.79 |
| 0.90 | 16.74 | 16.64 | 7.09 | 7.14 | 2.52 | 2.39 |
| 1.00 | 17.17 | 17.14 | 7.82 | 7.93 | 3.12 | 3.04 |
| 1.10 | 17.69 | 17.68 | 8.56 | 8.72 | 3.68 | 3.71 |
| 1.20 | 18.11 | 18.24 | 9.20 | 9.51 | 4.18 | 4.42 |
| 1.30 | 18.72 | 18.83 | 10.06 | 10.30 | 4.95 | 5.14 |
| 1.40 | 19.44 | 19.44 | 10.99 | 11.09 | 5.75 | 5.88 |
| 1.50 | 20.32 | 20.06 | 11.96 | 11.88 | 6.70 | 6.63 |
| 1.60 | 20.28 | 20.69 | 11.94 | 12.66 | 6.72 | 7.39 |
| 1.70 | 21.26 | 21.33 | 13.11 | 13.45 | 7.65 | 8.15 |
| 1.80 | 22.73 | 21.97 | 15.05 | 14.23 | 9.55 | 8.93 |
| 1.90 | 21.70 | 22.63 | 13.96 | 15.01 | 8.45 | 9.71 |
| 2.00 | 24.18 | 23.29 | 16.52 | 15.79 | 10.74 | 10.49 |

analogous to those presented graphically in Figure 6–2 in the previous study.

For options that are well out-of-the-money and low in time value, fair premiums computed from the detrended conditional distributions substantially exceed those obtained from standard Black-Scholes. This is true for both calls (shown in the table) and puts (not shown) up to a volatility level of about 80%. Deeply in-the-money options also display larger theoretical premiums over a similar range of volatilities when computed with the detrended distributions than when determined using Black-Scholes. At higher levels of volatility, between about 110% and 170%, the chosen options are relatively less in-the-money or out-of-the-money, and evidence smaller distribution-based premiums. Options that are at-the-money and high in time value, and hence associated with high levels of volatility, also tend to have less empirical value, as assessed by the conditional distribution methodology, than Black-Scholes would suggest. Overall, however, there is a surprisingly close agreement between premiums from standard Black-Scholes and those from the detrended conditional distributions for at-the-money options on stocks having moderate levels of volatility (between 30% and 100%). These observations are consistent with what might be expected given the differences between the log-normal distributions assumed by Black-Scholes and the leptokurtic ones actually observed in stock returns.

When examining the data, remember that "in-the-money" and "out-of-the-money" are relative terms that must be understood in the context of volatility: a $115 strike call on a $100 stock is deeply out-of-the-money (in the sense that the stock price is very unlikely to rise above $115 by expiration) when 10 days remain and volatility is low, i.e., less than 40%; on the other hand, in high-volatility situations, i.e., when volatility is greater than 100%, the same option would only be considered mildly out-of-the-money.

In addition to puts and calls, at-the-money straddles were investigated. Consistent with theoretical considerations, the premium for an at-the-money straddle has a straight-line relationship with estimated future volatility. This relationship is easily seen when smoothing polynomials are used. Virtually identical premiums (within the limits of statistical measurement error) are obtained whether the Black-Scholes or conditional distribution methodology is employed for the calculations. This was expected, given that the volatility estimator used in the study was designed to make Black-Scholes generate empirically correct premiums for at-the-money straddles.

Substantial noise, in the form of variability of premiums from volatility level to volatility level, is visible in the data from the detrended empirical distributions when volatility is high. The relatively small number of cases falling in the bins corresponding to high volatility levels is responsible for the observed noise. Smoothing polynomials are effective in this context for noise reduction, not to mention for interpolation between discrete levels of volatility. Such polynomials, along with the raw data, were plotted in the figures shown earlier but are not presented in Table 6–2 for reasons of space.

## Discussion

This study demonstrated how the first moment of a conditional distribution, its growth or trend, could be set to zero or otherwise adjusted. The process involved shifting, after the fact, the stock prices that tagged the bins in which the frequencies of occurrence were accumulated. When options were priced with unmodified empirical distributions (those that reflected any trend present in the data) substantial deviations from Black-Scholes were observed at high levels of volatility. One hypothesis was that the deviations were caused by a volatility-related trend component, such as the volatility payoff discussed near the end of Chapter 3. Removing the trends by setting the first moments to $100, the initial stock price, eliminated the extremely deviant premiums associated with high levels of volatility when options were appraised using conditional distributions. There is no question that trends, whatever their source, were responsible for the great deviations from Black-Scholes with increasing volatility levels. When trend is statistically partialled out, there is a much closer match between premiums derived from the conditional distribution methodology and from Black-Scholes for all volatility levels, option types, and degrees of moneyness.

Nevertheless, significant differences remain between premiums gleaned from the empirical distributions and those from Black-Scholes. These differences are quite consistent with what might be expected given the log-normal distribution of returns assumed by Black-Scholes and the more leptokurtic distribution of returns observed when real stock prices are examined. The disparities in premiums between the models are by no means trivial; for certain levels of volatility and moneyness, they are of clear practical significance. For example, consider a call struck at $115 and having 10 days left until expiration. Assume that this call is on a stock with an estimated future volatility of 60%. The fair premium for such a call is $0.76 when computed with Black-Scholes, and $0.93 when reckoned using the conditional distributions—a 22% spread. Few traders would want to ignore such a large divergence in theoretical fair value.

Having demonstrated the use of detrended conditional distributions in pricing options, this study raises the issue of whether or not the trends and growth observed in stock prices should always be partialled out or statistically removed. For example, it may not always be desirable to eliminate the effect of the volatility payoff on theoretical option prices. When options must be priced for speculative purposes, it may in fact be best to use the unaltered conditional distributions. Finally, it is worth pointing out that, even if trends have been removed from the raw data, growth and trend factors may be brought into the model explicitly, at a later stage, by setting the first moment of each conditional distribution to a theoretically appropriate or desired value.

## STUDY 4: SKEW AND KURTOSIS

The previous three studies examined volatility as a conditioning variable. Volatility, the second moment, has a known and dramatic influence on the fair value of an option. But what about the third and fourth moments, skew and kurtosis? This study adds skew and kurtosis as conditioning variables to the models of the previous studies. More precisely, the current study examines the skew or kurtosis that characterizes a given stock at a given time—the stock's historical skew or kurtosis—as a factor that may influence fair option premium. Just like historical volatility, historical skew or kurtosis can be computed and used as an input to an option pricing model. Note that the skew or kurtosis specific to a given stock at a given time is being considered here, not the skew and kurtosis characteristic of price movements generally. The latter are automatically taken into account when using empirically derived conditional distributions (unless these distributions have been modified in some way that removes the higher moments) since such distributions are derived from actual price movements.

How is historical skew or kurtosis computed? The measurement of historical skew or kurtosis is straightforward and may be accomplished with Equations 4.3 and 4.4 from Chapter 4. The only issue is what to use as a data sample, or sample of returns, in the calculations. In this study, the daily returns over the past $m$ bars (relative to the reference) were employed as the sample of values to which Equations 4.3 and 4.4 could be applied in order to obtain the desired statistics. It should be noted that historical measures for skew or kurtosis are substantially less reliable than historical measures for volatility. The lower reliability is due to the greater statistical error of measurement with the higher moments. To achieve the same reliability—or statistical error of estimate, the other side of the same coin—with skew or kurtosis that was achieved with volatility, a much larger sample of data points would have to be analyzed.

In this study, as in the previous one and those that follow, the conditional distributions used to price options are detrended. The reason is that it is easier to see the specific effects of skew, kurtosis, or whatever other conditioning variable is being investigated on option premiums in the absence of volatility-related growth than in its presence. Removal of trends from the conditional distributions moves the theoretical premiums computed from these distributions closer to those calculated with Black-Scholes, thus making comparisons easier. This is an example of the good old scientific method: to separate out and hold constant all influences other than those under investigation.

## Method

Except for some minor changes and additions, the analytic approach is similar to that of the previous study. First, a stock was chosen from the 2,246 stocks in the database. A reference bar was then selected from the 1,834 bars (not all active) available for the chosen stock. The reference bar was checked for validity. To be valid, a reference bar had to satisfy the two requirements used in the previous study: one, the stock had to be active and have an unadjusted price greater than $2 in the $m1 + m2$ bars preceding the reference; and two, there had to exist a period of at least $n$ bars following the reference from which a return or terminal price could be computed. In the analysis, $m1$ was 30, $m2$ was 70, and $n$ was 10—exactly as before. Given a valid reference bar, an estimate of future volatility based on two historical measures, together with three seasonal harmonics, was determined. The volatility estimate was the same as that used in the previous two studies. In addition to volatility, historical skew or kurtosis was assessed from the $m1 + m2$ bars that preceded the reference. Daily returns over the historical period were used in the calculation of these higher moments. There was then one of four possible conditions to verify: (1) That skew was positive or greater than zero, (2) That skew was negative or less than zero, (3) That kurtosis was positive or greater than zero, or (4) That kurtosis was negative or less than zero. If the specified condition was verified, the terminal price frequencies were accumulated for the stock. The accumulation of frequencies was carried out exactly as in the two foregoing studies. Finally, the next reference bar was selected and, when all valid reference bars for the chosen stock were processed, the next stock was picked. The sequence was repeated until every valid reference bar for every stock had been analyzed. The result was a set of conditional distributions for a specified level of either skew or kurtosis, and for each of the 20 levels of estimated future volatility. These distributions were then used to compute theoretical option premiums.

The whole procedure just described was repeated four times. Each time, one of the four possible conditions was specified. In this way, terminal stock price distributions for four possible skew or kurtosis situations (skew $> 0$, skew $< 0$, kurtosis $> 0$, kurtosis $< 0$) and for all 20 levels of predicted volatility (10%, 20%, . . . 200%) were computed along with derived option valuations.

It should be clear from the above discussion that only two levels of skew and two levels of kurtosis were analyzed, and not both at the same time. There were two reasons for handling skew and kurtosis in this manner. One reason was the large statistical error associated with estimation of the higher moments (see Chapter 4 for details), especially from the small samples involved when computing historical measures. For the same lookback period, measures of historical skew and kurtosis are much less reliable than, for instance, measures of historical volatility. Given the comparatively low reliability, two levels (positive or negative) appeared sufficient to test the hypothesis that historical skew and kurtosis would be useful in the appraisal of option value. A second reason for the coarseness of the analysis was to conserve degrees of freedom. A finer grained analysis would absorb more degrees of freedom and, given the low reliability of historical skew and kurtosis measurements, would not necessarily provide more in the way of useful information, at least in a preliminary investigation.

## Results

Theoretical premiums derived from the conditional distributions and from Black-Scholes are shown in Tables 6–3 and 6–4. Table 6–3 contains theoretical fair premiums for out-of-the-money calls having a strike price of $115 and 10 days left to expiration. Theoretical premiums for out-of-the-money puts struck at $85 appear in Table 6–4. Option premiums are shown in the tables for 19 levels of predicted volatility, ranging from 0.20 through 2.00, and for all four conditions involving skew and kurtosis; premiums were not displayed for all 20 levels of volatility because the lowest level (0.10) occurred too infrequently to yield meaningful results. In both tables, the first column (VLTY) contains the volatility levels corresponding to the bins used to accumulate the conditional distributions. Premiums for options on stocks having positive historical skew appear in the second column (SKEW > 0). The third column (SKEW < 0) contains the premiums for options on stocks that display negative skew. Fair premiums for options on stocks with positive historical kurtosis appear in the fourth column

(KURT > 0), while those for stocks with negative kurtosis are listed in the fifth column (KURT < 0). The sixth, and last, column contains fair premiums computed using Black-Scholes. Since neither skew nor kurtosis is an input to the standard Black-Scholes model, only one column (BS) of data for this model is required in the tables. The premiums in the last column are intended to serve as baselines for comparison with those derived from the conditional distributions and found in columns two through five.

***Skew, Kurtosis, and Out-of-the-Money Calls*** The data in Table 6–3 suggest that historical skew has only a minor

**TABLE 6–3**

Fair Premiums for Out-of-the-Money $115 Calls with 10 Days Left to Expiration as a Function of Volatility, Skew, and Kurtosis

| VLTY | SKEW > 0 | SKEW < 0 | KURT > 0 | KURT < 0 | BS |
|---|---|---|---|---|---|
| 0.20 | 0.04 | 0.03 | 0.04 | 0.01 | 0.00 |
| 0.30 | 0.09 | 0.07 | 0.09 | 0.04 | 0.02 |
| 0.40 | 0.24 | 0.21 | 0.23 | 0.13 | 0.14 |
| 0.50 | 0.51 | 0.49 | 0.52 | 0.38 | 0.39 |
| 0.60 | 0.93 | 0.94 | 0.95 | 0.73 | 0.76 |
| 0.70 | 1.46 | 1.38 | 1.45 | 1.21 | 1.24 |
| 0.80 | 1.95 | 1.88 | 1.94 | 1.80 | 1.79 |
| 0.90 | 2.52 | 2.52 | 2.55 | 2.19 | 2.39 |
| 1.00 | 3.05 | 3.23 | 3.17 | 2.66 | 3.04 |
| 1.10 | 3.66 | 3.71 | 3.72 | 3.25 | 3.71 |
| 1.20 | 4.12 | 4.27 | 4.18 | 4.24 | 4.42 |
| 1.30 | 4.93 | 4.95 | 4.96 | 4.66 | 5.14 |
| 1.40 | 5.64 | 5.85 | 5.71 | 6.15 | 5.88 |
| 1.50 | 6.90 | 6.35 | 6.58 | 8.10 | 6.63 |
| 1.60 | 6.87 | 6.39 | 6.65 | 7.06 | 7.39 |
| 1.70 | 8.41 | 6.22 | 7.74 | 6.16 | 8.15 |
| 1.80 | 10.09 | 8.53 | 9.41 | 10.35 | 8.93 |
| 1.90 | 9.65 | 6.85 | 8.17 | 14.26 | 9.71 |
| 2.00 | 9.45 | 12.04 | 10.69 | 11.93 | 10.49 |

impact upon fair call premiums when the underlying stock's volatility is low or moderate, although a more significant impact is observed when the stock's volatility is high. Out-of-the-money calls on stocks with positive skew, and with predicted volatility levels between 20% and 80%, tend to have slightly larger fair premiums than calls on similar stocks with negative skew. This is consistent with the notion that out-of-the-money calls should be worth more in the presence of positive skew due to the greater likelihood of a large gain in stock price. For volatility levels between 90% and 140%, fair call premiums tend to be slightly smaller when the underlying stocks are characterized by positive, rather than by negative, historical skew. However, substantially greater call premiums are found for stocks that have positive skew and predicted volatility levels between 150% and 190%.

The story is different with kurtosis. Historical kurtosis has a major effect upon theoretical call premiums at all levels of estimated future volatility. In low volatility situations, out-of-the-money calls generally have considerably larger theoretical premiums when the underlying stocks exhibit positive historical kurtosis (a leptokurtic distribution of returns in the recent past) than when they do not. The effect is present for volatility levels ranging from 20% to 130%. At higher levels of volatility, from 140% to 200%, the pattern is reversed: negative historical kurtosis is associated with greater option premiums. This is precisely the behavior that should be observed if kurtosis, like volatility, is a moderately stable—and thus somewhat predictable—characteristic of a stock.

At low volatility levels (20% to 50%), premiums computed with Black-Scholes fall at or below those based on distributions of returns from stocks having negative historical kurtosis; Black-Scholes premiums fall far below those computed for stocks having positive historical kurtosis. For stocks that exhibit higher levels of volatility, Black-Scholes premiums tend to fall between those computed for stocks with negative kurtosis and those calculated for stocks with positive kurtosis. The exact relationship varies with volatility and is most likely a function of general market kurtosis and of regression to the mean.

Remember that the market as a whole is characterized by positive kurtosis. Any negative historical kurtosis in a given stock's returns will, therefore, tend to regress to the more positive mean kurtosis with the arrival of the future. This is the same phenomenon observed with raw historical volatility and that can be expected to occur when using raw historical measures of any higher moments. In fact, there will be more regression to the mean with kurtosis than with volatility because of the former's lower reliability. At low levels of volatility, the positive future kurtosis that probably arises even with stocks that evidence negative historical kurtosis causes theoretical premiums for out-of-the-money options to exceed those suggested by Black-Scholes. At higher levels of volatility, the options become effectively less out-of-the-money and the impact of kurtosis becomes smaller. The theoretical premiums thus show less kurtosis-induced inflation and appear more in line with Black-Sholes.

***Skew, Kurtosis, and Out-of-the-Money Puts*** Table 6–4 contains the same kind of data as shown in Table 6–3, but for out-of-the-money puts rather than for out-of-the-money calls. Consistent with the idea that negative skew should enhance the value of an out-of-the-money put, puts on stocks with predicted volatilities between 20% and 80% have greater value when historical skew is negative than when it is positive. The magnitude of the premium difference, however, is fairly small. Put premiums are paradoxically larger under negative skew conditions when the underlying stock has an estimated future volatility between 130% and 190%.

Variations in kurtosis also lead to substantial differences in option premium. Positive kurtosis is associated with decidedly greater premiums than negative kurtosis for puts on stocks with volatilities between 20% and 90%. For volatilities between 110% and 200%, positive kurtosis is associated with lower put prices. As was the case for out-of-the-money calls, Black-Scholes underprices out-of-the-money puts at lower levels of volatility, regardless of historical kurtosis; at higher levels of volatility, Black-Scholes premiums fall between those associated with positive and those identified with negative kurtosis.

The relationships observed between put premiums and kurtosis are consistent with expectations. The fact that, when

**TABLE 6–4**

Fair Premiums for Out-of-the-Money $85 Puts with 10 Days Left to Expiration as a Function of Volatility, Skew, and Kurtosis

| VLTY | SKEW > 0 | SKEW < 0 | KURT > 0 | KURT < 0 | BS |
|---|---|---|---|---|---|
| 0.20 | 0.03 | 0.04 | 0.04 | 0.02 | 0.00 |
| 0.30 | 0.09 | 0.09 | 0.09 | 0.06 | 0.01 |
| 0.40 | 0.19 | 0.22 | 0.21 | 0.14 | 0.06 |
| 0.50 | 0.39 | 0.42 | 0.41 | 0.29 | 0.20 |
| 0.60 | 0.66 | 0.73 | 0.70 | 0.55 | 0.44 |
| 0.70 | 0.96 | 1.01 | 0.99 | 0.91 | 0.77 |
| 0.80 | 1.27 | 1.41 | 1.34 | 1.22 | 1.18 |
| 0.90 | 1.74 | 1.74 | 1.75 | 1.63 | 1.64 |
| 1.00 | 2.12 | 2.24 | 2.17 | 2.13 | 2.14 |
| 1.10 | 2.71 | 2.65 | 2.68 | 2.80 | 2.68 |
| 1.20 | 3.09 | 3.12 | 3.09 | 3.33 | 3.24 |
| 1.30 | 3.77 | 3.57 | 3.65 | 4.58 | 3.83 |
| 1.40 | 4.48 | 4.30 | 4.35 | 5.84 | 4.44 |
| 1.50 | 5.60 | 4.78 | 5.21 | 6.51 | 5.06 |
| 1.60 | 5.19 | 5.33 | 5.16 | 6.10 | 5.69 |
| 1.70 | 6.80 | 5.16 | 6.26 | 5.51 | 6.33 |
| 1.80 | 7.93 | 7.09 | 7.56 | 9.07 | 6.97 |
| 1.90 | 7.25 | 5.75 | 6.48 | 11.14 | 7.63 |
| 2.00 | 8.28 | 10.11 | 9.18 | 8.32 | 8.29 |

predicted volatility is low, even negative historical kurtosis is associated with larger distribution-based premiums than Black-Scholes premiums suggests either that (1) regression to the mean is so great that even negative historical kurtosis is followed by a positive average future kurtosis, or (2) that the negative skew that characterizes the stock market as a whole has sufficiently raised the value of far out-of-the-money (when considered in the light of volatility) puts.

### *Skew, Kurtosis, and In-the-Money Options*

Although the data are not presented, deeply in-the-money calls have the same amounts of time value as out-of-the-money puts with the same strike, and vice versa. Consequently, skew, kurtosis, and volatility affect in-the-money options of one kind much like they affect out-of-the-money options of the opposing kind.

Larger in-the-money put premiums are found with positive skew than with negative skew for volatility levels between 150% and 190%; smaller premiums tend to be found with volatility figures between 90% and 140%; and slightly larger premiums are observed at volatility levels between 20% and 80%. Premiums for in-the-money calls tend to be much larger when skew is negative than when it is positive for volatility levels between 130% and 190%; they are consistently somewhat smaller for levels of volatility between 20% and 80%.

When volatility is low, positive kurtosis is associated with larger theoretical premiums for both puts and calls, and negative kurtosis is associated with smaller premiums. When volatility is high, positive kurtosis is associated with smaller premiums and negative kurtosis with larger premiums.

***At-the-Money Options*** Because the conditional distributions were detrended, and no interest or dividend payments were entered into the model, at-the-money calls and at-the-money puts have identical premiums.

At lower volatility levels (less than 140%), premiums are virtually unchanged across skew conditions. Positive skew is associated with greater call premiums than is negative skew at higher levels of volatility (between 150% and 190%). Black-Scholes premiums tend to fall in or near the range defined by the premiums associated with the two distinct levels of skew. The latter finding makes sense since at-the-money option premiums are less affected by general market skew or kurtosis than are the premiums of either deeply in-the-money or deeply out-of-the-money options.

Kurtosis exerts a stronger and more consistent influence upon the value of at-the-money options than does skew. At-the-money option premiums are greater with positive historical kurtosis than with negative kurtosis at lower levels of volatility (below 110%). At higher volatility levels (above 120%), the pattern is reversed; positive historical kurtosis is associated with lower premiums than negative historical kurtosis.

The difference in the premiums between the two conditions is considerable and of practical relevance to traders.

Much random variation or "noise" is apparent in the premiums derived from the conditional distributions at volatility levels in excess of about 140%. The cause is the relatively small number of cases that contribute to the high-volatility conditional distributions from which the premium estimates are computed. The random estimation error is greater under the negative than under the positive kurtosis condition because, at all levels of volatility, the number of instances where kurtosis is less than zero is by far smaller than the number of cases where historical kurtosis is greater than zero.

## Discussion

The observations made in this study are quite encouraging. Fairly simple historical measures of skew and kurtosis were employed, no corrections were made for regression to the mean, and only a coarse, two-level analysis was performed for each historical measure. Despite these limitations, effects large enough to be important to traders were observed.

Skew has only a modest effect on option value at low levels of volatility. For out-of-the-money options on low-volatility stocks, the effect of skew differs from puts to calls and is what might naively be expected on the basis of distributional shape; positive skew yields relatively greater call premiums, while negative skew is associated with larger put premiums. There is little variation in premium between the two skew conditions for at-the-money options when volatility is low. At high volatility levels, skew has a stronger and more consistent effect on theoretical premiums: all options, regardless of moneyness or type, evidence greater value when historical skew is positive than when it is negative. This latter observation suggests that historical skew might have some predictive capacity, not only for future skew, but also for future volatility.

By way of contrast, kurtosis has a tremendous, easily demonstrable effect on the worth of options at all levels of volatility. Even simple historical kurtosis, a measure that has low reliability and is thus a fairly poor predictor of kurtosis in the near future, can forecast sizable variations in fair value. How strong is the effect of kurtosis? Consider a $100 stock with an historical volatility of 40%. With a volatility of 40%, the theoretical premium for a call struck at $115 is $0.23 when historical kurtosis is positive, but only $0.13 when it is negative. Black-Scholes prices the same call at $0.14, ignoring kurtosis. A fair call premium of $0.95 is found under positive kurtosis conditions, and a premium of $0.73 under negative kurtosis conditions, when the underlying stock's volatility is 60%. When historical kurtosis is positive, the fair value of the call exceeds the Black-Scholes estimate of $0.76 by 25%. As easily seen, these are not minuscule differences of interest only to academicians. Positive kurtosis means greater value for all options (both puts and calls, regardless of moneyness) at low levels of volatility; it means lower fair premiums for options at higher levels of volatility.

Although the effects of kurtosis on the value of in-the-money and out-of-the-money options are consistent with theoretical expectations based on the shape of the underlying price distributions, the similar effects observed on the worth of at-the-money options are curious and were not anticipated. One might speculate on the cause. Could the effect of historical kurtosis on fair value be a result of its action on future volatility and not a direct function of kurtosis itself? When estimated future volatility is low, positive historical kurtosis might be associated with an actual future volatility that exceeds the predicted value. And, while the detrending employed in the analysis removes any variation in the volatility payoff, it does not remove variation in future volatility. Deviations from predicted future volatility that are related to differences in historical kurtosis might explain some of the observations made in the current study.

Whether future volatility or the volatility payoff differs with the level of kurtosis (or skew) can easily be tested with the companion software and is left as an exercise for the reader. Theoretically, the volatility payoff should be greater with greater kurtosis, even when volatility is held constant, because of the longer tails in the price distribution. Whether volatility itself is affected is a matter of empirical investigation. There is no reason why models for volatility and for the volatility payoff, such as those developed in Chapter 5, cannot include measures of skew and kurtosis among the independent variables.

It should be noted that neither skew nor kurtosis appears as an input to any popular volatility or pricing model. Nevertheless, in the light of the results just presented, an advanced volatility or pricing model would do well to include some measure of skew and of kurtosis. Both skew and kurtosis are easily handled as model inputs when using the conditional distribution methodology and other techniques discussed later in this book. The next step along this line of research is to develop better estimates for future skew and kurtosis, just as was done in Chapter 5 for volatility. Given the intrinsically lower reliability of historical measurements of the higher moments, such efforts are likely to be worthwhile.

## STUDY 5: TRADING VENUE

This study does for trading venue what the previous study did for skew and kurtosis. Along with volatility, venue is used as a conditioning variable. The venues examined are the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the so-called "over-the-counter" markets (NASD). Stocks trading on each of these exchanges or venues are, in the current study, analyzed by having their returns binned, price distributions computed, and theoretical premiums for options trading on them derived. The reason for this study is that venue, like skew and kurtosis, may affect theoretical option premiums. The possibility of option value being affected by stock venue was suggested by tests reported by the authors in *Futures* (Katz and McCormick, 2001b,c) and by observations made when trading; options on NASD stocks consistently appear to have higher implied volatilities, even when the historical volatilities of the underlying stocks are the same—a phenomenon that could be due to a higher mean volatility in the NASD coupled with a tendency for historical volatility to exhibit mean reversion. It should be noted that the current study uses an approach like the one employed in the aforementioned articles appearing in *Futures*.

### Method

The analysis was conducted in the same manner as in Study 4, but with one exception: the selection of cases based on skew or kurtosis was replaced with selection based on venue or exchange. Analyses were carried out for three venues: the NYSE, the AMEX, and the NASD. For each of these venues, distributions were accumulated and theoretical option premiums based on terminal premium expectation were computed. Premiums were computed for both puts and calls over a range of strikes and levels of volatility. For details concerning the calculations, consult Studies 3, 4, and 5, above.

## Results

Only results for the NYSE and the NASD are discussed. Results for the AMEX are not presented since there were too few stocks with prices greater than $2 trading on that exchange (and present in the Worden Bros. TC-2000 database) to yield a sufficient number of data points for stable and meaningful statistics or theoretical option premiums to be computed.

Table 6–5 displays theoretical premiums broken down by moneyness (out-of-the-money, in-the-money), volatility (20% to 200% in steps of 10%), and source of premium (NYSE conditional distributions, NASD conditional distributions, Black-Scholes).

### *Out-of-the-Money Options, Detrended Distributions*

As Table 6–5 reveals, both puts and calls have lower theoretical premiums for stocks trading on the NYSE than for stocks trading on the NASD when volatility levels are low (25% to 60%) and systematic trend is absent. At low levels of volatility, Black-Scholes yields theoretical premiums even smaller than those computed from the NYSE data. The differences observed in the premiums are substantial, often reaching more than 50% of the option's value! At higher volatility levels, theoretical premiums for options on NYSE stocks exceed those for options on NASD stocks, with a few exceptions that are most likely artifacts resulting from the limited samples of data points available at some levels of volatility. Black-Scholes premiums are generally below the NYSE premiums, but frequently above the NASD premiums.

In-the-money options are not discussed because, when the first moment of the distribution of returns is zero, i.e., when there

**TABLE 6–5**

Fair Premiums Based on Detrended Conditional Distributions as a Function of Trading Venue for Out-of-the-Money Options with 10 Days Left to Expiration on a Nominal $100 Stock

| | *Out-of-the-Money $115 Calls* | | | *Out-of-the-Money $85 Puts* | | |
|---|---|---|---|---|---|---|
| **VLTY** | **NYSE** | **NASD** | **BS** | **NYSE** | **NASD** | **BS** |
| 0.2 | 0.04 | 0.04 | 0.00 | 0.04 | 0.06 | 0.00 |
| 0.3 | 0.07 | 0.16 | 0.02 | 0.08 | 0.12 | 0.01 |
| 0.4 | 0.20 | 0.31 | 0.14 | 0.19 | 0.24 | 0.06 |
| 0.5 | 0.44 | 0.59 | 0.39 | 0.38 | 0.43 | 0.20 |
| 0.6 | 0.89 | 0.97 | 0.76 | 0.66 | 0.71 | 0.44 |
| 0.7 | 1.40 | 1.44 | 1.24 | 1.00 | 0.98 | 0.77 |
| 0.8 | 2.00 | 1.91 | 1.79 | 1.44 | 1.31 | 1.18 |
| 0.9 | 2.80 | 2.48 | 2.39 | 1.82 | 1.73 | 1.64 |
| 1.0 | 3.52 | 3.08 | 3.04 | 2.36 | 2.16 | 2.14 |
| 1.1 | 4.11 | 3.66 | 3.71 | 3.15 | 2.66 | 2.68 |
| 1.2 | 4.97 | 4.14 | 4.42 | 3.65 | 3.08 | 3.24 |
| 1.3 | 5.72 | 4.89 | 5.14 | 4.19 | 3.68 | 3.83 |
| 1.4 | 6.27 | 5.72 | 5.88 | 4.53 | 4.44 | 4.44 |
| 1.5 | 5.84 | 6.75 | 6.63 | 3.52 | 5.43 | 5.06 |
| 1.6 | 8.99 | 6.53 | 7.39 | 6.79 | 5.16 | 5.69 |
| 1.7 | 9.41 | 7.56 | 8.15 | 6.90 | 6.22 | 6.33 |
| 1.8 | 12.22 | 9.44 | 8.93 | 12.27 | 7.56 | 6.97 |
| 1.9 | 5.78 | 8.56 | 9.71 | 6.40 | 6.72 | 7.63 |
| 2.0 | 16.52 | 10.27 | 10.49 | 10.07 | 9.12 | 8.29 |

is no trend, in-the-money calls behave (as far as time value) precisely like out-of-the-money puts (with the same strike), and in-the-money puts act precisely like out-of-the-money calls. Since the conditional distributions on which this subset of results is based were detrended, trend is effectively absent and the put-call parity that is involved in the aforementioned relationship between in-the-money and out-of-the-money options holds.

The effect of venue on option premiums appears to be very much like that of kurtosis: both positive kurtosis and the NASD venue are associated with relatively higher premiums at lower levels of volatility, and with relatively lower premiums at higher volatility levels; negative kurtosis and the NYSE trading venue are associated with the opposite pattern—higher relative premiums at lower volatility levels, and lower ones at higher levels.

***At-the-Money Options, Detrended Distributions*** Table 6–6 shows theoretical fair premiums for at-the-money options on stocks that have no systematic trends. The data are broken down by option type (put, call), venue (NYSE, NASD, BS), and volatility level (0.2, 0.3, ...1.6). The pattern formed by the premiums is extremely similar to the one observed for the out-of-the-money

**TABLE 6–6**

Fair Premiums Based on Detrended Conditional Distributions as a Function of Trading Venue for At-the-Money Options with 10 Days Left to Expiration on a Nominal $100 Stock

| | *At-the-Money $100 Calls* | | | *At-the-Money $100 Puts* | | |
|---|---|---|---|---|---|---|
| **VLTY** | **NYSE** | **NASD** | **BS** | **NYSE** | **NASD** | **BS** |
| 0.2 | 1.87 | 2.04 | 1.59 | 1.87 | 2.04 | 1.59 |
| 0.3 | 2.40 | 2.50 | 2.38 | 2.40 | 2.50 | 2.38 |
| 0.4 | 3.13 | 3.26 | 3.18 | 3.13 | 3.26 | 3.18 |
| 0.5 | 3.91 | 4.06 | 3.97 | 3.91 | 4.06 | 3.97 |
| 0.6 | 4.79 | 4.83 | 4.77 | 4.79 | 4.83 | 4.77 |
| 0.7 | 5.62 | 5.58 | 5.56 | 5.62 | 5.58 | 5.56 |
| 0.8 | 6.50 | 6.29 | 6.35 | 6.50 | 6.29 | 6.35 |
| 0.9 | 7.44 | 7.05 | 7.14 | 7.44 | 7.05 | 7.14 |
| 1.0 | 8.28 | 7.78 | 7.93 | 8.28 | 7.78 | 7.93 |
| 1.1 | 9.15 | 8.53 | 8.72 | 9.15 | 8.53 | 8.72 |
| 1.2 | 10.06 | 9.15 | 9.51 | 10.06 | 9.15 | 9.51 |
| 1.3 | 10.69 | 10.01 | 10.30 | 10.69 | 10.01 | 10.30 |
| 1.4 | 11.41 | 10.97 | 11.09 | 11.41 | 10.97 | 11.09 |
| 1.5 | 10.38 | 12.03 | 11.88 | 10.38 | 12.03 | 11.88 |
| 1.6 | 14.21 | 11.77 | 12.66 | 14.21 | 11.77 | 12.66 |
| 1.7 | 14.38 | 13.04 | 13.45 | 14.38 | 13.04 | 13.45 |
| 1.8 | 18.99 | 14.88 | 14.23 | 18.99 | 14.88 | 14.23 |
| 1.9 | 12.28 | 14.02 | 15.01 | 12.28 | 14.02 | 15.01 |
| 2.0 | 20.20 | 16.23 | 15.79 | 20.20 | 16.23 | 15.79 |

options in Table 6–5. Theoretical premiums for options on NYSE stocks are less than premiums for options on NASD stocks. The differences, however, are smaller at low volatility levels than were seen with out-of-the-money options. Black-Scholes baselines (BS) tend to fall in between the premiums of options on NYSE stocks and those of options on NASD stocks. The large discrepancies between the Black-Scholes and conditional distribution premiums that were so evident with the out-of-the-money options are not seen with at-the-money options. The latter finding of less discrepant Black-Scholes premiums makes sense in that general market kurtosis should have the most influence on in-the-money and out-of-the-money options, and the least influence on at-the-money options.

Note that the puts and calls in Table 6–6 have identical premiums. When interest and dividends are ignored (or zero) and all trends are removed, the expected terminal stock price will equal the initial price or, equivalently, the expected return will be zero. In such a case, the premium of a call will have precisely the same amount of time value as that of a put with an identical strike and expiration. This is put-call parity, which holds regardless of the shape of the underlying distribution of returns, just so long as that distribution's expectation is zero. Since at-the-money options have no intrinsic value, only time value, their premiums will be identical, as can be observed in Table 6–6.

***Out-of-the-Money Options, No Detrending*** Although not presented in the table, the theoretical premiums of out-of-the-money options based on raw (not detrended) distributions were also examined. These option premiums reflect not only the shape of the distribution of returns from the stocks on the two exchanges, but also any systematic trends in the prices of these stocks, including trends related to the venue and to the action of volatility.

At levels of volatility less than 60%, premiums for options on NASD stocks are greater than those for options on NYSE stocks, which are, in turn, greater than those calculated with Black-Scholes. The higher premiums derived from the conditional distributions most likely result from the general presence of kurtosis in the marketplace. The pattern is observed for both puts and calls and mirrors the pattern seen with premiums based on the detrended distributions and presented in Table 6–5.

At levels of volatility greater than 70%, calls on NYSE stocks have the highest premiums, followed by calls on NASD stocks, and by Black-Scholes, which again produces the lowest premiums. Compared to the results from the detrended distributions, the extent to which the NYSE premiums exceed the NASD premiums is greater, and the extent to which the NASD premiums exceed the Black-Scholes premiums is much greater. The larger theoretical call premiums discovered with real stocks when trends are not removed are almost certainly due to the volatility payoff effect discussed earlier.

Puts at higher levels of volatility evidence inconsistent premium differences between the venues, and smaller premiums for real stocks (on either venue) than they should have according to the Black-Scholes model. Again, the presence of a volatility-related trend component is responsible for the difference between these results and those seen when detrended distributions were employed in the analysis. The trend component is sufficient (and in the correct direction) to wash out the differences between the venues for puts on higher volatility stocks and to cause Black-Scholes to grossly overvalue these puts. Whether trends should be left in the model is a question for further study on larger data samples extending further back in time.

## Discussion

Theoretical option premiums clearly differ with the trading venue of the underlying stocks. Furthermore, such differences are not merely a result of differences in the bullishness or bearishness of stocks on a given exchange, i.e., the overall trend: they appear even when trend is partialled out by means of statistical manipulations of the distributions. The observed differences in premium across venue or exchange are systematic and large enough to be of concern to traders and hedgers.

It appears that some of the venue effect can be attributed to differences in skew or kurtosis and possibly in volatility. Different venues or exchanges might have stocks that trade with different characteristic levels of skew or, especially, kurtosis. This was not directly tested in the study, but would be easy to examine in any of several ways. Perhaps if skew and kurtosis were partialled out, venue might not affect premium.

Skew and kurtosis were not included as conditioning variables in this study as too many degrees of freedom would have been consumed. However, the inclusion of skew and kurtosis would have made it possible to examine venue for any effect beyond that explained by the other variables. Degrees of freedom should not be a problem in a more constrained regression model, even if nonlinear. Polynomial regressions and neural network models are studied in the next chapter.

As well as skew and kurtosis, volatility might be a factor in the venue effect. For one thing, mean volatility differs greatly between the NYSE and the NASD, with the latter exchange having stocks with much higher average levels. Combined with regression to the mean, these differences in average volatility can distort volatility predictions such as those used in the analysis as the primary conditioning variable. It must be noted that the predictive model for volatility was developed using data from all venues represented in the database—the NYSE, AMEX, and NASD—and might, therefore, have systematic, venue-specific biases built into it.

The venue effect can be incorporated into a pricing model in any of several ways. One way is to develop different models for different venues or exchanges. This is implicitly done when venue is treated as a conditioning variable in a model based upon conditional distributions of returns. The venue variable can also be included in the model used to estimate future volatility. Inclusion of venue in the volatility forecast might be sufficient to eliminate much of the variation in theoretical premiums observed between the venues studied. Inclusion of venue as part of the volatility model would result in the consumption of far fewer degrees of freedom, if it was feasible. Venue, like other variables, can also be handled in the context of a general nonlinear approach such as that involving polynomial, neural, or hybrid pricing models. Finally, venue might be an unnecessary variable to consider were the effects of skew, kurtosis, and possibly other intervening variables accounted for.

## STUDY 6: STOCHASTIC CROSSOVER

It is clear that statistical variables like skew and kurtosis, and such market variables as venue, affect the worth of options, even when trend factors are statistically held constant. But what about so-called "technical" variables, the kinds of indicators used by technical analysts and traders? The indicators that come to mind are moving averages and oscillators of various sorts. One such indicator, an oscillator, is the well-known "Lane's Stochastic."

*Lane's Stochastic* measures where the most recent price of a security falls within the range of prices seen in the recent past. It is called an oscillator because its value fluctuates up and down, quasi-cyclically, within a limited range; in this instance, between 0.0 and 1.0 (or between 0% and 100% as preferred by some traders). Lane's Stochastic is easily calculated for any series of quotes having high, low, and closing prices. Let

$$U_i = \text{highest of } H_{i-m+1}, H_{i-m+2}, \ldots H_i \tag{6.5}$$

be the highest of the last $m$ highs, and

$$L_i = \text{lowest of } L_{i-m+1}, L_{i-m+2}, \ldots L_i \tag{6.6}$$

be the lowest of the $m$ most-recent lows. If $C_i$ is the most recent close, then

$$S_k = \frac{\sum_{i=k-2}^{k}(C_i - L_i)}{\sum_{i=k-2}^{k}(U_i - L_i)} \tag{6.7}$$

is Lane's Stochastic at the $k$-th bar in the data series. Notice the built-in 3-bar smoothing found in the traditional oscillator, also known as the *Fast Stochastic* or *Fast-D*. The *Slow Stochastic*, known as *Slow-D*, is obtained by taking a moving average of the Fast Stochastic.

When using Lane's Stochastic, the technical trader will often pay attention to threshold lines placed at the 20% and 80% levels. These are the traditional oversold and overbought thresholds used with Lane's Stochastic. When the oscillator reading is below 20%, the market is said to be oversold and may be ready for a bounce; when the reading is above 80%, the market or stock is considered overbought and may soon experience a correction. Often the technical trader will establish a long position when the oscillator crosses from below to above the 20% line, and establish a short position on a cross from above 80% to below. Whether one can make a profit doing this is doubtful; it is just too easy, and that means many traders will try to profit by the method, thereby making it impossible for anyone to actually do so, even if the indicator actually has (or once had) some validity. Tests presented in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000) demonstrate that, used as described above, Lane's Stochastic is useless for directional trading. The same statement can be made for almost every other technical indicator, when used in the standard manner.

However, can a technical indicator be of use in the appraisal of options? Although most technical indicators appear to be of little use where trends are concerned, perhaps they can be of value where nontrend variables are involved. In other words, the status of indicators like Lane's Stochastic may be prognostic of near-future volatility, skew, or kurtosis, and other option-relevant variables. Why? It is because technical indicators are generally used by directional traders whose trading efforts drive any inefficiencies (with respect to these indicators) from the market. As far as the authors are aware, technical indicators of the kind under consideration are not popular with those who use options to trade volatility; therefore, any inefficiencies in the options arena that are related to readings on such indicators might have survived and are just waiting to be exploited. It is easy to imagine differences in near-future volatility, skew, and kurtosis that might follow attacks on support or resistance, oscillator threshold crossings, and other events tracked by technical analysts. Of course, whether or not there is anything useful in the technical analyst's arsenal for traders of options is an empirical question. It is this question that the current study begins to address.

## Method

The analysis followed the lines of previous studies and took place in two computational blocks.

***First Computational Block*** In the first computational block, stock returns were binned and the distributions were determined. Two sets of distributions were constructed: one set for stocks with a Slow Stochastic cross above the 20% line at the reference bar; the other for stocks with a cross below the 80% line at the reference bar.

Calculation of the first set of distributions began with the initialization of all elements of the distribution array to zero and with the retrieval of a stock from the database. A valid reference bar was then selected. A reference bar was considered valid when the following conditions held: (1) the price of the stock, uncorrected for splits, was greater than $2; (2) there were at least $m1 + m2$ bars preceding the reference and $n$ bars following it; and (3) that Slow Stochastic had just crossed from below 20% to above. In the analysis, $m1$ was 30, $m2$ was 70, and $n$ was 10—the same as in the previous two studies. The Slow Stochastic, used to evaluate condition (3), was computed with a 9-bar exponential moving average. Once a valid reference bar was located, the immediately preceding $m_1$ bars were used to calculate a short-term historical volatility and the $m2$ bars before that were used for a longer-term volatility. The average range measure was used in these calculations. Once determined, the two measures of historical volatility and the date of the current or reference bar served as inputs to a regression model that provided an estimate of the expected volatility over the future $n$-bar holding period.

The volatility predicted by the regression estimator was used to determine an index into the row of the distribution array. The index was calculated (using Fortran notation) as

```
iv = Int (0.5 + 19 * (v - vmin)/(vmax - vmin))
```

where `v` was the forecast volatility, and where `vmin` and `vmax` defined the range. Hence, `iv` was zero for a volatility level between 0.06 and 0.14, 1 for a volatility level between 0.15 and 0.24, up to 19 for a level between 1.95 and 2.04. An index into the column of the same array was determined from the return from the stock over the holding period. To calculate this array index, the price of the stock at the reference bar was rescaled to a nominal $100. The terminal price (the price $n$ bars after the reference) was then found and the identical rescaling was applied. A 10% gain over the holding period would, therefore, result in a $110 terminal price, and a 5% loss would produce a terminal price of $95. The column index into the distribution array was then computed as

```
ip = Int (0.5 + 99 * (p - pmin)/(pmax - pmin))
```

where `p` was the terminal price as described above, and where `pmin` and `pmax` defined the range.

The two array indices were then checked. If these indices fell into the desired ranges, the value of the array element which they referenced was incremented by one. In this manner, the frequency distributions of terminal price for each level or range of volatility was accumulated. The ranges for which data were accumulated were $0 \le \texttt{iv} \le 19$ (or, equivalently, $0.06 \le \texttt{v} \le 2.04$) for volatility, and $0 \le \texttt{ip} \le 99$ (or, equivalently, $\$19 \le \texttt{p} \le \$251$) for terminal price. These ranges encompassed the vast majority (more than 99.9%) of all data points.

Finally, the next valid reference bar was chosen and, when all reference bars for the selected stock had been processed, the next stock was retrieved from the database. The sequence was repeated until all the data had been analyzed.

Each row in the array generated by the calculations carried out in the first computational block corresponded to a level or range of volatility, and each element (column) of a given row corresponded to a range of terminal price. The actual values of the elements were frequencies of occurrence. Each row of the array can, therefore, be understood to contain a histogram of terminal price frequency for a specified level of volatility. Specifically, the array generated contained, for each possible level of stock volatility, the raw frequency distributions for the terminal stock prices that follow a Slow Stochastic cross over the 20% line occurring at the reference bar.

Once the array was determined for the 20% crossover condition, the analysis was repeated for the 80% crossover. This involved repeating the steps outlined above, but with one alteration: when checking for a valid reference bar, condition (3) now required that the Slow Stochastic had just crossed from above 80% to below. Otherwise, the procedure was identical in every respect. The product was a second array that contained the distributions of terminal stock price for each level of volatility conditional upon the Slow Stochastic having crossed below 80% at the reference bar.

***Second Computational Block*** In this block, theoretical (expectation-based) option premiums were calculated for each level of volatility (0.2 to 2.0 in steps of 0.1), for both indicator patterns (crossed above 20%, crossed below 80%), for the two standard types of options (puts, calls), and for several strikes ($85, $90, . . . $115). Black-Scholes premiums were also determined for each option type, level of volatility, and strike price. In all cases, a nominal $100 stock price was assumed. The Black-Scholes premiums were intended to serve as baselines. All results were written to an intermediate text file. This file was then loaded into an Excel spreadsheet, and the relevant data assembled into several tables, two of which appear in the text below.

The manner in which theoretical option value can be computed from frequency distributions is covered in Chapters 2 and 3. Such computation is also discussed in this Chapter in Study 1. The calculations involved in the statistical removal of the trend component from a distribution are discussed in Study 3, earlier in this Chapter.

## Results

Theoretical premiums for out-of-the-money options (strike = $115 for calls, $85 for puts) on a nominal $100 stock, broken down by option type (put, call), volatility (0.2, 0.3, . . . 1.6), derivation (observed or raw distributions, detrended distributions, Black-Scholes baselines), and Slow Stochastic pattern (PAT1 = crosses above 20% threshold, PAT2 = crosses below 80% threshold), are presented in Table 6–7.

***Out-of-the-Money, Detrended Distributions*** When calculations are performed with the detrended distributions, the premiums of the out-of-the-money options are, with few exceptions, substantially greater with a Stochastic Oscillator that crosses from an oversold level to above 20% than with a Stochastic that crosses from an overbought level to below 80%. The exceptions mentioned only occur for one or two levels of volatility, and are almost certainly due to a small number of cases in the corresponding bins and, hence, large statistical errors of estimation. In other words, out-of-the-money options are worth more when the Stochastic Oscillator has just crossed above the 20% line (Pattern 1); they are worth less (relatively) when the oscillator has just dropped below the 80% line (Pattern 2). Weak markets go with more valuable options, both puts and calls, than do strong markets.

**TABLE 6–7**

Theoretical Premiums for Out-of-the-Money Options as a Function of Oscillator Pattern and Estimated Future Volatility

| | *Out-of-the-Money Calls* | | | | | *Out-of-the-Money Puts* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Observed* | | *Detrended* | | | *Observed* | | *Detrended* | | |
| **VLTY** | **PAT1** | **PAT2** | **PAT1** | **PAT2** | **BS** | **PAT1** | **PAT2** | **PAT1** | **PAT2** | **BS** |
| 0.2 | 0.04 | 0.02 | 0.04 | 0.02 | 0.00 | 0.07 | 0.02 | 0.07 | 0.02 | 0.00 |
| 0.3 | 0.08 | 0.06 | 0.09 | 0.06 | 0.02 | 0.13 | 0.04 | 0.12 | 0.05 | 0.01 |
| 0.4 | 0.24 | 0.21 | 0.25 | 0.20 | 0.14 | 0.29 | 0.12 | 0.29 | 0.14 | 0.06 |
| 0.5 | 0.51 | 0.46 | 0.51 | 0.40 | 0.39 | 0.44 | 0.23 | 0.44 | 0.27 | 0.20 |
| 0.6 | 1.08 | 1.00 | 1.10 | 0.82 | 0.76 | 0.84 | 0.35 | 0.83 | 0.47 | 0.44 |
| 0.7 | 1.59 | 1.58 | 1.54 | 1.26 | 1.24 | 1.00 | 0.55 | 1.05 | 0.78 | 0.77 |
| 0.8 | 1.95 | 2.54 | 2.02 | 2.00 | 1.79 | 1.59 | 0.71 | 1.52 | 1.13 | 1.18 |
| 0.9 | 2.61 | 2.93 | 2.54 | 2.50 | 2.39 | 1.83 | 1.13 | 1.90 | 1.51 | 1.64 |
| 1.0 | 4.14 | 3.45 | 3.53 | 2.88 | 3.04 | 2.19 | 1.37 | 2.71 | 1.89 | 2.14 |
| 1.1 | 3.93 | 3.50 | 3.71 | 3.27 | 3.71 | 2.47 | 2.33 | 2.72 | 2.56 | 2.68 |
| 1.2 | 4.99 | 3.50 | 4.77 | 3.66 | 4.42 | 3.46 | 1.76 | 3.71 | 1.60 | 3.24 |
| 1.3 | 6.26 | 6.33 | 5.36 | 5.76 | 5.14 | 3.21 | 3.87 | 4.13 | 4.50 | 3.83 |
| 1.4 | 5.04 | 6.53 | 4.68 | 3.28 | 5.88 | 4.13 | 2.32 | 4.48 | 4.21 | 4.44 |
| 1.5 | 7.44 | 4.18 | 6.33 | 4.98 | 6.63 | 4.63 | 4.31 | 5.74 | 3.08 | 5.06 |
| 1.6 | 7.36 | 4.13 | 6.68 | 4.53 | 7.39 | 4.25 | 6.09 | 5.16 | 5.69 | 5.69 |

How great are the differences in option value between the two Oscillator patterns? Quite sizable, according to the data in Table 6–7, and certainly great enough to be of interest to options traders. As an example, consider a stock with a volatility of 60%, annualized. If this stock has just had its Stochastic reading move up from oversold territory, an out-of-the-money call with a strike of $115 would theoretically be worth $1.10; the same call on a stock exhibiting the same volatility, but having a Stochastic Oscillator reading that has just fallen below the overbought threshold, should be worth only $0.82. Black-Scholes prices the call at $0.76. A put on the stock will have an even greater variation in value across the conditions: $0.83, $0.47, and $0.44, for Pattern 1, Pattern 2, and Black-Scholes, respectively. These differences cannot be ignored; they could easily be exploited for a profit by even a retail trader.

For out-of-the-money calls, Black-Scholes premiums are consistently smaller than the lower Pattern 2 premiums when volatility is less than 90%. The same applies to out-of-the-money puts when volatility is less than 70%. Above these levels of volatility, Black-Scholes yields premiums that are sometimes below the lowest of the two patterns, and sometimes between the two pattern levels, depending on volatility.

***Out-of-the-Money, Raw Distributions*** The differences in theoretical premiums between the oscillator patterns are similar to those observed with the detrended distributions when the raw data are examined; however, the extent to which the Pattern 1 premiums exceed the Pattern 2 premiums has been magnified for the puts and attenuated for the calls. The attenuation of the Stochastic Oscillator effect for the calls makes the observed differences in theoretical premium less consistent over volatility levels.

When comparisons are made to Black-Scholes, the conditional distribution premiums (regardless of oscillator pattern) are large when volatility is less than or equal to 100%. For Pattern 1, they are much larger with the raw conditional distributions than with either the detrended distributions or with Black-Scholes. In other words, it is with Pattern 1 that the volatility payoff or trend effect really appears in the results.

Pattern 2 seems to eliminate this volatility-induced trend, even at high volatility levels—the premiums do not grow much larger than the Black-Scholes baseline. The volatility effect does attenuate put premiums at higher volatility levels. This reminds the authors of how cheap OEX puts were at the top just prior to the 1987 crash. Volatility was high and the Stochastic Oscillator was in overbought territory and had begun to drop below the 80% line. The puts were behaving consistent with the findings of this study. Yet, in fact, the puts had great value in that particular instance. However, statistically, puts do have less relative worth when volatility is high, and the actual market prices at the time reflected this statistical tendency. The deviations from Black-Scholes are immense for some levels of volatility and pattern. The actual worth, as determined by trends and all, of an out-of-the-money put when the Stochastic Oscillator has crossed below 80% and the volatility is 90% is $1.13; Black-Scholes prices the same put at $1.64, over 45% greater in value! For Pattern 1, calls also evidence much greater value at most levels of volatility than Black-Scholes would suggest.

***At-the-Money Options*** The first thing to note when examining the theoretical premiums for at-the-money options (see Table 6–8), as determined from the detrended distributions, is that they are identical for both puts and calls. This reflects the fact that puts and calls of the same strike and expiration have the same time value when there are no trends in the data, and when dividends and interest are ignored, as in the present instance. Since at-the-money options have no intrinsic value, and since the options have the same strike, they have the same total premium.

At all levels of volatility, there is a very consistent pattern of Pattern 1 having greater premiums for both puts and calls than Pattern 2. Black-Scholes tends to yield premiums that fall between those of the two patterns. Many of the deviations in premium that are observed seem tradable.

Although Stochastic Oscillator patterns were found useless for trend forecasting and directional trading in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000), the use of the oscillator was not examined there in the context of varying levels of volatility. It seems that Pattern 1 may be

**TABLE 6–8**

Theoretical Premiums for At-of-the-Money Options as a Function of Oscillator Pattern and Estimated Future Volatility

| | *At-the-Money Calls* | | | | | *At-the-Money Puts* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Observed* | | *Detrended* | | | *Observed* | | *Detrended* | | |
| **VLTY** | **PAT1** | **PAT2** | **PAT1** | **PAT2** | **BS** | **PAT1** | **PAT2** | **PAT1** | **PAT2** | **BS** |
| 0.2 | 2.02 | 2.00 | 2.00 | 1.83 | 1.59 | 1.99 | 1.66 | 2.00 | 1.83 | 1.59 |
| 0.3 | 2.40 | 2.35 | 2.63 | 2.21 | 2.38 | 2.89 | 2.07 | 2.63 | 2.21 | 2.38 |
| 0.4 | 3.37 | 3.19 | 3.42 | 2.95 | 3.18 | 3.48 | 2.72 | 3.42 | 2.95 | 3.18 |
| 0.5 | 4.07 | 4.17 | 4.08 | 3.67 | 3.97 | 4.10 | 3.20 | 4.08 | 3.67 | 3.97 |
| 0.6 | 5.05 | 5.35 | 5.11 | 4.50 | 4.77 | 5.18 | 3.66 | 5.11 | 4.50 | 4.77 |
| 0.7 | 5.99 | 6.47 | 5.85 | 5.35 | 5.56 | 5.69 | 4.24 | 5.85 | 5.35 | 5.56 |
| 0.8 | 6.42 | 7.86 | 6.61 | 6.29 | 6.35 | 6.85 | 4.70 | 6.61 | 6.29 | 6.35 |
| 0.9 | 7.41 | 8.06 | 7.25 | 6.94 | 7.14 | 7.06 | 5.74 | 7.25 | 6.94 | 7.14 |
| 1.0 | 9.69 | 8.86 | 8.44 | 7.55 | 7.93 | 7.14 | 6.09 | 8.44 | 7.55 | 7.93 |
| 1.1 | 9.02 | 8.78 | 8.57 | 8.29 | 8.72 | 8.04 | 7.79 | 8.57 | 8.29 | 8.72 |
| 1.2 | 10.51 | 7.54 | 10.14 | 7.86 | 9.51 | 9.69 | 8.40 | 10.14 | 7.86 | 9.51 |
| 1.3 | 12.05 | 11.54 | 10.47 | 10.46 | 10.30 | 8.80 | 9.15 | 10.47 | 10.46 | 10.30 |
| 1.4 | 11.16 | 16.21 | 10.54 | 8.40 | 11.09 | 9.92 | 4.99 | 10.54 | 8.40 | 11.09 |
| 1.5 | 13.96 | 8.29 | 12.11 | 9.59 | 11.88 | 10.40 | 11.84 | 12.11 | 9.59 | 11.88 |
| 1.6 | 13.08 | 9.66 | 11.73 | 10.45 | 12.66 | 10.33 | 11.25 | 11.73 | 10.45 | 12.66 |

associated with some downward trend at low levels of volatility, and with a definite upward trend (evidenced by higher call prices) at high volatility levels—an apparent interaction between Stochastic pattern and the volatility payoff. If these trends are stable and appear in other samples, it would suggest that these Oscillator patterns could be used at high levels of volatility for directional trading, especially with options. For example, given a stock with a volatility of 100%, the presence of Pattern 1, and an at-the-money call with a market price that is close to what Black-Scholes indicates is the fair price, $7.93, then purchasing the call might lead to a profitable trade—the call is theoretically worth over $9 based on future expectation, as computed from the conditional distributions.

## Discussion

There is no doubt that at least one technical indicator, the Stochastic Oscillator, provides information that is useful in the appraisal of options. In general, both puts and calls have greater value when the Stochastic Oscillator crosses above 20% than when it crosses below 80% and when trends are assumed to be absent. The same pattern applies, although with somewhat less consistency, when trends are present, i.e., when they have not been statistically removed. It seems that, with certain combinations of volatility and oscillator pattern, the trend (volatility-induced, perhaps) can exert a large influence on the worth of an option.

Other technical patterns, such as those occurring when a stock's price approaches recent levels of support or resistance, when moving averages cross, and so on, should be investigated. Even though these patterns might not be very useful for directional traders speculating on the stock, they might provide a significant edge for the options trader betting on volatility or kurtosis. Technical factors may even be useful for trend forecasting when other variables are included in the model, e.g., volatility. More importantly, technical indicators may be prognostic of the shape of the distribution of future stock price movements and, therefore, of the value of options trading on the affected stocks, even if they cannot forecast direction.

## SUMMARY

It was demonstrated in this chapter how the conditional distribution methodology could be used to let the market reveal the theoretical fair value of an option. The method of conditional distributions has a number of desirable properties. For one thing, the method implicitly takes into account the actual price distributions that occur in the market, rather than distributions that are *assumed* to accurately describe the behavior of stock prices, but which may do so inaccurately. When using conditional distributions to price options, the fact that the distribution of returns from real stocks is not log-normal presents no problem. The use of conditional distributions also automatically corrects for such things as regression to the mean in volatility and in other critical model inputs. Moreover, the impact of variables like skew, kurtosis, venue, and even the status of technical indicators are easily assessed within the context of a pricing approach based on empirically derived stock price distributions. The conditional distribution methodology used in this chapter also allows theoretical modifications and adjustments, such as detrending or incorporation of the effects of well-understood variables (like the rate of risk-free interest), to be easily accomplished. Finally, the method is transparent: the investigator can easily examine the shape of a distribution, the effect of a conditioning variable, and so on. This is in contrast, e.g., to neural network pricing models, which are studied in the next chapter. All in all, the method of conditional distributions was found to be flexible and powerful, especially when combined with the smoothing and interpolation methods discussed later in this book. In addition, the method provides the ability to easily examine the impact of different variables and conditioning factors on the value of an option; this is essential for the trader or hedger attempting to gain an edge in the marketplace.

Despite its power and flexibility, the method of conditional distributions is practical, fast, and easily implemented on modern computers. A complete set of conditional distributions can be stored in an array or tensor, and used to price a wide variety of options. The calculations involved can be carried out quickly and efficiently, given some attention to the details of coding.

However, there is a demon that plagues the conditional distribution methodology. This is the degrees of freedom demon that enters the picture whenever statistical estimation based on real data samples is involved. The very flexibility of the approach, and the lack of constraints on the model's functional form, implies an avaricious consumption of degrees of freedom. Add too many conditioning variables to the model, and the estimation error grows like Jack's proverbial beanstalk. This is the problem of combinatorial explosion.

The degrees of freedom demon can be tamed in various ways. One way is to employ some kind of simple data smoothing. Polynomial smoothing was actually employed in many of the charts presented in this and earlier chapters. Another way to tame the demon is through the use of multivariate nonlinear modeling techniques, such as general polynomial modeling or neural networks. The use of polynomial models and neural networks, both to enhance the conditional distribution methodology and to directly price options, is examined in the next chapter.

The study of option appraisals based on conditional distributions can reveal much about the markets. Consider the studies in this chapter. They demonstrate that there is a lot to learn about the effects of different variables on theoretical option premiums and that such investigations are not merely of academic interest but can benefit the active options trader. For example, the first study clearly demonstrated the effect of regression to the mean when standard historical volatility was used in pricing options. The second study revealed that the so-called "volatility payoff," seen in stock returns, also affects fair premium for options at high levels of volatility. In the third study, the use of detrending was illustrated and, with the effect of a systematic trend-volatility relationship removed, the effect on option premiums of the general kurtosis in the market as a whole could be observed. Kurtosis was demonstrated to raise the observed premiums for both puts and calls at low volatility levels. The third study directly examined the use of historical measures of skew and kurtosis in the pricing of options. This study demonstrated that, just as with historical volatility, historical skew and historical kurtosis were related to fair premium and could thus contribute to a pricing model. Again, the effects were large enough to be of concern to those who use options to either speculate or hedge. The fifth and sixth studies demonstrated that fair option premiums were dependent upon exchange or trading venue and on the status of technical trading indicators, respectively. The effect of trading venue might reduce to differences in skew or kurtosis, and may not be essential as an input to a model that already incorporates the higher moments; however, the Stochastic Oscillator patterns (examined in Study 6) do contribute information that is independent of skew or kurtosis, at least of the historical variety, and may have some prognostic usefulness with respect to volatility and volatility-trend interactions. This study, at the very least, makes it evident that other technical variables should be explored as possible contributors to an advanced pricing model.

There are several issues raised in this chapter that need further exploration. One is how to best handle trends. Should trends be removed from the data sample? Whether or not to remove trends or to leave them in is an important question for further study. The removal of trends implies a theoretical stance; that any trends found are merely a sample artifact, and are not to be considered a real and consistent feature of stock price movements. Larger and more varied samples need to be studied. If trends, including those caused by volatility (as in the volatility payoff), really exist in the marketplace, then they should probably not be removed (or specific kinds of trends should be reinserted, e.g., the volatility payoff) when pricing options for speculative purposes.

Another issue is how to best tame the degrees of freedom demon. Perhaps the distributions of returns on which option appraisals are based can be described compactly in terms of just a handful of coefficients—for instance, the first several moments. If this is the case, then instead of having to estimate an entire distribution, only a small number of parameters would need to be determined for each set of conditions of interest. Estimation of a few coefficients consumes far fewer degrees of freedom than does estimation of a complete distribution. The use of polynomial or neural regressions might also be a way to obtain a good fit with fewer free parameters and, hence, to conserve degrees of freedom while reducing combinatorial explosion. Some empirical work needs to be done to discover the best approach with regard to the degrees of freedom and other estimation issues.

The impact of each of a wide range of possibly relevant variables on theoretical fair premium also needs to be explored. Finding relevant and predictive variables can quickly pay off in better option models. Better models should reveal more inefficiencies that the astute trader can exploit for a profit.

Finally, the prices of real options need to be examined in order to see how they compare to the theoretical fair prices found by various pricing models. So far, the actual market prices of options have not been examined. Do such prices reflect the value indicated by models such as those analyzed in this chapter? Do they reflect Black-Scholes and Cox-Ross-Rubinstein? Or do they reflect something in between? The low put prices at the top before the 1987 crash imply that traders respond to the correct probabilities and statistical expectation, as calculated here, and not to Black-Scholes estimates. Does this mean that there is no opportunity to find inefficiencies of the kind unveiled by the models used herein? Are option prices in the market consistent with these models rather than with the more popular ones? Do real option prices reflect phenomena like regression to the mean? These questions are studied when actual market prices for options are analyzed in a later chapter. The answers to questions such as these have direct bearing on whether models like those considered in this chapter can highlight inefficiencies that can be exploited for a profit.

## SUGGESTED READING

Basic coverage of histogram analysis and binning procedures can be found in *Statistics* (Hays, 1963). The approach to option pricing used in this chapter was first discussed by the authors in “Market Realities and Option Pricing” and “More Intelligent Option Pricing” in *Futures* (Katz and McCormick, 2001b, c). An analysis of Lane’s Stochastic and other technical indicators can be found in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000).

CHAPTER 7

# Neural Networks, Polynomial Regressions, and Hybrid Pricing Models

The problem involved in the development of an option pricing model is primarily one of constructing a continuous, nonlinear function that maps a set of input or independent variables, perhaps represented as a vector, to an output that approximates the theoretical fair value of an option.

## CONTINUOUS, NONLINEAR FUNCTIONS

What is a continuous, nonlinear function? A nonlinear function can be defined by first considering the nature of a linear function. Any function f that satisfies

$$\mathrm{f}(a\mathbf{x}) = a\mathrm{f}(\mathbf{x}) \tag{7.1}$$

and

$$\mathrm{f}(\mathbf{x} + \mathbf{y}) = \mathrm{f}(\mathbf{x}) + \mathrm{f}(\mathbf{y}) \tag{7.2}$$

where $a$ is a scalar quantity, and $\mathbf{x}$ and $\mathbf{y}$ are vectors or functions, is *linear*. In two dimensions, a linear relationship can be represented by a line; in three dimensions, by a plane; and in any finite number of dimensions greater than three, a hyperplane. A *nonlinear* function is one that does not satisfy Equations 7.1 and 7.2, and that, even when it is of finite dimension, cannot be spatially represented by a line, plane, or hyperplane.



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

An example of a linear function is the standard regression estimator

$$\hat{\mathbf{y}} = \mathbf{xa} \tag{7.3}$$

in which $\hat{\mathbf{y}}$ is the column vector containing the estimates or predictions, $\mathbf{x}$ is the data matrix on which the estimates are based, and $\mathbf{a}$ is the column vector of regression weights. In a standard regression model, the columns of $\mathbf{x}$ correspond to the model input variables (including, perhaps, an intercept) and the rows correspond to the cases or data points. A regression model may be construed as a line, plane, or hyperplane that has been fit, in a least-squares sense, to a specified set of data.

An example of a nonlinear function is the square root taken over the domain of positive real numbers. A plot of the square root appears, not as a straight line, but as a curve. Another nonlinear function is one close to the heart of almost every options enthusiast. Black-Scholes maps volatility, and a host of other inputs (also known in functionland as *arguments*), to a theoretical fair price. To represent a nonlinear function like Black-Scholes spatially requires a curved, multidimensional surface.

It is obvious that an option pricing model requires a nonlinear function. Just look at the charts sprinkled throughout this book in which theoretical premium is shown as a function of volatility or of some other influential variable. In all cases, except for at-the-money options, the relationship between premium and volatility is curvilinear. Stock price and strike always have a curvilinear relationship to fair premium. Only interest rate has an influence on premium that is close to linear, at least when examined visually.

By now it should be clear what is meant when a function is described as nonlinear or curvilinear. But what about continuity? What does it mean to say that some function is continuous over some domain?

Roughly speaking, a *continuous* function is one for which the change in its output decreases with decreasing change in its input. All linear functions are continuous. The square root is continuous over the domain of positive real numbers. Black-Scholes, too, defines a continuous mapping when volatility, time left, stock price, and strike are all strictly greater than zero, and when risk-free interest is greater than or equal to zero. The Heaviside Function, H($x$), which takes on the value of 0 for all $x \leq 0$ and 1 for all $x > 0$, is *discontinuous*; it has a discontinuity at 0. When plotted, continuous functions do not have jumps, gaps, or breaks; discontinuous functions do.

It is logical that an option pricing model must be defined by a continuous function. It would not make much sense for an infinitesimal change in, say, the volatility entered into a model to result in a massive change in the model's output, the theoretical price of the option. At any point in the volatility spectrum, the less volatility is perturbed, the smaller the change in premium that should result. The same can be said for all other inputs, treated individually or as a vector. When plotted, curves that describe the relationships between various influential variables and fair option value appear to be (and probably should be) smooth and unbroken. Within the limits of sampling or estimation error, such smoothness can be seen in the various charts that appear throughout this book. This substantiates the claim that a continuous function underlies the relationship between option prices and most of the variables that affect them.

Continuity has a lot to do with a particular kind of smoothness. Option models need to be continuous, but one can rationally demand more: that they be continuous in their first and second derivatives, except at expiration (when time left is zero). The reasoning here is that not only should an option model be "smooth" in relating its inputs to its output, but that it should also have Greeks that are smoothly related to its inputs. The reader may recall that the Greeks are derivatives (in the calculus sense of the word) of fair premium (the model's output) with respect to various inputs (e.g., time left, stock price, and volatility). Perhaps even more should be demanded: that *all* derivatives be continuous. Continuity of derivatives of all orders is associated with a special kind of smoothness. A function that is continuous in all of its derivatives is said to be *analytic* and can be expanded at any point in its domain as a locally convergent power series. In small regions of input space, such functions can be well approximated by polynomials—an important fact in the context of what follows.

Given the above discussion, the nature of a continuous, nonlinear function should be clear. It should also be clear that what is required for pricing options is a continuous, nonlinear function that has as inputs variables like volatility, time left, stock price, and strike, and that has the expected (in a statistical sense) terminal price of the option—its fair value—as an output. But how can such a function be constructed?

## CONSTRUCTION OF A PRICING FUNCTION

One way to construct a pricing function is to derive it purely from theoretical considerations. Most extant option pricing models have been constructed in this manner. The Black-Scholes and Cox-Ross-Rubinstein models are good examples of pricing functions based on theoretical analyses.

But what if one wants to be heretical and to do it empirically by fitting, in a least-squares sense, an appropriate kind of function to the data? The first thing that comes to mind is regression. Standard regression, however, whether univariate or multivariate, is linear; it fits a line, plane, or hyperplane to the data. This will not do; the relationship between expected terminal option premium and the various input variables is decidedly nonlinear and a linear approximation, such as that provided by standard regression, will give a poor fit to the data. So what is one to do? The answer lies in general nonlinear models.

There are general-purpose data modeling techniques that are designed to solve problems that involve continuous, nonlinear mappings. The two most popular—and, in this case, most relevant—are neural networks and polynomial regressions.

## POLYNOMIAL REGRESSION MODELS

The basic polynomial regression involves solving for the coefficients in an equation, one that contains terms which represent the various powers and possibly cross-products of the input variables. In the univariate case, when there is only one input, a polynomial regression takes the form of

$$\hat{y} = \mathrm{f}(x) = \sum_{i=0}^{n} a_i x^i \tag{7.4}$$

where $\hat{y}$ is the predicted value of the dependent variable, $n$ is the order of the polynomial model, $x$ is the input or independent variable, and $a_0, a_1, \ldots a_n$ are the regression coefficients. Naturally, the $a_i$ are chosen to minimize $\|y - \hat{y}\|^2$, the sum of the squared prediction or model errors. In the bivariate (two independent variable) case, a polynomial regression model appears as

$$\hat{y} = \sum_{k=0}^{n} \left[ \sum_{i=0}^{k} a_{i,k-i} x_1^i x_2^{k-i} \right] \tag{7.5}$$

where $x_1$ and $x_2$ are the two inputs to the model, $n$ is the order of the polynomial, and $a_{ij}$ is the coefficient for the $x_1^i x_2^j$ term. Clarity might be greater for the second-order bivariate model ($n = 2$), which can easily be expanded as

$$\hat{y} = a_{0,0} + a_{1,0} x_1 + a_{0,1} x_2 + a_{1,1} x_1 x_2 + a_{2,0} x_1^2 + a_{0,2} x_2^2 \tag{7.6}$$

The expansion makes it easy to see the intercept term, the linear terms, the interaction term, and the squared (second-order) terms.

It may come as a surprise that the coefficients (the $a_i$ or $a_{ij}$) in the above models are, in fact, ordinary regression weights. By preparing a data matrix with various powers and cross-products of the independent variables, an ordinary linear regression package can be used to find a set of coefficients that minimize the sum of the squared errors. Polynomial models, like those defined here, can be solved using ordinary least-squares regression techniques because, although clearly nonlinear with respect to the independent variables, these models are definitely linear in their coefficients.

Although it is perfectly sound, mathematically speaking, to perform a polynomial regression by the application of standard regression methodology to data consisting of various powers and cross-products of the independent variables, this is almost never the best way to proceed. Polynomial models, especially those of a higher order that involve many terms, are almost certain to exhibit severe collinearity. Solving directly for the coefficients may be impossible or may lead to a near-total loss of precision due to the accumulation of roundoff errors that occurs when working with ill-conditioned matrices in a finite-precision environment (in other words, on a computer). There are many ways, however, to finesse a solution without encountering a serious loss of precision or a computer crash. One exemplary way is to replace the simple powers of the independent variables with Chebyshev Polynomials. The basic idea is to orthogonalize the various terms in the model to as great an extent as is possible and to thereby avoid the problems of collinearity and ill-conditioning that might otherwise arise. These issues, however, are more computational than theoretical in that a regression package with infinite precision arithmetic applied to a table containing simple powers and cross-products of the input variables would produce results identical to those of the more sophisticated methods.

Some problems respond very well to solutions in terms of polynomial approximations. Over a small enough region of their domain or input space, continuous, nonlinear functions with continuous derivatives can often be well-approximated even by low-order polynomials. To get a tight fit over a larger region of a function's domain may require polynomial models of a higher order.

The approximation of functions by polynomials is a venerable practice and polynomial regression models are employed in many fields of science. The ability to fit a basic polynomial to a set of data points is even built into the charting feature of Microsoft's Excel spreadsheet program where it is often used for the purpose of smoothing and thus making more apparent any relationship that exists in the graphed data. At one time in the early 1990s, there were even specialized software products that offered composite polynomial models as an alternative to neural networks.

The reader has already encountered a variety of polynomial regression models in previous chapters. Simple polynomial regressions involving a single independent variable were used to smooth data in many of the charts. In Study 1 of Chapter 5, e.g., a simple second-order polynomial was found to describe effectively the relationship between historical and future volatility. A more complex example can be found in the multivariate polynomial regression model developed in Study 4 of Chapter 5 for the prediction of near-term future volatility. In the latter model, long- and short-term historical volatility measures were among the inputs. These inputs were analyzed in terms of a polynomial having linear, second-order, and cross-product terms. The model produced excellent results. No doubt, polynomial models can do a fine job of approximating a variety of nonlinear functions or relationships.

The main problem with polynomial modeling is not that of achieving a good fit to the data; rather, it is the irksome, yet familiar, problem of undesirable curve-fitting and excessive consumption of available degrees of freedom. Nevertheless, a good polynomial model may require far fewer degrees of freedom, and be less subject to combinatorial explosion in the degrees of freedom consumed with increasing numbers of variables, than a less constrained model, such as the one implicit in the use of conditional distributions. At least with standard regression, and by extension with polynomial models, there are statistics to help the developer assess the extent to which curve-fitting and an excessive consumption of degrees of freedom are problems in any given situation. One nice statistic useful in this context is the prediction sum of squares (PRESS) statistic, which is fully described in *Classical and Modern Regression with Applications* (Myers, 1986).

## NEURAL NETWORK MODELS

Neural network technology arose from attempts to emulate the behavior of neural tissue in living systems by implementing, on the computer, structures composed of simulated neurons and synapses (neuronal interconnections). Neural network technology endeavors to emulate the kind of information processing and decision making that presumably occurs in living organisms.

Neural networks come in a variety of flavors, depending on their *architecture*, i.e., the particular ways in which the simulated neurons are interconnected and the internal behavior of these simulated neurons, i.e., their signal processing behavior, *transfer functions*, and learning rules. The most popular kind of neural network, and the most important in the current context, is the *multilayer*, *feed-forward perceptron*. There are many other kinds of neural network configurations, including learning vector quantization (LVQ), Kohonen, as well as various adaptive resonance and recurrent models.

Neural architectures differ in the way they learn: some networks employ "supervised learning" and others "unsupervised learning." *Supervised learning* is when the network is shown something and guided to produce a correct solution by being shown that solution. In other words, a kind of *paired-associate learning* is in operation: there are inputs that the network sees and a desired output for every set of inputs; for every set of inputs, the task of the network is to learn to produce the desired outputs. In contrast, *unsupervised learning* involves a network that simply takes presentations of inputs and learns to organize them into patterns as the network sees fit. If an analogy is made to statistics, unsupervised networks would be more akin to cluster and factor models, whereas the supervised networks would be closer to various forms of regression and discriminant analysis, albeit in nonlinear form.

Research on neural networks began, on a theoretical level, in the 1940s. At that time, however, computer technology to adequately implement the theory was not available. Around the time when computer technology had become sophisticated enough to accommodate neural network research, Minsky and Papert, in their book *Perceptrons* (1969), brought such research to an abrupt standstill: They "proved" that a special kind of two-layer neural network could not, in any way, solve the *exclusive or* problem; this was enough, given their status as MIT professors, to discourage further study of the subject for many years. The field did not recover from that blow until a form of gradient descent optimization (*back-propagation* in neuro-speak) was applied to finding the connection weights in neural networks containing more than two layers and employing sigmoid transfer functions. Since three-layer networks can readily solve the *exclusive or* problem (as well as many others), the objections that Minsky and Papert expressed were rendered irrelevant and research began again, in earnest.

## Feed-Forward Networks

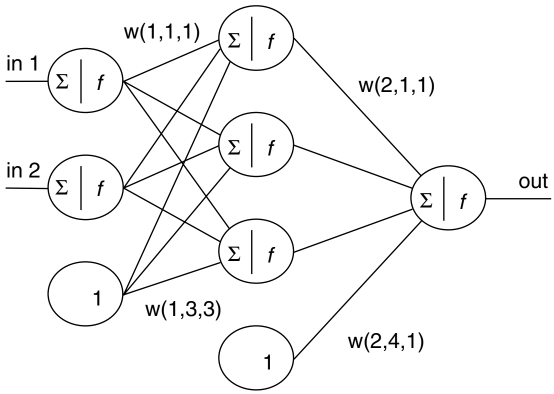
A *feed-forward neural network* is composed of neurons arranged in layers, like the nodes in a binomial pricing tree. The *input layer* is the first layer found in such a network: from the outside world, it receives data (inputs or independent variables) from which some inference is to be drawn or a prediction is to be made. This layer is massively connected to the next layer, which is usually called a *hidden layer* since it has no connections to the outside world. The hidden layer's outputs are, in turn, fed to the next layer, which may be another hidden layer, in which case the process repeats, or it may be the *output layer*. Each neuron in the output layer produces an output from the network. There may be one or more outputs. The outputs constitute the inferences, predictions, estimations, or decisions made by the network.

Networks can vary in size from only a few input variables to hundreds or thousands of them and from only three layers to dozens. The size usually depends on the nature and complexity of the problem at hand. In almost all cases, a three-layer or four-layer network (i.e., a network with one or two hidden layers) will perform well; exceeding four layers is only rarely of benefit and will add to the complexity and training time. In general, the layer that immediately follows the input layer will have anywhere from two to three neurons, to perhaps 10 or 20 times the number of input neurons, depending on the problem. In a four-layer network, the layer that immediately precedes the output neuron may have anywhere from two or three neurons, to perhaps ten times the number of output neurons. These numbers may vary, depending on the problem.

Figure 7–1 depicts the structure of a typical three-layer, feed-forward neural network that contains two active input or first-layer neurons, three active neurons in the second (middle) layer, and one output neuron. In the figure, neurons are depicted as circles and each column represents a layer. There are two kinds of neurons: active neurons and threshold, or bias, neurons. Within each *active neuron*, summation takes place; i.e., all stimuli targeted at that neuron are added together. The summation process is symbolized by the $\sum$ inside the circle that defines each active neuron. The *transfer function*, symbolized by $f$, is then applied; it defines how the neuron responds to varying levels of stimulation coming from its inputs. *Threshold neurons* work differently; they supply a constant output that acts much like an intercept in a regression model. These neurons are inactive in

**FIGURE 7–1**

A Three-Layer Feed-Forward Neural Network with Two Inputs, Three Active Hidden-Layer Neurons, and One Output




the sense that they do not receive or respond to any input signals. They are depicted in the figure as circles containing the numeral 1. The outputs from the neurons in any layer are fed via interconnections that vary in strength and polarity (as specified by connection weights) to neurons in the next layer. Neurons in each layer of a standard feed-forward network are massively connected to those in the next. These connections are shown in Figure 7–1 as lines, each of which is associated with a coefficient or *connection weight*. A few connections weights—symbolized by w($i$,$j$,$k$), where $i$ is the source layer index, $j$ is the source neuron index, and $k$ is the destination neuron index—are also on display in the diagram. Larger neural networks, and those that possess more than three layers, have the same general architecture as the small three-layer network just examined. Although neural network technology is readily available, many tricks are required to use it successfully. Neural networks learn from past experience. The system developer must therefore take the role of a teacher and provide the network with adequate training examples; i.e., he or she must provide an adequate sample of past experiences from which the system can learn. One difficulty in developing a neural network is finding and effectively "massaging" the data into training examples or *facts* that highlight meaningful patterns so that the neural network can learn efficiently and not become confused or get put astray.

As previously mentioned, training a neural network involves repeatedly presenting pairs of inputs and desired outputs (facts) to the network. This is done so that the network can learn, based on many examples, how to make accurate decisions, predictions, estimations, or classifications.

During the training process, the *connection weights*, which define the strength and polarity of the connections between the layers of the network, are repeatedly adjusted in order to maximize the correspondence between the network's actual outputs and the *targets* (desired outputs). The *back-propagation learning algorithm*, a kind of gradient descent minimization, is the most frequently used approach for determining how the network is to adjust the weights in response to a training fact. It is through such adjustments to the connection strengths, repeated again and again, that learning occurs.

A new neural network is typically initialized with a random set of weights; training then begins. The network engages in a kind of hypothesis testing in which it makes guesses about the targets or desired outputs. It does this based on the data just received and the theories, or "constructs," about the data that it has thus far formed. If a guess is wrong, the learning algorithm adjusts the weights in such a way as to make the network's output better agree with the target the next time around. Training then moves on to the next fact, which consists of input data for the network and a target against which the network's output can be compared. Training continues as additional facts are fed to the network and corrections made to the weights, and hence to the network's internal construction of its world. This process is repeated as long as the network trains. As training proceeds, the correspondence between the network's outputs and the targets usually improves and so the quality of the decisions or forecasts increases. Finally, there comes a point when additional training yields little or no improvement in the results. When this happens, *convergence* is said to have occurred. Training is then terminated and the network is examined to evaluate its performance on the data on which it was trained, as well as on data it has never seen before (the latter to discover whether the network can *generalize* or successfully apply what it has learned to new situations). If the training is judged to have been successful and the network is producing accurate classifications, decisions, estimates, or forecasts, then the work is done.

Neural networks are plagued by the demon known as *curve-fitting*. Curve-fitting, also referred to as *over-fitting*, has a lot to do with the number of degrees of freedom that are available when fitting a model to a set of data points. Consider multivariate regression: If the number of independent variables (including the intercept) is equal to the number of cases to which the model is fitted, a perfect squared multiple correlation will result, even when the data are completely random! This is an extreme example of curve-fitting. Excessive curve-fitting occurs when the number of free parameters (connection weights) is too large relative to the number of facts or cases. A disproportionate multiple correlation results, which then shrinks (often dramatically) when tested on a second, independent sample.

There is a formula to correct for *shrinkage* in a multiple correlation that is obtained from a standard multivariate regression. Here it is:

$$r_c = \sqrt{1 - \frac{n-1}{n-p}(1 - r^2)} \tag{7.7}$$

In this equation, $r_c$ is the multiple correlation corrected for shrinkage, $n$ is the total number of data points or facts to which the model has been fit (or trained, in neuro-speak), $p$ is the number of free parameters in the model, and $r$ is the observed (uncorrected) multiple correlation in the fitting sample. The way to avoid excessive curve-fitting or shrinkage in a multiple regression is to ensure that there are a large number of data points, or facts, relative to the number of free parameters in the model. With neural networks, the same logic applies. The secret to avoiding an excessively curve-fit solution is to minimize the number of free parameters and to train on a sufficiently large set of facts. The fewer the free parameters and greater the number of facts, the less the undesirable curve-fitting that will be experienced. The connection weights in a neural network are akin to the weights in a multivariate regression. A feed-forward neural network model is, in fact, a specific form of nonlinear multiple regression. As with standard (linear) multiple regression, the same shrinkage correction formula works acceptably well to estimate what the correlation will be between a neural network's output and the target in an independent sample (one not used for training), given the correlation achieved in the training sample. The shrinkage correction formula—which, by the way, may also be used in the context of multivariate polynomial regressions—can help the model developer avoid excessive curve-fitting. Its use in neural network development is thoroughly discussed and illustrated in Chapter 11 of *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000).

When properly trained, neural networks make admirable pattern recognizers. They have been successfully applied to a wide range of problems. Neural networks assess credit risk, recognize written and spoken words, spot tumors in mammograms, process sonar and radar signals, filter out noise in electronic communications systems, and much more. Neural networks are also known to be universal mappers or approximators; in theory (and, so it seems, in practice), they can be trained to provide as good an approximation as desired, given a sufficient number of neurons and training examples, to any well-behaved continuous function (and even to some less well-behaved functions) over a specified domain. This ability makes the neural network an ideal tool for pricing options. However, the ability of a neural network to model such a wide range of relationships comes at great expense in terms of degrees of freedom.

## HYBRID MODELS

Given sufficient data, a neural network model can theoretically solve a difficult nonlinear estimation problem, such as that of pricing an option. And, although a single polynomial regression model of reasonable order probably cannot solve the option pricing problem over its entire input domain, a set of such models, each operating over a properly chosen subdomain, almost certainly can. However, the best way to solve a difficult nonlinear estimation problem like that of pricing options might not be through the direct (and naive) application of a neural network or a set of polynomial regression models. In cases where it is feasible, a smarter approach might be to finesse a solution by use of a properly designed hybrid model.

Why would the extra intellectual difficulty of designing and custom programming a complex hybrid model be preferred over the ease and simplicity of a basic neural network or polynomial regression? Because it might be possible to obtain an equally good (if not better) fit to the data with far fewer free parameters in the model.

Obviously, the smaller the number of free parameters in the model, the fewer the degrees of freedom the model consumes. A model that consumes fewer degrees of freedom is less curve-fit, more robust, and less subject to shrinkage. In addition, a hybrid model that can be made to more "naturally" fit the problem to be solved by the incorporation of hints or content-relevant theory will produce better overall results—especially in the face of limited (read finite) data samples. The idea is similar to that which lies behind the use of so-called "biased" statistics; inclusion of a bias that reflects some a priori expectation regarding the data can result in a dramatic reduction in a statistic's standard error, especially over regions of input space where data points are sparse and estimation is poor.

How might one go about constructing a hybrid model for pricing options? Perhaps by starting with Black-Scholes. Although Black-Scholes fails to take into account skew or kurtosis, and does not correct for regression to the mean in historical volatility, it does provide a reasonable first approximation to an option's fair value. Perhaps with a bit of encouragement, Black-Scholes can be made to appraise options much more accurately. How might Black-Scholes be coaxed to yield more accurate theoretical option premiums? One way that comes to mind is through the appropriate adjustment or "tweaking" of the numbers entered into its volatility input.

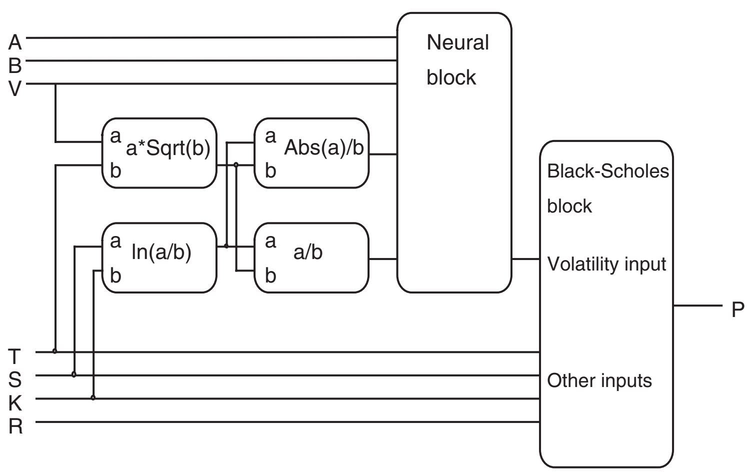
Skew, kurtosis, and regression to the mean can all be construed as factors that affect fair premium through an alteration of the effective (not the actual) volatility of the underlying security. The impact on effective volatility of these factors is not uniform, but rather it depends in a complex, nonlinear way on such things as moneyness, actual volatility, skew, kurtosis, venue, and perhaps even technical indicators. However, if the data sent to the Black-Scholes volatility input could somehow be massaged, then perhaps the model could be induced to yield far more valid theoretical price estimates. In fact, with proper adjustments to the volatility figures plugged into Black-Scholes, the model might even be persuaded to account for the effects of skew, kurtosis, and regression to the mean!

But how can the volatility input to Black-Scholes be intelligently adjusted? The answer is obvious: through the use of a neural network and some additional processing blocks. The hybrid model in Figure 7–2 combines Black-Scholes with a neural network and some simple data preprocessing to produce a better estimator of fair premium, i.e., a superior option pricing model.

Figure 7–2 may be regarded as a kind of electronic circuit diagram. At the left hand side of the diagram are the input signals: skew (A), kurtosis (B), historical (or, better yet, predicted) volatility (V), time left (T), stock price (S), strike price (K), and risk-free interest rate (R). The output (P), an estimate of the fair premium, appears at the right hand side of the diagram.

**FIGURE 7–2**

Block Diagram of a Simple Hybrid Model for Pricing Options



A Black-Scholes subsystem lurks inside the hybrid model shown in Figure 7–2. Four of the five standard inputs to Black-Scholes (time left, stock price, strike, and interest) are received by this subsystem directly from the outside world. The remaining Black-Scholes input, volatility, comes not from outside but from another component in the hybrid model, the neural subsystem. The neural subsystem gets its inputs from both the outside world (skew, kurtosis, and historical volatility) and from some simple computational blocks that provide the neural component with two distinct measures of moneyness. The preprocessing blocks pick their input data right off the hybrid entity's input lines.

It should be noted that the hybrid model in Figure 7–2 cannot be "trained" or fitted to the data with a standard neural network package—not even with a sophisticated one. The reason is that the correspondence between the target and the neural component's output is not what must be maximized; what must be maximized is the correspondence between the target and the output from the Black-Scholes block. Furthermore, Black-Scholes cannot simply be run in reverse, i.e., inverted, to obtain the volatility figures necessary to hit the targets and a neural network then trained to estimate those volatility figures. Such an approach will not work because Black-Scholes produces an output that responds nonlinearly to volatility for all options other than those that are exactly at-the-money. Given a nonlinear function f, and a random variable $y$ that is not perfectly predictable, it will generally be the case that $f(E(f^{-1}(y))) \neq E(y)$. Yet, it is $E(y)$, the expectation of the target (the terminal option price), and not $f(E(f^{-1}(y)))$ that the hybrid system must estimate if options are to be properly valued. Anyhow, the Black-Scholes relationship between volatility and theoretical premium is not even invertible when in-the-money options and terminal prices less than the intrinsic value are involved, as will be the case for many of the individual data points or facts required to train (using standard back-propagation) a neural component of a system, like the one depicted in Figure 7–2, in an effort to get the entire hybrid model to estimate the expectation of an option's theoretical price and, hence, to correctly appraise an option's true worth. The fact that standard back-propagation cannot be used to directly train the neural component of a hybrid model, such as that shown in Figure 7–2, does not in any way imply that neural connection weights that minimize the sum of the squared errors for the entire model cannot be found. It merely means that some custom mathematical programming must be done to solve the problem. The problem of finding an optimal set of connection weights for the neural component can readily be solved by any of a variety of methods, including gradient descent minimization—a process quite similar to standard back-propagation.

A hybrid model like that shown in Figure 7–2 will be encountered later in this chapter in the context of an empirical study.

## GENERAL OVERVIEW

What follows is a succession of studies intended to explore the application of polynomial regressions, neural networks, and hybrid models to the problem of pricing options. A variety of questions are asked in these studies and attempts are made to answer them on the basis of experimental evidence. Among the questions considered are: Given a sufficient set of training examples or data points, can a neural network be trained to generate theoretical option price as its output? Can a simple neural network impersonate Black-Scholes? How about a polynomial regression? Can a multivariate regression involving polynomials correctly describe the relationship between theoretical option price (the expectation of the price of an option at expiration) and the variables that determine it? How well can polynomial regressions and neural networks smooth and interpolate pricing data adduced from, say, conditional distributions? Can the application of general nonlinear models in this manner reduce the noise present in such data and thereby improve the accuracy of the results? Can a hybrid model be trained to accurately price options over the whole range of possible input values encountered in the markets? How severe is the problem of curve-fitting given the available data and the number of free parameters required for a good fit? Do the various approaches examined differ with respect to their tendency to curve-fit when applied to the pricing problem? By what means can the degrees of freedom consumed be kept small without sacrificing a good fit to the data when using a general nonlinear model?

## DATA

The stock data used in the studies that follow were extracted from the Worden TC-2000 database. Each quote in the extracted data consisted of figures for price (open, high, low, and close), volume, and split factor. The price and volume figures in the data were back-adjusted for splits; however, the split factor allowed for recovery of the original prices and volume when necessary. For each stock, these quotes were arranged into a series. Each series consisted of 1,836 bars spanning the period from January 2, 1996 to April 14, 2003. Not every stock was active over the entire period and so, in some cases, the data series was padded. Bars that came before a stock was alive, or after it had died and had permanently ceased to trade, were marked by volume figures that were precisely zero. Quotes for days when a stock did not trade, but was clearly still alive, were marked by volume figures that were greater than zero, but far smaller than even a single transaction would produce. The opening price of the first active bar defined the price fields of all earlier bars and the closing price of the last active bar defined the price fields of all subsequent bars, so that no bar had a price of zero; this enabled vectorized routines to be used without encountering divide-by-zero or domain errors when, e.g., calculating historical or future volatilities. Each data series was examined for errors. Small errors that could be reasonably corrected were corrected. Larger or more persistent errors were handled by deletion of the series for the stock exhibiting them; only nine series suffered this fate. Overall, there were very few errors in the data that needed correction. There were a total of 2,237 series or stocks after all bad data had been corrected or eliminated. The complete set of data series was placed in a simple binary database file for fast and convenient access. Further details on the extraction and preparation of this data can be found in Chapter 4.

Theoretical option premiums derived from conditional distributions were used in some of the studies as standards for comparison. These were obtained from spreadsheets prepared in the course of writing Chapter 6. The spreadsheets contained all the data generated by the C language programs that performed the analyses—by far more data than were actually presented in this text.

Finally, some of the data sets used in the following studies take the form of plasmodes. A *plasmode* is a data set specifically constructed to be consistent with a given theory or with the assumptions made by a particular analytical method. The term originated in the world of factor analysis, where artificial data sets were sometimes assembled for testing purposes. These data sets were designed to have an underlying organization consistent with the factor model and to possess a known factor structure. The idea was to see whether some particular method of factor analysis could recover the hidden structure embedded in the data. Using plasmodes, various methods of factor extraction and rotation could be compared in terms of their ability to correctly identify the already known factors. Plasmodes happen to be quite useful in the present context. One plasmode soon to be encountered is the Black-Scholes plasmode employed in Study 1.

## SOFTWARE

Several software packages, libraries, and programming languages were used for the analyses reported herein. Most calculations were performed using ISO-standard C, a highly portable language. In some instances, however, the *Visual Basic for Applications* language that is built into Microsoft's Excel spreadsheet program was employed. Libraries written in C language provided routines for everything from database access to regression analysis. Some of these routines were translated to Excel's Visual Basic language to make them available and convenient to use in that environment. N-Train, a sophisticated neural network package designed (by Katz) to operate with high levels of numerical stability, was employed to train and test standard feed-forward perceptrons. A special library provided all the functionality needed to train and test hybrid models of the kind shown in Figure 7–2. All tables and presentation graphics were prepared with Microsoft's Excel spreadsheet program.

## STUDY 1: NEURAL NETWORKS AND BLACK-SCHOLES

The purpose of this study was to determine whether a standard, feed-forward neural network—a multilayer perceptron with a linear input layer and sigmoidal hidden and output layers—could learn to imitate Black-Scholes. Why bother training a complex neural network to do what Black-Scholes already does more simply when the goal is to do better? Because, if a neural network cannot reproduce the functional relationships implicit in Black-Scholes, it is quite unreasonable to expect it to be able to correctly interpret the relationships that would be implicit in a better, more advanced pricing model. However, if a neural network can learn to emulate Black-Scholes, perhaps it can also learn to approximate the relationships demonstrable in data taken from real stocks.

Assuming that a neural network can learn Black-Scholes, how large must it be to achieve a good fit? How many degrees of freedom will it consume? Are there any particular difficulties in training such a neural network? These are some additional questions addressed in the current study.

### General Methodology

The experiment began with the construction of a plasmode. The plasmode used in the current study consisted of a set of data points— "facts" in the jargon of neural engineering—constructed using the standard Black-Scholes pricing model. Each data point or fact consisted of three independent or input variables, and one dependent or target variable. The input variables were strike price, time left, and volatility. Stock price was assumed to be fixed at a nominal $100.00 and risk-free interest was assumed to be zero for the purpose of this experiment. The target was the Black-Scholes theoretical premium for a standard call option.

Inputs for the plasmode were constructed by stepping volatility from `vxmin` to `vxmax` in increments of `vxinc`; time left from `tlmin` to `tlmax` in increments of `tlinc`; and strike price from `skmin` to `skmax` in increments of `skinc`. For each possible combination of inputs, Black-Scholes was applied to obtain the associated target, the theoretical price for a call option. A large set of facts that evenly covered the input space defined by the ranges of the input variables was thereby generated.

A final step in the construction of the plasmode was to subtract from each target the intrinsic value of the call in order to obtain the call's time value. Intrinsic value requires no statistical estimation to evaluate; it is easily and unambiguously calculated as *max* (0, *stock$* – *strike$*) for a call or *max* (0, *strike$* – *stock$*) for put. It is not intrinsic value, but time value that is of critical interest. Time value is the component of option premium that requires something more than simple arithmetic to evaluate and is difficult to model in the real world. The idea was that by subtracting intrinsic value, a well-understood and easily determined variable, the target's range would be reduced and the neural network would be forced to focus on the most important relationships present in the data. There is no disadvantage to having a model that estimates time value only; intrinsic value may easily be added back to any estimate of time value, neural or otherwise, in order to obtain an option's theoretical premium.

Once the plasmode had been constructed, attempts were made to train several neural networks to fit the data.

## Test of a Small Neural Network

In the first test, an attempt was made to train a small, three-layer neural network to mimic Black-Scholes over a relatively wide range of input values. For this test, volatility ran from 0.10 (`vxmin`) to 2.00 (`vxmax`) in steps of 0.05 (`vxinc`); time left ran from 5 (`tlmin`) to 20 (`tlmax`) bars in steps of 1 (`tlinc`); and strike price ran from $70.00 (`skmin`) to $130.00 (`skmax`) in steps of $2.50 (`skinc`). This resulted in a plasmode containing 15,600 facts. The neural network in this test had three active input neurons, 16 active neurons in a single hidden layer, and one output neuron.

All training and testing of the neural model was performed using N-Train, a neural network software package developed and maintained by the authors. Facts were scaled and shuffled using default parameters. A new neural network was then created. Standard (default) transfer and error functions were specified: neurons in the input layer employed linear functions, those in the hidden and output layers employed logistic functions, and the error function was standard least-squares. The logistic function, $f(x) = 1/(1 + \exp(-x))$ has a sigmoidal form and is one of the most popular neural transfer functions. The learning rate was set to 1, momentum was turned off, the connection weights were randomized with a dispersion of 1 and a seed of 93 for the Gaussian random number generator, and the maximum number of training passes was set to 99,999 (the `setparms` arguments for this were `–lr1`, `–mm1`, `–rw1s93`, and `–mr99999`). Learning rate multipliers, training tolerances, and other parameters were left at default values. It should be noted that, with these settings, the neural network is trained on all facts on every pass through the fact set. These settings were chosen as they tend to yield the most stable convergence when a neural network must provide a high-accuracy fit to the data, certainly a requirement here.

The neural network underwent over 15,000 training passes on the plasmode data. It was trained to full convergence. Once trained, the neural network was polished. Polishing was accomplished by running a few additional passes over the data with the learning rate reduced to 0.01, a very small number. The idea behind polishing is to remove any small bias that may be created by the process of training.

### Results with a Small Neural Network

A multiple correlation of 0.999724 between the neural output and the Black-Scholes target was achieved for the fully trained network. Correction for shrinkage brought this correlation down to 0.999723, given a fit to 15,600 cases of a model with 64 degrees of freedom ($3 * 16 + 16 * 1$ connection weights). A correlation this close to a perfect 1.0 may appear quite impressive and is certainly rarely seen when working with neural networks. However, even such an impressive correlation is not quite sufficient in the present context. The root-mean-square error for the model was \$0.0855 and the worst absolute error was \$1.465, certainly unacceptable.

It should be noted that most neural network software packages are incapable of training a neural network to anywhere near the required level of precision or even to the unacceptable level of precision achieved in this test. Finding a set of optimal connection weights in a neural network is an ill-conditioned problem that requires special attention to the accumulation of roundoff error, numerical stability, and other computational issues, if correlations much beyond 0.99 are to be achieved. Most packages give little, if any, attention to computational issues and, at best, use "single precision" floating point arithmetic, which is barely adequate to achieve even modest correlations with difficult data. To go much beyond a 0.99 correlation, a learning algorithm coded in "double precision," and designed to maximize numerical stability and to control for loss of precision in intermediate calculations, is essential. N-Train goes so far as to perform some calculations, such as inner products, in greater than double precision arithmetic in order to minimize the accumulation of roundoff errors and to achieve good numerical stability. Although neural networks are not usually employed for high accuracy approximations, they can unquestionably be used for this purpose when correctly implemented.

Returning to the test results, the worst errors were not evenly distributed among the data points, but tended to be seen at the edges of the input space, especially at that edge represented by the lowest time and volatility levels where the peak in time premium was sharpest when plotted against the strike price.

A small sample of the option premium data appears in Table 7–1. At the left hand side of the table are the strike prices (STRIKE) and at the top are three categories of volatility and time remaining. The premiums appear in pairs of columns. The first column of each pair (NOUT) contains the premium estimates produced by the neural network; the second column (TARG) contains the Black-Scholes target to which the network was trained. The full premium for a call option is the time value premium (which appears in the table) plus the intrinsic value (the positive difference between the nominal stock price of $100.00 and the strike).

As can be seen in Table 7–1, by far the worst errors (the most distressing differences between the neural output and

**TABLE 7–1**

Correspondence between Neural Output and Black-Scholes for a Three-Layer Feed-Forward Network with 16 Hidden-Layer Neurons

| | TIME = 5 VLTY = 0.20 | | TIME = 15 VLTY = 0.50 | | TIME = 15 VLTY = 1.00 | |
|---|---|---|---|---|---|---|
| STRIKE | NOUT | TARG | NOUT | TARG | NOUT | TARG |
| 70.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.15 | 0.27 |
| 72.50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.32 | 0.41 |
| 75.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.56 | 0.61 |
| 77.50 | 0.00 | 0.00 | 0.01 | 0.02 | 0.87 | 0.89 |
| 80.00 | 0.00 | 0.00 | 0.02 | 0.04 | 1.24 | 1.24 |
| 82.50 | 0.00 | 0.00 | 0.08 | 0.10 | 1.69 | 1.68 |
| 85.00 | 0.00 | 0.00 | 0.19 | 0.22 | 2.26 | 2.23 |
| 87.50 | 0.00 | 0.00 | 0.42 | 0.42 | 2.95 | 2.89 |
| 90.00 | 0.01 | 0.00 | 0.78 | 0.74 | 3.76 | 3.68 |
| 92.50 | 0.04 | 0.00 | 1.29 | 1.24 | 4.58 | 4.58 |
| 95.00 | 0.11 | 0.01 | 1.94 | 1.93 | 5.45 | 5.62 |
| 97.50 | 0.24 | 0.16 | 2.92 | 2.86 | 6.82 | 6.78 |
| 100.00 | 0.34 | 0.93 | 3.90 | 4.04 | 8.26 | 8.07 |
| 102.50 | 0.19 | 0.18 | 2.98 | 2.96 | 7.08 | 6.98 |
| 105.00 | 0.08 | 0.02 | 2.14 | 2.11 | 5.97 | 6.01 |
| 107.50 | 0.03 | 0.00 | 1.53 | 1.47 | 5.15 | 5.15 |
| 110.00 | 0.01 | 0.00 | 1.04 | 0.99 | 4.37 | 4.39 |
| 112.50 | 0.00 | 0.00 | 0.68 | 0.65 | 3.70 | 3.73 |
| 115.00 | 0.00 | 0.00 | 0.43 | 0.42 | 3.14 | 3.15 |
| 117.50 | 0.00 | 0.00 | 0.25 | 0.26 | 2.66 | 2.65 |
| 120.00 | 0.00 | 0.00 | 0.14 | 0.16 | 2.25 | 2.22 |
| 122.50 | 0.00 | 0.00 | 0.07 | 0.09 | 1.88 | 1.86 |
| 125.00 | 0.00 | 0.00 | 0.03 | 0.06 | 1.54 | 1.55 |
| 127.50 | 0.00 | 0.00 | 0.01 | 0.03 | 1.24 | 1.28 |
| 130.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.96 | 1.06 |

the target) are found in the leftmost time left and volatility category, which lies at one extreme of the input space. When time left is short and volatility is low, the neural network has great difficulty in modeling the sharp peak in the Black-Scholes pricing data. There are nine instances where either the neural or

Black-Scholes price is greater than \$0.00. In four of these instances (44%), the error is greater than \$0.05. The root-mean-square error over all nine cases is a sizable \$0.20, while the largest error is a whopping \$0.60.

The errors are much smaller, but still not quite acceptable, in the middle category of time left and volatility. Here, there are 23 cases where either the neural or Black-Scholes price is greater than \$0.00. Four cases (17%) demonstrate an error larger than \$0.05. The largest absolute error is still quite sizable at \$0.14, but the root-mean-square error is an almost tolerable \$0.04.

In the rightmost category, at the other edge of the model's input space, the errors are again more substantial. Here, there are 25 prices greater than \$0.00, a largest error of \$0.19, eight errors (32%) greater than \$0.05, and a root-mean-square error of \$0.072.

## Test of a Larger Neural Network

Perhaps the problem in the previous test was that the neural network was simply too small to fully model the fine nuances found in the Black-Scholes plasmode. No doubt, a better fit to the data can be obtained with a larger, more complex neural model. The current test investigates whether a more complex neural model can, indeed, provide a more acceptable fit to the plasmode data than that achieved by the smaller neural network in the previous test.

In the current test, a four-layer neural network with three active neurons in the first or input layer, 26 active neurons in the second layer, eight active neurons in the third layer, and one neuron in the fourth or output layer was trained. The data were identical to what was used in the test of the smaller neural network. Scaling was also the same as with the smaller network (the N-Train default), except for the maximum output, which was raised from 18.50165 to an even 20.0 in order to reduce the unnecessary curvature imposed by the presence of an artificial boundary condition where none was required. The minimum output was left at zero, as this is a valid boundary condition for option price. Transfer and error functions were N-Train defaults: linear transfer functions for the input layer, sigmoid (logistic function) transfers for all other layers, and an error function equal to the square of the difference between the neural network's output and the target. The global learning rate was initially set to 1.0, but both the global learning rate and the layer-by-layer learning rate multipliers were adjusted now and then over the course of training in order to hasten convergence. Connection weights were initialized with a Gaussian random generator set to a dispersion of 1.0, and seeded with 93, just as was the case for the smaller network tested earlier. No momentum was employed and training took place on every fact in the plasmode on each pass through the data. The neural network was trained to full convergence, which required many tens of thousands of passes over the plasmode data to achieve. Given the size of the network, and the accuracy to which it was to be trained, tens of thousands of training passes should not be viewed as excessive or unexpected. Once trained, the network was polished by performing a few additional training runs with the learning rate set to a very small number. Finally, the trained neural network was tested. The process involved using the trained network to generate an option price estimate for each fact in the plasmode and then comparing each such price estimate to the corresponding Black-Scholes target price.

## Results with a Larger Neural Network

In contrast to the smaller neural network, which produced unacceptable results in the previous test, the larger network in the current test performed admirably. It must be stated that this larger network consumed 294 ($= 3 * 26 + 26 * 8 + 8 * 1$) degrees of freedom, but it yielded an extraordinary multiple correlation of 0.99999941 between its output and the Black-Scholes target. The root-mean-square error was a measly \$0.00397, while the worst absolute error in the entire sample of 15,600 data points was now only \$0.0999. Both error measurements were far smaller than those calculated for the less complex neural model in the earlier test. The data presented in Table 7–2 bear this out.

Table 7–2 is virtually identical to Table 7–1, except that the neural network theoretical price estimates (NOUT) are those

**TABLE 7–2**

Correspondence between Neural Output and Black-Scholes for a Four-Layer Feed-Forward Network with 26 and 8 Hidden-Layer Neurons

| | *TIME = 5 VLTY = 0.20* | | *TIME = 15 VLTY = 0.50* | | *TIME = 15 VLTY = 1.00* | |
|---|---|---|---|---|---|---|
| **STRIKE** | **NOUT** | **TARG** | **NOUT** | **TARG** | **NOUT** | **TARG** |
| 70.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.27 | 0.27 |
| 72.50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.41 | 0.41 |
| 75.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.62 | 0.61 |
| 77.50 | 0.00 | 0.00 | 0.02 | 0.02 | 0.89 | 0.89 |
| 80.00 | 0.00 | 0.00 | 0.05 | 0.04 | 1.24 | 1.24 |
| 82.50 | 0.00 | 0.00 | 0.10 | 0.10 | 1.68 | 1.68 |
| 85.00 | 0.00 | 0.00 | 0.22 | 0.22 | 2.23 | 2.23 |
| 87.50 | 0.00 | 0.00 | 0.42 | 0.42 | 2.89 | 2.89 |
| 90.00 | 0.00 | 0.00 | 0.75 | 0.74 | 3.68 | 3.68 |
| 92.50 | 0.00 | 0.00 | 1.23 | 1.24 | 4.58 | 4.58 |
| 95.00 | 0.02 | 0.01 | 1.94 | 1.93 | 5.62 | 5.62 |
| 97.50 | 0.17 | 0.16 | 2.86 | 2.86 | 6.78 | 6.78 |
| 100.00 | 0.93 | 0.93 | 4.03 | 4.04 | 8.07 | 8.07 |
| 102.50 | 0.19 | 0.18 | 2.96 | 2.96 | 6.98 | 6.98 |
| 105.00 | 0.03 | 0.02 | 2.12 | 2.11 | 6.01 | 6.01 |
| 107.50 | 0.00 | 0.00 | 1.46 | 1.47 | 5.15 | 5.15 |
| 110.00 | 0.00 | 0.00 | 0.99 | 0.99 | 4.39 | 4.39 |
| 112.50 | 0.00 | 0.00 | 0.65 | 0.65 | 3.73 | 3.73 |
| 115.00 | 0.00 | 0.00 | 0.42 | 0.42 | 3.15 | 3.15 |
| 117.50 | 0.00 | 0.00 | 0.26 | 0.26 | 2.65 | 2.65 |
| 120.00 | 0.00 | 0.00 | 0.16 | 0.16 | 2.22 | 2.22 |
| 122.50 | 0.00 | 0.00 | 0.10 | 0.09 | 1.86 | 1.86 |
| 125.00 | 0.00 | 0.00 | 0.06 | 0.06 | 1.55 | 1.55 |
| 127.50 | 0.00 | 0.00 | 0.03 | 0.03 | 1.28 | 1.28 |
| 130.00 | 0.00 | 0.00 | 0.02 | 0.02 | 1.06 | 1.06 |

produced by the more complex, four-layer neural model. The differences between the neural outputs (NOUT) and the Black-Scholes prices (TARG) are much smaller than in Table 7–1, attesting to the vastly better fit achieved by the larger neural network model.

For a time remaining of 5 bars, and a volatility of 20%, there are five prices greater than \$0, not one of them displaying an error greater than \$0.01. The root-mean-square error is \$0.011 for these figures. There are 23 prices greater than \$0 when time left is 15 and volatility 50%. Not one error greater than \$0.01 is in evidence and the root-mean-square error is only \$0.004. When time left is 15 and volatility 100%, only one error of \$0.01 appears, no error is greater, and the root-mean-square error is a negligible \$0.002. Overall, the four-layer neural network trained in this test does a commendable job of emulating Black-Scholes, with errors generally far less than the size of a single tick.

### Discussion

Although the smaller, three-layer neural network trained in the first test was unable to adequately price options according to Black-Scholes, the larger, four-layer network trained in the second test performed very well. The match between Black-Scholes prices and those obtained from the larger neural model was quite good; almost all errors were well below one tick or price increment (\$0.05 for most options on most exchanges) in size. For practical purposes, the four-layer network could serve as a drop-in replacement for Black-Scholes. In comparison to the smaller, three-layer neural network, the four-layer one had no difficulty in modeling the sharp peak in time premium as a function of strike price, that is associated with conditions of low volatility and little time left. The results indisputably demonstrate that a four-layer neural model of moderate complexity can, when trained properly, accurately impersonate Black-Scholes over a relatively wide range of input values, and price options trading on securities having log-normal price distributions.

Given that a sufficiently complex neural network can imitate Black-Scholes, chances are that such a network can also model real-market price expectations reasonably well. The question is whether or not the neural network would also model the noise in the data, rather than smoothing out the random variations to reveal the true relationships that presumably exist between the model inputs and future price expectation, the target. With a large number of free parameters, like that present in the neural model examined in the current study, fitting of the irrelevant patterns in the data, the noise, as well as the relevant ones, is a distinct possibility. In other words, excessive curve-fitting may be an issue in developing a neural model to price options.

One way to handle the issue of curve-fitting and of an excess consumption of degrees of freedom is to include some prior knowledge in the model. This is loosely the idea behind Bayesian and "biased" statistics. In the present situation, incorporating the knowledge implicit in Black-Scholes might do the trick. Data points based on Black-Scholes, evenly distributed throughout the data space, might be added to the fact set on which the neural network would be trained. The addition of such data points would, for one thing, increase the total number of training facts. But, more importantly, it would increase the relative density of data points in regions of the data space where real-market data points are sparse. The effect, imagined visually, would be a response surface that, in regions of data space densely populated with real-market data points, slices right through those points, but that, in regions where real-market data points are scanty, is pulled towards (or defaults to) the nearby areas populated by the Black-Scholes data points. In regions of data space where there are many real data points, the Black-Scholes bias would have only a minuscule impact on the solution; however, in the absence of sufficient outside information, the bias would reduce the tendency of the response surface to oscillate wildly or to wander far away from where it should be—something that often happens around the fringes of the data space, where nonlinear models seem to shoot for the sky or go off the deep end.

If a neural network can be trained to accurately price options, whether directly or by a "biased statistics" approach like the one discussed above, an efficient, practical pricing model will be the result. A neural network may take a massive amount of computation, and hence hours or even days, to fully train; once trained, however, pricing options with a neural network can be almost as fast as pricing options with a standard model like Black-Scholes and by far faster than pricing options with Cox-Ross-Rubinstein.

The reader may be thinking, "That is not true. The networks tested above had only strike, volatility, and time left as inputs; a usable model would also require stock price and interest rate." In fact, the absence of these inputs presents no problem. It will be recalled from earlier in this work that, given certain reasonable assumptions, a stock's price may be rescaled to a nominal $100, the strike price of the option similarly rescaled, the option price computed with the rescaled stock and strike prices, and the resultant option price then transformed back to the original scaling. Adjustment for risk-free interest may also be performed without changes to the basic model: factor the growth due to interest in a risk-neutral world into the stock price, price the option using the basic model, then discount the resultant theoretical option premium for the growth, due to interest, expected for any security in a risk-neutral world. Voila! The two inputs missing from the model have now been taken care of so that any standard option may be priced.

It should be noted that several tests (not reported in this study) were performed in an effort to find a way to conserve degrees of freedom and to speed up the training process. One such test used a larger three-layer neural network; the gain in performance over the smaller three-layer model discussed earlier was not sufficient to justify further exploration given that vastly better results could readily be obtained from a four-layer model. Another test involved a two-step procedure: First, a small, simple neural network was quickly trained to provide a rough initial approximation to the Black-Scholes target. Second, another net was coached to estimate the error made by the first neural network. The results were no better than those achieved with greater simplicity by more extensive training of a single, somewhat larger neural model. Finally, some additional tests involved experimentation with various error and transfer functions, learning parameters, and other related elements in an effort to speed convergence. Although convergence could be coaxed to take place marginally faster, all the hands-on attention required was not worth the effort. What it boils down to is that model fit is determined primarily by the number of connection weights—free parameters—and that training must continue until full convergence is achieved, a lengthy process with no work-arounds (except, perhaps, more efficient training algorithms).

None of the additional, unreported tests involved the "biased statistics" approach outlined earlier in this chapter. This approach may ameliorate the deleterious effect of an excessive number of free parameters and yield an overall more valid and robust solution, but it will neither result in faster training (in fact, training will be slower as a result of the additional data points in the fact set) nor reduce the actual number of free parameters in the model.

## STUDY 2: POLYNOMIAL REGRESSIONS AND BLACK-SCHOLES

Neural networks are evidently able to emulate Black-Scholes with a high degree of precision. Can polynomial models do as well? Are polynomial models any better with respect to the degrees of freedom necessary for a good fit to the data? What are the advantages, or disadvantages, if any, of polynomial regression models over neural networks? These are some of the questions that the current study addresses.

### Method

The current study follows the same basic logic as the previous one, except that a multivariate polynomial regression, rather than a neural network, was fitted to the Black-Scholes plasmode or data set.

The plasmode used in the current study was the same as that used in the previous one, with two alterations: one, the range of strike price was reduced from the previous range of $70 to $130 to a range of $75 to $125 in the current investigation; and two, the range of volatility was reduced from the previous range of 10% to 200% to a range of 20% to 200%. The response of option premium to strike is hard to model at low levels of volatility and time left; it was for this reason that the ranges were reduced to make it somewhat easier for a polynomial model of modest order to provide a reasonably good fit to the data. Even a low-order polynomial can model an analytic function well over a sufficiently small region of its domain. Achieving a good fit over a larger span or region may require polynomials of much higher order and, therefore, may demand that many more free parameters be optimized. Each fact or data point in the data set consisted of three independent variables and one dependent or target variable. The three independent variables were volatility, time left (measured in bars), and strike price. A nominal $100 stock price was assumed, as was a zero interest rate. The target was the same as in the previous investigation, the theoretical option price derived from Black-Scholes.

The fitting of the model involved the following steps. First, a fact was retrieved from the data set. The volatility (one independent variable) was multiplied by the square root of the time left (another independent variable) and the result was then rescaled to have a range extending from –1 to +1. In C language, the rescaled time-and-volatility composite was calculated as

```
x1 = 2.0 * (vx * sqrt (tl) - 0.2 * sqrt (5.0))/
     (2.0 * sqrt (20.0) - 0.2 * sqrt (5.0)) - 1.0
```

where $-1 \le$ `x1` $\le 1$, `vx` was the original volatility, and `tl` was the time left. Chebyshev Polynomials of orders 0 to 13 were then evaluated at `x1`, the appropriately scaled composite of volatility and time. The evaluations for polynomials of orders 0 to 13 were placed in `tx1[0]` through `tx1[13]`, respectively. The recurrence relationship

$$
\begin{aligned}
T_0 &= 1 \qquad\qquad (7.8)\\
T_1 &= x\\
T_n &= 2xT_n - T_{n-2} \quad \text{(for all } n > 1\text{)}
\end{aligned}
$$

where $T_n$ represents the Chebyshev Polynomial of order $n$, and $x$ represents the value at which the polynomial is to be evaluated, was used to determine the Chebyshev Polynomials efficiently and with high precision.

Next, the strike price was rescaled to a range of –1 to +1 and placed in `x2`. In C, the rescaled strike was computed as

```
x2 = 2.0 * (sk - 75.0)/(125.0 - 75.0) - 1.0
```

where $-1 \le$ `x2` $\le 1$ and `sk` was the strike price taken from the fact. Again, a set of Chebyshev Polynomials was evaluated, this time at x2, the rescaled strike. The results were placed in tx2[0] through tx2[17]. Since the relationship of premium to strike has been found harder to model over the range of concern than that of premium to volatility, polynomials of orders up to 17 were analyzed.

All cross-products of the elements of tx1 and tx2 were then computed, with the results placed in x[0] through x[251]. In C language, the code to accomplish this looked like

```
for (i2 = 0; i2 <= 17; i2++)
  for (i1 = 0; i1 <= 13; i1++)
    x[i1 + 13 * i2] = tx1[i1] * tx2[i2];
```

Since the Chebyshev Polynomial of order 0 is always 1, no separate computations needed to be performed for terms involving only x1 or x2, i.e., for those terms not actually involving true cross-products.

A new, expanded fact, consisting of x[0] through x[251] as independent variables, and the target from the original fact as the dependent variable, was the result. This expanded fact was saved in a standard fact file. Finally, another fact was retrieved from the Black-Scholes plasmode and the sequence of steps described above was repeated. The process continued until all data points in the plasmode had been analyzed, expanded, and written to the output fact file.

At this point, a multivariate linear regression was performed on the expanded facts. All calculations were done in double precision arithmetic in order to avoid the accumulation of roundoff error that may easily occur when working with large arrays and data sets.

Chebyshev Polynomials were employed in the model because they have two very desirable properties: (1) they are orthogonal, at least under certain conditions, and (2) their range is the interval from −1 to +1. Both these properties help to avoid ill-conditioning and loss of precision in the determination and use of a regression model. Were simple powers and cross-products of the independent variables used to compute the regression, a total or near-total loss of precision would almost certainly be encountered for models of the order contemplated—assuming an attempt to invert a singular matrix did not interrupt or crash the regression procedure. Using Chebyshev Polynomials, models of amazingly high order can be fitted to large data sets without encountering ill-conditioned matrices and related problems. The better scaling and the reduced collinearity associated with orthogonal (uncorrelated) terms or independent variables are the keys.

Naturally, the orthogonality of the Chebyshev Polynomials strictly holds only under very specific conditions that are rarely, if ever, met by empirical data. However, when employed in the manner described above, these polynomials are close enough to orthogonal for making stable and accurate computation of a regression model using standard double precision arithmetic feasible.

Apart from contributing to the accuracy of the solution, the –1 to +1 output range of the Chebyshev Polynomials confers another benefit: the additional error, or change in the regression estimate of the dependent variable, incurred by dropping a term from the model is limited to the size of the deleted term's coefficient. When looking over the regression coefficients, it is easy to see the relative contribution of each of the terms to the model. This facilitates understanding. It also facilitates simplification through an appropriate choice of model order for each independent variable and through a process of pruning in which terms with coefficients deemed sufficiently small are dropped.

Time left and volatility were combined to reduce the complexity of the problem. The inclusion of all three independent variables directly from the fact set would have required many more free parameters in the regression. Polynomial models have the disadvantage that they are much more subject to combinatorial explosion in the number of free parameters with increasing numbers of inputs than are neural models. The way in which time and volatility were combined into a single variable relies on the assumption that successive price movements or returns are statistically independent—an assumption implicit in the Black-Scholes model and only "slightly violated" by real-world prices (see Chapter 4).

## Results

The polynomial model fit the Black-Scholes data like a tight glove. Over the 12,432 facts in the sample, the correlation between Black-Scholes and the estimate made with the polynomial model was an incredible 0.99999999, with a root-mean-square error of only \$0.000409 and a largest error of \$0.0124. Even when the model was fitted over a wider range of strikes (\$70 to \$130), the root-mean-square error remained small at \$0.00129 and the maximum error rose only to \$0.0415. As before, the largest error occurred with volatility and remaining time at their lowest levels, where the peak in theoretical premium as a function of strike is sharpest.

Table 7–3 shows the polynomial regression estimates (PREST) and the Black-Scholes targets (TARG) as a function of strike price (STRIKE) for three combinations of time left and volatility. As can easily be seen, the match between the regression model and Black-Scholes is superb. The quality of the fit is clearly evident in the root-mean-square and largest absolute errors that were calculated from the data in Table 7–3. With 5 bars left and 20% volatility, the root-mean-square error was \$0.0063 and the largest absolute error was \$0.0124. The root-mean-square error was \$0.0003 and the largest error was \$0.0005 when 15 bars (trading days) remained until expiration and volatility was 50%. With 15 bars left and 100% volatility, the error figures were all \$0.0001 (or less), reflecting an exceptional fit of the polynomial model to the Black-Scholes data.

The excellent fit to the Black-Scholes data was achieved with a regression model having 252 coefficients or free parameters. Many of these coefficients were quite small, especially those associated with terms involving higher order polynomials. The model could easily be simplified by pruning (removing) those terms that contribute little to the final estimate of fair premium.

## Discussion

No doubt, a polynomial regression can emulate Black-Scholes very well. And, by way of comparison to a neural network, a polynomial model is fast to train—minutes instead of hours or days.

However, to achieve a good fit with a polynomial model having roughly the same number of free parameters as a neural model, a reduction in the number of inputs to the model is required. In order to simplify the model in the current instance,

**TABLE 7–3**

Correspondence between Theoretical Call Premiums Computed with a Polynomial Regression Estimator and with Black-Scholes as a Function of Strike, Volatility, and Time

| | TIME = 5 VLTY = 0.20 | | TIME = 15 VLTY = 0.50 | | TIME = 15 VLTY = 1.00 | |
|---|---|---|---|---|---|---|
| STRIKE | PREST | TARG | PREST | TARG | PREST | TARG |
| 75.00 | 25.00 | 25.00 | 25.01 | 25.01 | 25.61 | 25.61 |
| 77.50 | 22.50 | 22.50 | 22.52 | 22.52 | 23.38 | 23.38 |
| 80.00 | 20.00 | 20.00 | 20.04 | 20.04 | 21.24 | 21.24 |
| 82.50 | 17.50 | 17.50 | 17.60 | 17.60 | 19.18 | 19.18 |
| 85.00 | 15.00 | 15.00 | 15.21 | 15.21 | 17.23 | 17.23 |
| 87.50 | 12.50 | 12.50 | 12.91 | 12.92 | 15.39 | 15.39 |
| 90.00 | 9.99 | 10.00 | 10.74 | 10.74 | 13.68 | 13.68 |
| 92.50 | 7.51 | 7.50 | 8.74 | 8.74 | 12.08 | 12.08 |
| 95.00 | 5.01 | 5.01 | 6.93 | 6.93 | 10.62 | 10.62 |
| 97.50 | 2.66 | 2.66 | 5.36 | 5.36 | 9.28 | 9.28 |
| 100.00 | 0.95 | 0.93 | 4.04 | 4.04 | 8.07 | 8.07 |
| 102.50 | 0.17 | 0.18 | 2.96 | 2.96 | 6.98 | 6.98 |
| 105.00 | 0.02 | 0.02 | 2.11 | 2.11 | 6.01 | 6.01 |
| 107.50 | 0.01 | 0.00 | 1.47 | 1.47 | 5.15 | 5.15 |
| 110.00 | –0.01 | 0.00 | 0.99 | 0.99 | 4.39 | 4.39 |
| 112.50 | 0.00 | 0.00 | 0.65 | 0.65 | 3.73 | 3.73 |
| 115.00 | 0.00 | 0.00 | 0.42 | 0.42 | 3.15 | 3.15 |
| 117.50 | 0.00 | 0.00 | 0.26 | 0.26 | 2.65 | 2.65 |
| 120.00 | 0.00 | 0.00 | 0.16 | 0.16 | 2.22 | 2.22 |
| 122.50 | 0.00 | 0.00 | 0.09 | 0.09 | 1.86 | 1.86 |
| 125.00 | 0.00 | 0.00 | 0.06 | 0.05 | 1.55 | 1.55 |

time left and volatility had to be combined into a single variable. Of course, combining time and volatility in the manner done in this study is only legitimate if successive returns in the marketplace are statistically independent—that successive returns are statistically independent is an assumption made by Black-Scholes, but one that is not necessarily valid. If time remaining and volatility had not been combined into a single variable, the polynomial regression would have required estimates for thousands of coefficients, rather than for a mere 252 of them.

And what if skew or kurtosis is to be incorporated into a polynomial regression as input variables? Oops. Up goes the number of free parameters by another factor of 10 or more. That is the main problem with multivariate polynomial models: the rapid explosion in the number of coefficients or free parameters (and degrees of freedom consumed) with any increase in the number of model inputs. Neural networks are much more graceful in the way they respond to the addition of input variables to the model and can generally provide an equally good fit to the data.

The conclusion is that if polynomial regressions are to be used to price options, substantial simplification of the problem must be made a priority. There are many ways in which such simplification may be accomplished. Time left and strike, e.g., could be eliminated as input variables; instead, separate polynomials could be developed for each particular combination of time left and strike. There would then be headroom, in terms of degrees of freedom, for the addition of other variables like skew or kurtosis as model inputs. Naturally, this means an untidy profusion of polynomials, rather than a clean, unitary model. Although polynomials can be used to price options and to smooth and interpolate the empirical data on which they (the polynomials) are based, neural networks may be a better choice in the context of option pricing, given their ability to handle more inputs without requiring as vast an increase in the number of parameters that must be determined when fitting the model to the data.

## STUDY 3: POLYNOMIAL REGRESSIONS ON REAL-MARKET DATA

In Studies 1 and 2, neural networks and polynomial regressions were shown to be fully capable of pricing options, at least according to Black-Scholes. But, can such general-purpose nonlinear models price options based on the behavior of stock returns in the actual marketplace? Can they smooth out the noise and capture the essential relationships in the data? The next three studies, beginning with the current one, attempt to answer these questions.

The current study explores the use of a simple polynomial regression to capture the relationship between terminal option price expectation, the target, and time left, estimated future volatility, and strike, the model inputs. In other words, a polynomial regression model to appraise options based on actual stock price behavior is developed and evaluated. To maintain simplicity, only call options are considered; if calls can be successfully modeled by polynomials, so can puts.

## Method

The analysis took place in three computational units or blocks. The first computational unit involved preparation of the fact set or training data. The second unit performed polynomial expansions on the facts. And the third unit fitted a multivariate linear regression to the expanded facts.

***First Computational Unit*** Code implementing the first unit embodied the following algorithm or sequence of actions. First, all data series for one of the 2,246 stocks were retrieved from the same binary stock database used in earlier chapters. Details regarding construction of the stock database can be found in the introductory sections of Chapter 4. Next, a valid reference bar was chosen from the 1,834 available bars. A reference bar was valid if (1) there were at least $m1 + m2$ prior bars over which the stock was active, (2) there were at least $n$ bars beyond the reference over which the stock was active, and (3) the lowest original (unadjusted for splits) closing price over the previous $m1$ bars was greater than or equal to $2. In this analysis, $m1$ was 30, $m2$ was 70, and $n$ was 10.

When a valid reference bar had been located, an estimate of near-future volatility was computed with the multivariate regression model developed in Study 4 of Chapter 5. The estimated future volatility was used to determine the first of two array indices. In C-style notation, this array index was calculated as

```
ivx = (int)floor(0.5 + (nvx - 1)*(vx - bvxmn)/(bvxmx - bvxmn))
```

where `ivx` was the array row index, `nvx` was the number of volatility levels to be used in the analysis, `vx` was the estimated near-future volatility, `bvxmn` was the center of the lowest volatility level, and bvxmx was the center of the highest volatility level. The settings used for the scaling parameters in the above formula were 37 for nvx, 0.20 for bvxmn, and 2.00 for bvxmx. These parameter values imply an initial bin centered at 20% estimated future volatility, an increment of 0.05 to get from one volatility level or bin to the next, and a final bin centered at 200% volatility. Time left (tl) was stepped from 5 to 20 in increments of 1. For each time left, the second of two array indices (itl) was determined. A time left of 5 indexed column 0 (itl = 0), 6 indexed column 1, and so on up to a time left of 20, which indexed column 15 (itl = 15). In addition to the array column index, the terminal price of the stock, scaled to a nominal $100 at the reference bar, was computed as

```
tfin = exp(-(k/252.0)*rfi[i])*100.0*cls[i+k]/cls[i]
```

where tfin was the desired terminal stock price, 252.0 was the typical number of bars in a year, k was the time left (in bars), rfi[i] was the risk-free interest at the reference bar, cls[i+k] was the close at the k-th bar following the reference, and cls[i] was the close at the reference bar. When performing this calculation, k was set to (int)tl, the time left. Case counts were then accumulated in one array and sums of terminal stock prices in another array, using ivx and itl as the row and column indices, respectively.

Once all possible values of time left were considered, the next valid reference bar was chosen and the sequence repeated. When the supply of valid reference bars for the given stock was exhausted, data for another stock was retrieved from the database. Processing continued in this fashion until all the stock data had been analyzed. The result was one array that contained the number of instances that each combination of time left and estimated future volatility were observed, and another array that contained the sum of the terminal prices for each combination of time left and volatility. The case counts and sums in the two arrays were used to determine the mean, or statistical expectation, for terminal stock price as a function of time left and estimated future volatility. The mean or expectation is, in this context, a measure of trend when it is compared to the nominal $100 initial stock price.

Once the trend was determined for each combination of time left and volatility, another pass was made over the stock data. Again, a stock was selected from the database, a valid reference bar was chosen, and near-future volatility was estimated based on data that was historical with respect to the reference. At this point, time left was stepped from 5 to 20, in increments of 1, and strike price was stepped from $75 to $125, in increments of $2.50.

The volatility estimate, time left, and strike price were used to create three array indices. These three indices were then employed to address the elements of two arrays. In the first of the two arrays, the instance counts were accumulated. In the second of the two arrays, the terminal prices of theoretical options (in this case, calls) trading on the stocks were summed.

The terminal price of a call having a given strike and remaining time was computed as follows: First, the terminal stock price, discounted for interest, was calculated as in the previous pass. Next, an adjustment was made for trend. This involved subtracting the quantity `100.0–aefin(ivx,itl)` from the terminal stock price; the array element `aefin(ivx,itl)` was the expectation, computed in the previous pass, for terminal stock price at a given level of volatility and time. The terminal price of the call (not the stock) was then easily found as the greater of zero or `s–sk`, where `s` was the terminal stock price, adjusted for trend, and `sk` was the strike.

After the element addressed in the first array by the three indices had been incremented and the corresponding element in the second array had the terminal option price added to its value, the next combination of time left and strike were considered. Once all times and strikes were analyzed, the next valid reference bar was chosen and, when there were no more such bars for the selected stock, the next stock was selected and its data retrieved. The sequence continued until all stocks, bars, times, and strikes had been processed.

The results were two three-dimensional arrays that possessed all the information needed to compute the empirical expectation for the terminal price of a call as a function of volatility, time left, and strike—as reflected by the array indices. It should be noted that the expectation thus computed is identical to that which would have resulted from application of the conditional distribution methodology, given a sufficiently fine-grained binning of terminal stock price. The sum of the terminal option prices in each element of the second array was over-written with the corresponding expectation obtained by dividing that sum by the count contained in the equivalent element of the first array.

The final step in the first computational unit was to write the expectations, along with the volatility, time left, and strike variables, to a new fact file against which a pricing model could be trained. This was accomplished by stepping each of the three array indices through its full range of values. For each possible combination of index values, the corresponding levels of volatility, time left, and strike were written to the fact file, followed by the terminal price expectation obtained from the array element addressed by the given combination of index values. Each fact, therefore, consisted of three inputs and one target. Facts were written with strike moving most rapidly and volatility least rapidly.

***Second Computational Unit*** In this unit, a new set of facts was constructed from the set prepared in the first computational unit by replacing the three independent variables in each original fact with new, more numerous sets of input variables. The new input variables were computed as cross-products of various orders of Chebyshev Polynomials evaluated at the original input variables. In mathematical notation,

$$x_{ij} = T_i(.2\sqrt{5}, 2\sqrt{20}, v\sqrt{t}) \cdot T_j(75, 125, k) \tag{7.9}$$

where the $x_{ij}$ were the new input variables, $v$ was the original volatility input, $t$ was the time left, and $k$ was the strike price. In this equation, $T_n(a,b,d)$ represents the $n$-th order Chebyshev Polynomial, rescaled for $a \le d \le b$ and evaluated at $d$. An $x_{ij}$ was computed for all $i$ and $j$ that satisfied the following inequalities: $0 \le i \le 13$, $0 \le j \le 17$, and $i + j \le 17$. Consequently, there were 160 $x_{ij}$ or new inputs. Since $T_0 \equiv 1$ it was not necessary to explicitly compute the new variables that were not true cross-products.

The reader may wonder why volatility and time were combined into a single value. The reason was simplification.

Were these two variables not combined, three distinct polynomials would have had to be evaluated and multiplied; in other words,

$$x_{ijl} = T_i(.2,2,v)\cdot T_j(5,20,t)\cdot T_l(75,125,k) \tag{7.10}$$

This would have led to a much larger number of new variables in the model. In fact, 802 variables would have been necessary for a model of roughly the same order, i.e., with $0 \le i \le 13$, $0 \le j \le 13$, $0 \le l \le 17$ and $i + j + l \le 17$.

The target in each expanded fact was the same terminal option price expectation found in the original, unexpanded fact on which the new fact was based.

***Third Computational Unit*** In this unit, a standard multiple linear regression was carried out on the fact set generated by the second computational unit. In other words, a set of $w_{ij}$ were determined in order to obtain an estimate of option premium

$$\begin{aligned} \hat{y} &= \sum_{ij} w_{ij} x_{ij} \\ &= \sum_{ij} w_{ij} T_i(.2\sqrt{5}, 2\sqrt{20}, v\sqrt{t}) T_j(25,125,k) \end{aligned} \tag{7.11}$$

for each fact that minimized the sum of the squared errors taken over all of the facts. The code implementing the regression was carefully written to minimize numerical instability and the accumulation of roundoff errors in the face of large, ill-conditioned matrices.

The third computational unit, the regression procedure, produced two output files. One file contained the regression report, which included various statistics, such as the regression weights (the $w_{ij}$ in the above discussion), variance inflation factors, and $t$-tests. The other file was written in standard fact file format and contained the regression-estimated option premiums and the target premiums (the terminal option price expectations) against which the model was trained or fitted. The output produced by the third computational block was loaded into an Excel spreadsheet for visualization, further analysis, and presentation.

### Results

A fairly good fit of the model to the data was achieved. The correlation between the actual target and the model's estimate of the target was 0.99936, the root-mean-square error was $0.2920, and the maximum absolute error was $1.7405, over a total of 12,432 cases or data points.

When compared to the results obtained from fitting a similar model to Black-Scholes in Study 2, these error and correlation statistics may not look so great. One must remember, however, that the facts or data points in the current study were derived from real stock prices and not from a "perfect," noiseless equation. Since the number of stock prices on which each fact was based was finite, the estimate of future option price expectation, which forms each fact's target, is just an estimate—a number that will almost always be too high or too low, sometimes dramatically so, and only rarely be correct within the intended precision of the model. Even if the model being tested were perfect, precisely capturing the true relationships underlying the data, less than perfect correlation and error statistics would be anticipated.

A more important issue in assessing the model, rather than whether the statistics look great, merely acceptable, or even bad, is whether the poor statistical showing reflects systematic or model error, or whether it is a consequence of random variation in the facts, i.e., noise, which the model has wisely ignored or "smoothed out." The tables and chart presented below help address that issue.

Table 7–4 shows the polynomial regression estimate of fair premium (PR), the Black-Scholes estimate (BS), and the target computed from the price activity of real stocks (CD) for each of three combinations of time and volatility over a range of strike prices (STRIKE).

For a time left of 5 bars and volatility of 50%, the polynomial regression has no error greater than $0.02 for in-the-money (strike < $97.50) and out-of-the-money (strike > $102.50) options. Black-Scholes, on the other hand, has six errors greater than $0.02, four of which are greater than $0.05. For at-the-money options (strikes between $97.50 and $102.50), Black-Scholes does somewhat better than the polynomial model. Given that there are numerous data points contributing to the target

**TABLE 7–4**

Theoretical Call Premiums Computed with Polynomials, Black-Scholes, and Conditional Distributions as a Function of Strike for Three Combinations of Time and Volatility

| | TIME = 5 VLTY = 0.25 | | | TIME = 15 VLTY = 0.50 | | | TIME = 15 VLTY = 1.00 | | |
|---|---|---|---|---|---|---|---|---|---|
| STRIKE | PR | BS | CD | PR | BS | CD | PR | BS | CD |
| 75.00 | 25.00 | 25.00 | 25.00 | 25.19 | 25.03 | 25.20 | 26.42 | 26.23 | 26.40 |
| 77.50 | 22.49 | 22.50 | 22.50 | 22.76 | 22.57 | 22.77 | 24.29 | 24.14 | 24.27 |
| 80.00 | 20.00 | 20.00 | 20.01 | 20.36 | 20.14 | 20.37 | 22.24 | 22.13 | 22.23 |
| 82.50 | 17.50 | 17.50 | 17.51 | 18.01 | 17.77 | 18.02 | 20.29 | 20.21 | 20.27 |
| 85.00 | 15.00 | 15.00 | 15.01 | 15.72 | 15.48 | 15.74 | 18.43 | 18.39 | 18.41 |
| 87.50 | 12.51 | 12.50 | 12.52 | 13.51 | 13.29 | 13.54 | 16.67 | 16.68 | 16.66 |
| 90.00 | 10.05 | 10.00 | 10.04 | 11.43 | 11.24 | 11.46 | 15.03 | 15.07 | 15.02 |
| 92.50 | 7.59 | 7.52 | 7.58 | 9.49 | 9.36 | 9.53 | 13.50 | 13.57 | 13.50 |
| 95.00 | 5.19 | 5.11 | 5.19 | 7.75 | 7.66 | 7.77 | 12.09 | 12.18 | 12.09 |
| 97.50 | 3.06 | 2.98 | 3.01 | 6.21 | 6.16 | 6.23 | 10.79 | 10.89 | 10.81 |
| 100.00 | 1.49 | 1.41 | 1.37 | 4.90 | 4.87 | 4.90 | 9.61 | 9.71 | 9.65 |
| 102.50 | 0.58 | 0.51 | 0.53 | 3.82 | 3.78 | 3.80 | 8.53 | 8.63 | 8.59 |
| 105.00 | 0.21 | 0.14 | 0.20 | 2.94 | 2.88 | 2.90 | 7.56 | 7.65 | 7.63 |
| 107.50 | 0.10 | 0.03 | 0.08 | 2.24 | 2.16 | 2.19 | 6.68 | 6.76 | 6.77 |
| 110.00 | 0.05 | 0.00 | 0.04 | 1.70 | 1.59 | 1.64 | 5.89 | 5.96 | 5.99 |
| 112.50 | 0.01 | 0.00 | 0.02 | 1.29 | 1.15 | 1.23 | 5.18 | 5.24 | 5.30 |
| 115.00 | 0.00 | 0.00 | 0.01 | 0.98 | 0.82 | 0.92 | 4.55 | 4.60 | 4.68 |
| 117.50 | 0.00 | 0.00 | 0.01 | 0.74 | 0.57 | 0.69 | 3.99 | 4.02 | 4.12 |
| 120.00 | –0.01 | 0.00 | 0.00 | 0.57 | 0.40 | 0.51 | 3.49 | 3.51 | 3.62 |
| 122.50 | –0.01 | 0.00 | 0.00 | 0.43 | 0.27 | 0.39 | 3.05 | 3.06 | 3.19 |
| 125.00 | –0.01 | 0.00 | 0.00 | 0.33 | 0.18 | 0.30 | 2.67 | 2.66 | 2.80 |

expectations for this combination of time and volatility, “noise” in the target is minimal and deviations of model estimates (whether polynomial or Black-Scholes) from target values represent primarily model error. Black-Scholes gives a better fit for these strikes because the premiums of near- and at-the-money options are least affected by deviations from normality in the underlying stock price distributions (i.e., by violation of the Black-Scholes assumption that returns have a log-normal distribution) and because the polynomial model has some difficulty handling the sharp turn, where the strike moves through the at-the-money region, in an otherwise gentle curve that describes the relationship of premium to strike. The root-mean-square error over all strikes was $0.032 for the polynomial regression model and $0.033 for Black-Scholes; the average absolute errors were $0.02 and $0.025, respectively.

With 15 bars left and a volatility of 50%, the root-mean-square error was $0.038 for the polynomial model and $0.151 for Black-Scholes. Measured by the average absolute deviation, these errors were $0.033 and $0.130, and by the maximum absolute deviation, $0.07 and $0.26. In this instance, the polynomial model dramatically outperforms Black-Scholes.

Examining the figures in Table 7–4 reveals systematic error in both the polynomial and Black-Scholes models. The polynomial model tends to underestimate the value of in-the-money calls and tends to overestimate the value of out-of-the-money calls. Black-Scholes underestimates the value of both in-the-money and out-of-the-money calls, with the underestimation being noteworthy in the out-of-the-money case and very severe in the in-the-money case. The underestimation of the value of both in-the-money and out-of-the-money options by Black-Scholes is a model error; it is exactly what might be expected from the application of a pricing model that assumes a log-normal distribution of returns to a market that exhibits a longer-tailed, leptokurtic distribution of price movements. The model error seen with the polynomial regression is much smaller; it reflects the limitation imposed by model order on the fit that can be achieved to the data.

When options with 15 bars of life remaining on stocks that have a volatility of 100% were examined, the root-mean-square error was found to be $0.080 and $0.088 for the polynomial and Black-Scholes models, respectively. The corresponding figures for the average absolute error were $0.061 and $0.076; for the maximum absolute error they were $0.14 and $0.17. Again, the polynomial model performs better than Black-Scholes, although less dramatically so than at a volatility level of 50%. The reduced contrast between the polynomial model and Black-Scholes is probably a result of the greater noise or estimation error in the data points. Greater noise can be assumed to exist in the facts corresponding to a volatility of 100% than in those

**TABLE 7–5**

Theoretical Call Premiums Computed with Polynomials, Black-Scholes, and Conditional Distributions as a Function of Volatility for Three Strikes

| | TIME = 10 STRIKE = 80 | | | TIME = 10 STRIKE = 100 | | | TIME = 10 STRIKE = 120 | | |
|---|---|---|---|---|---|---|---|---|---|
| VLTY | PR | BS | CD | PR | BS | CD | PR | BS | CD |
| 0.20 | 20.00 | 20.00 | 20.02 | 1.65 | 1.59 | 1.71 | –0.01 | 0.00 | 0.01 |
| 0.25 | 20.02 | 20.00 | 20.02 | 2.05 | 1.99 | 1.98 | 0.02 | 0.00 | 0.02 |
| 0.30 | 20.05 | 20.00 | 20.04 | 2.43 | 2.38 | 2.35 | 0.05 | 0.00 | 0.03 |
| 0.35 | 20.07 | 20.00 | 20.06 | 2.80 | 2.78 | 2.74 | 0.07 | 0.01 | 0.05 |
| 0.40 | 20.09 | 20.01 | 20.09 | 3.17 | 3.18 | 3.15 | 0.10 | 0.03 | 0.10 |
| 0.45 | 20.13 | 20.02 | 20.14 | 3.56 | 3.58 | 3.58 | 0.16 | 0.08 | 0.16 |
| 0.50 | 20.18 | 20.04 | 20.19 | 3.96 | 3.97 | 3.98 | 0.25 | 0.14 | 0.25 |
| 0.55 | 20.25 | 20.08 | 20.28 | 4.38 | 4.37 | 4.39 | 0.37 | 0.24 | 0.38 |
| 0.60 | 20.34 | 20.13 | 20.35 | 4.80 | 4.77 | 4.84 | 0.52 | 0.36 | 0.55 |
| 0.65 | 20.44 | 20.20 | 20.42 | 5.21 | 5.16 | 5.21 | 0.70 | 0.51 | 0.72 |
| 0.70 | 20.54 | 20.29 | 20.51 | 5.62 | 5.56 | 5.58 | 0.89 | 0.68 | 0.89 |
| 0.75 | 20.65 | 20.40 | 20.62 | 6.01 | 5.96 | 6.00 | 1.10 | 0.88 | 1.12 |
| 0.80 | 20.76 | 20.52 | 20.73 | 6.39 | 6.35 | 6.35 | 1.31 | 1.10 | 1.31 |
| 0.85 | 20.89 | 20.66 | 20.86 | 6.77 | 6.75 | 6.71 | 1.54 | 1.33 | 1.50 |
| 0.90 | 21.03 | 20.82 | 20.97 | 7.14 | 7.15 | 7.11 | 1.77 | 1.58 | 1.78 |
| 0.95 | 21.18 | 20.99 | 21.12 | 7.51 | 7.54 | 7.41 | 2.02 | 1.85 | 1.98 |
| 1.00 | 21.35 | 21.17 | 21.27 | 7.89 | 7.94 | 7.85 | 2.27 | 2.13 | 2.32 |
| 1.05 | 21.53 | 21.37 | 21.51 | 8.27 | 8.33 | 8.27 | 2.53 | 2.42 | 2.54 |
| 1.10 | 21.72 | 21.57 | 21.60 | 8.65 | 8.73 | 8.47 | 2.80 | 2.72 | 2.68 |
| 1.15 | 21.93 | 21.79 | 21.82 | 9.04 | 9.12 | 8.96 | 3.07 | 3.03 | 3.05 |
| 1.20 | 22.14 | 22.01 | 21.95 | 9.42 | 9.52 | 9.17 | 3.35 | 3.35 | 3.10 |
| 1.25 | 22.35 | 22.24 | 22.22 | 9.80 | 9.91 | 9.60 | 3.63 | 3.68 | 3.47 |
| 1.30 | 22.57 | 22.48 | 22.46 | 10.17 | 10.31 | 10.04 | 3.91 | 4.01 | 3.81 |
| 1.35 | 22.80 | 22.73 | 22.63 | 10.54 | 10.70 | 10.56 | 4.19 | 4.35 | 4.26 |
| 1.40 | 23.02 | 22.98 | 22.94 | 10.90 | 11.09 | 10.89 | 4.48 | 4.70 | 4.53 |
| 1.45 | 23.25 | 23.24 | 23.17 | 11.25 | 11.49 | 11.28 | 4.77 | 5.05 | 4.88 |
| 1.50 | 23.49 | 23.51 | 23.77 | 11.61 | 11.88 | 12.14 | 5.07 | 5.40 | 5.79 |
| 1.55 | 23.73 | 23.77 | 24.11 | 11.96 | 12.27 | 12.07 | 5.38 | 5.76 | 5.44 |
| 1.60 | 23.98 | 24.05 | 23.42 | 12.32 | 12.67 | 11.49 | 5.69 | 6.12 | 5.14 |
| 1.65 | 24.24 | 24.32 | 23.82 | 12.68 | 13.06 | 12.15 | 6.01 | 6.49 | 5.79 |
| 1.70 | 24.50 | 24.60 | 24.53 | 13.03 | 13.45 | 13.06 | 6.33 | 6.86 | 6.41 |
| 1.75 | 24.75 | 24.89 | 25.41 | 13.39 | 13.84 | 14.41 | 6.64 | 7.23 | 7.27 |

**TABLE 7–5**

(*Continued*)

| | TIME = 10 STRIKE = 80 | | | TIME = 10 STRIKE = 100 | | | TIME = 10 STRIKE = 120 | | |
|---|---|---|---|---|---|---|---|---|---|
| VLTY | PR | BS | CD | PR | BS | CD | PR | BS | CD |
| 1.80 | 25.01 | 25.17 | 25.22 | 13.73 | 14.23 | 14.32 | 6.95 | 7.61 | 7.49 |
| 1.85 | 25.25 | 25.46 | 25.29 | 14.05 | 14.62 | 14.44 | 7.23 | 7.99 | 7.71 |
| 1.90 | 25.48 | 25.76 | 25.87 | 14.36 | 15.01 | 15.43 | 7.50 | 8.37 | 8.52 |
| 1.95 | 25.69 | 26.05 | 25.07 | 14.64 | 15.41 | 14.29 | 7.75 | 8.75 | 7.25 |
| 2.00 | 25.89 | 26.35 | 26.07 | 14.89 | 15.80 | 15.37 | 7.97 | 9.13 | 8.06 |

corresponding to a volatility of 50% due to the far smaller number of distinct returns contributing to each of the latter facts.

Errors were substantial for out-of-the-money options, but quite small for in-the-money options, when appraisals of value were made with the polynomial model. Black-Scholes produced sizable errors at both ends of the moneyness spectrum, as well as at some points in between.

Table 7–5 shows the target premiums computed as the terminal option price expectations (CD), together with premiums derived from Black-Scholes (BS) and from the polynomial regression (PR). These premiums are presented for three strikes (an in-the-money strike of \$80, an at-the-money strike of \$100, and an out-of-the-money strike of \$120) over the full range of estimated future volatility (VLTY). In all cases, premiums were for options with 10 bars of life remaining.

Random variation of the target premiums about their "true" values is much more evident in Table 7–5 than in Table 7–4, especially at high levels of volatility. This is because the samples of stock returns that contribute to the expectation estimates for terminal option price become small as volatility increases. For example, at 50% volatility, a sample of 294,275 data points is available for estimating the terminal price expectation of an option. The sample size drops to 31,809 at 100% volatility and to 2,091 at 150% volatility. By the time 200% volatility is reached, a sample of only 153 stock returns is available from the database employed in this study. Despite the greater random variation of the target at higher levels of volatility, systematic pricing errors can still be observed in the figures in Table 7–5.

With in-the-money options (strike of $80), the polynomial model provides a close fit to the data for volatilities up to 85%, overestimates prices for volatility levels between 90% and 140%, and gives visually correct appraisals when volatility is between 145% and 200%—although, in this range of volatility, the premiums are scattered fairly widely about the model-estimated values due to the small samples on which these premiums are based. By way of contrast, Black-Scholes greatly underestimates option worth at lower levels of volatility (between 30% and 100%) and overestimates option value at higher levels of volatility (between 120% and 200%).

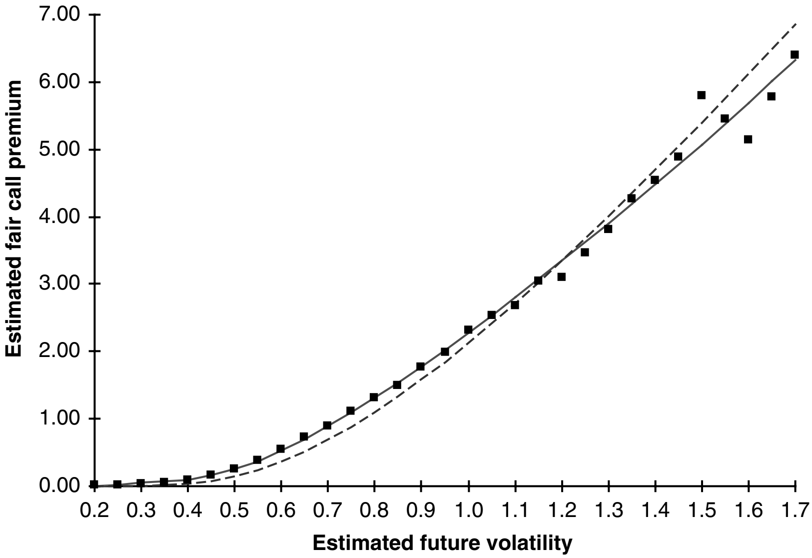
A similar pattern appears with out-of-the-money options. Again, the polynomial model yields accurate premiums and a tight fit (small absolute errors) to the data at lower levels of volatility, underestimates fair value at high levels of volatility, and overestimates option worth over a small range in between. Black-Scholes significantly underestimates the worth of out-of-the-money options at lower levels of volatility and greatly overestimates the value at higher levels. The root-mean-square error over all volatility levels was $0.228 for the polynomial model and $0.269 for Black-Scholes, in the case of in-the-money call options. The corresponding figures were $0.349 and $0.367 for at-the-money options and $0.295 and $0.405 for out-of-the-money options. The average absolute errors were $0.141 and $0.185, $0.203 and $0.228, and $0.163 and $0.257, for the polynomial and Black-Scholes models over the three strikes or levels of moneyness; the corresponding maximum absolute errors were $0.66 and $0.98, $1.07 and $1.18, and $1.02 and $1.50.

As is easily seen, the overall error, regardless of the statistic by which it was measured, is worse for Black-Scholes than for the polynomial regression model. The relatively poor performance of Black-Scholes is exaggerated for out-of-the-money and in-the-money options, where the market's violation of the lognormal assumption of Black-Scholes has the greatest impact.

Finally, Figure 7–3 illustrates graphically the relationships between volatility and premium as determined by the direct

**FIGURE 7–3**

Call Prices Estimated by Polynomial Regression, Black-Scholes, and Terminal Expectation, as a Function of Volatility




evaluation of option terminal price expectation (small square markers), by the polynomial regression model (solid line), and by Black-Scholes (broken line). The data in the figure are for a strike price of $120 and a time left of 10 bars.

It is easy to observe that the polynomial model (solid line) provides a very good fit to the data points (small rectangles) at lower levels of volatility. As volatility rises, the data points become more dispersed about the solid line, but that line, derived from the polynomial model, does seem to trace out something close to the true underlying relationship that probably exists between premium and volatility in the market—i.e., the relationship that would have been tracked by the square markers were they based on samples large enough to eliminate the visible estimation error. The Black-Scholes model (broken line), however, systematically underprices the calls for volatility levels below about 100% and overprices them for volatility levels greater than about 120%.

As evident in Figure 7–3, the polynomial model does an exceptional job of smoothing out the noise in the empirical data and, thus, of revealing the true relationship between fair premium and estimated future volatility.

What does the model actually look like? Table 7–6 contains the regression weights for various terms in the model. In the table, VT(0) through VT(13) represent Chebyshev Polynomials of orders 0 to 13, evaluated at $av\sqrt{t} - b$, where $v$ is the volatility, $t$ is the time left (in bars), and $a$ and $b$ are scaling constants chosen to make the value of the expression fall between –1 and +1. The SK(0) through SK(17) represent another set of Chebyshev Polynomials evaluated at $aK + b$, where $K$ is the strike price and $a$ and $b$ are again scaling coefficients (different than those above, but serving the same purpose). Each regression weight in Table 7–6 is applied to the cross-product of the polynomial that corresponds to the table row with the polynomial that corresponds to the table column. The weighted cross-product terms are then added to obtain the model output, an estimate of the fair premium of a call.

Scanning Table 7–6 reveals that the largest weights appear in the upper-left corner and are associated with cross-products of Chebyshev Polynomials having low order. These cross-products contribute the most to the model. As the focus moves away from the upper-left corner of Table 7–6, the size of the weights declines, implying smaller contributions to the model's output from cross-products involving higher order polynomials. In fact, cross-products with weights less than, say, 0.001, could probably be dropped from the model with little impact on the overall fit. Chebyshev Polynomials (and their cross-products) have the neat property of evaluating to a number between –1 and +1. This means that the maximum absolute error introduced by omitting a term can be no greater than the absolute size of that term's regression weight.

## Discussion

This study clearly demonstrates that a pricing model based on a polynomial regression can do a better job of describing the relationship between variables such as time, volatility, and

**TABLE 7–6**

Regression Weights for a Simple Option Pricing Model Based on Fitting Terminal Option Price Expectation Data with Chebyshev Polynomials

| | VT(0) | VT(1) | VT(2) | VT(3) | VT(4) | VT(5) | VT(6) |
|---|---|---|---|---|---|---|---|
| SK(0) | 13.2336 | 5.4363 | –0.5219 | –0.8564 | –0.1078 | –0.0146 | 0.0432 |
| SK(1) | –11.7348 | 0.7639 | –0.1758 | –0.1526 | 0.0040 | –0.0330 | 0.0182 |
| SK(2) | 2.7245 | –1.9231 | 0.6140 | –0.0171 | –0.0424 | 0.0446 | 0.0025 |
| SK(3) | –0.0731 | –0.0301 | 0.0419 | –0.0265 | –0.0075 | 0.0103 | –0.0128 |
| SK(4) | –0.2263 | 0.3684 | –0.2336 | 0.1014 | –0.0295 | –0.0052 | 0.0025 |
| SK(5) | 0.0107 | 0.0002 | 0.0012 | 0.0132 | 0.0012 | 0.0026 | 0.0077 |
| SK(6) | 0.0617 | –0.1154 | 0.0866 | –0.0544 | 0.0294 | –0.0110 | 0.0029 |
| SK(7) | –0.0024 | 0.0025 | –0.0006 | –0.0020 | 0.0011 | –0.0004 | –0.0006 |
| SK(8) | –0.0237 | 0.0484 | –0.0366 | 0.0273 | –0.0180 | 0.0114 | –0.0041 |
| SK(9) | 0.0013 | –0.0017 | –0.0008 | –0.0005 | –0.0009 | –0.0001 | –0.0004 |
| SK(10) | 0.0112 | –0.0133 | 0.0196 | –0.0168 | 0.0095 | –0.0059 | 0.0040 |
| SK(11) | 0.0015 | 0.0035 | 0.0004 | 0.0013 | 0.0007 | 0.0005 | 0.0000 |
| SK(12) | –0.0045 | 0.0177 | –0.0062 | 0.0043 | –0.0047 | 0.0060 | |
| SK(13) | 0.0016 | 0.0006 | 0.0001 | 0.0000 | 0.0001 | | |
| SK(14) | 0.0015 | 0.0020 | 0.0036 | –0.0054 | | | |
| SK(15) | 0.0006 | 0.0006 | –0.0006 | | | | |
| SK(16) | –0.0006 | 0.0050 | | | | | |
| SK(17) | 0.0004 | | | | | | |

| | VT(7) | VT(8) | VT(9) | VT(10) | VT(11) | VT(12) | VT(13) |
|---|---|---|---|---|---|---|---|
| SK(0) | 0.0413 | 0.0461 | 0.0579 | 0.0432 | 0.0084 | –0.0069 | –0.0261 |
| SK(1) | 0.0505 | 0.0464 | 0.0631 | 0.0502 | 0.0223 | 0.0093 | 0.0010 |
| SK(2) | 0.0358 | 0.0234 | 0.0127 | 0.0122 | 0.0016 | 0.0010 | 0.0046 |
| SK(3) | –0.0090 | –0.0064 | –0.0124 | –0.0073 | –0.0089 | –0.0065 | –0.0049 |
| SK(4) | –0.0104 | –0.0015 | –0.0048 | –0.0029 | 0.0015 | 0.0008 | –0.0001 |
| SK(5) | 0.0020 | 0.0034 | 0.0026 | 0.0025 | 0.0030 | 0.0020 | |
| SK(6) | 0.0017 | –0.0011 | 0.0015 | –0.0001 | –0.0004 | | |
| SK(7) | 0.0012 | 0.0006 | 0.0013 | 0.0007 | | | |
| SK(8) | 0.0018 | –0.0005 | –0.0006 | | | | |
| SK(9) | –0.0004 | –0.0003 | | | | | |
| SK(10) | –0.0035 | | | | | | |

strike on the one hand, and real-market expectation for an option's terminal price on the other, than can Black-Scholes.

Polynomial models have virtues that go beyond their ability to capture the complex relationships involved in pricing options. Among these virtues are simplicity and transparency, especially when compared to models based on neural networks. Polynomial models also appear to effectively smooth out noise in the data on which they are based. The model developed in this study clearly demonstrates the smoothing ability of a polynomial regression of appropriate order. Finally, polynomial models are easy to train: the required regression weights can be computed in minutes, even for enormous data sets, rather than the hours or even days that are sometimes required for a neural model.

The one irksome problem with polynomial models is the combinatorial explosion in the number of coefficients that occurs with increasing numbers of model inputs. Neural networks suffer less from this problem, although they may consume more degrees of freedom when only a few inputs are necessary. Neural networks, therefore, are an alternative that may be worth considering when there are many independent variables to be included in the model. They are examined in Study 4.

## STUDY 4: BASIC NEURAL PRICING MODELS

Can simple neural networks price real-market options? In other words, can they provide meaningful estimates of terminal option price expectation? How good are neural networks at capturing essential pricing relationships whilst ignoring the inevitable noise in the data on which they must be trained? Finally, how does the performance of a neural model compare to that of a polynomial model, like the one explored in Study 3? These are the questions considered in the current investigation.

### Method

The same fact set that was prepared with the first computational unit in Study 3 was used in the current one. Each of the 12,432 facts in the fact set consisted of figures for three independent variables and one target or dependent variable. The three independent variables were estimated future volatility, time remaining (in bars), and strike price (in dollars). Estimated future volatility ranged from 20% to 200% in increments of 5%; time left ranged from 5 to 20 bars in increments of 1 bar; and strike price ranged from $75.00 to $125.00 in increments of $2.50. The target was the expected (mean) option price at option expiration, computed from detrended stock returns, given the volatility, time, and strike. For full details regarding preparation of the fact set, consult the discussion of methodology in Study 3, above.

Instead of creating another fact set by expanding the inputs in terms of Chebyshev Polynomials and then performing a regression analysis on the expanded facts (as was done in Study 3), the facts were directly employed to train a neural network. The neural network was created and trained using the N-Train neural network development package.

The steps in creating and training the model were as follows. First, a default scaling control file was created using the `makescl` command with arguments specifying three independent variables and one dependent (target) variable. The facts were then scaled with the `scale` command, which produced the file `train.scl.` This file (the "internal scaling file," in N-Train parlance) was then opened in a text editor and the upper limit for the network's output was changed from 30.876 to 35.000 in order to avoid saturation at the high end of the premium spectrum. Although 30.876 was the largest target premium found in the fact set, it does not, in any way, represent a hard upper limit on the premium of an option. On the other hand, zero does represent a hard lower limit on option premium and so this datum in `train.scl` was not altered. After editing the internal scaling file, the `getfacts` command was issued to load the fact set into the N-Train program. Once loaded, the facts were thoroughly shuffled using the `shuffle` command.

The next step was to employ the `newnet` command to create a four-layer neural network with three active inputs, 26 active second- layer neurons, eight active third-layer neurons, and 1 output. The network size and architecture were the same as for the larger neural model tested against Black-Scholes in Study 1, earlier in this chapter. Using `setparms`, the learning rate was set to 1, the maximum run count was set to 99,999, and the momentum factor was set to 1 (no momentum). The same command was also used to randomize the connection weights with a seed of 93 and dispersion of 1, using random numbers having a Gaussian distribution. Training was then initiated with the `train` command. At various points in the lengthy training process, adjustments were made to the learning rate and to the layer-specific learning rate multipliers in an effort to speed convergence. The network was trained to full convergence; a process that took many tens of thousands of training passes over the fact set, not to mention a quite substantial amount of processor time. Finally, the neural network was polished by performing some additional training passes at a very low learning rate. When training was complete, the trained network was run on the fact set using `runnet`. This produced a file that contained, for each fact, the neural network's estimate of the target's premium (on the basis of the model's inputs), followed by the target premium itself. Various statistics were computed, tables prepared, and charts drawn using the data in the `runnet` output file and in the original fact file.

### Results

The trained neural model produced estimates of target premium that, overall, were much closer to the actual targets than the estimates produced by the polynomial model examined in Study 3. The correlation between the model estimate and the actual target, computed over all facts, was 0.999914; the root-mean-square error was $0.1073 (compare this to $0.292 for the polynomial regression); and the maximum absolute error was $0.6812 ($1.7405 for the polynomial model). Overall, the neural model looks very good. As will be seen, however, the very tight fit to the data was, to some extent, achieved by modeling the noise, as well as the true dependencies underlying the data. In other words, the degrees of freedom consumed by the neural model may have exceeded what the data could supply without risk of undesirable curve-fitting.

Table 7–7 shows fair call premiums generated by the neural model (NN), side-by-side with premiums computed with Black-Scholes (BS), and from actual stock prices (CD, the target). These premiums are shown for three combinations of time left (TIME) and estimated future volatility (VLTY) over a range of

**TABLE 7–7**

Theoretical Call Premiums Computed with a Neural Network, Black-Scholes, and Conditional Distributions as a Function of Strike for Three Combinations of Time and Volatility

| | TIME = 5 VLTY = 0.25 | | | TIME = 15 VLTY = 0.50 | | | TIME = 15 VLTY = 1.00 | | |
|---|---|---|---|---|---|---|---|---|---|
| STRIKE | NN | BS | CD | NN | BS | CD | NN | BS | CD |
| 75.00 | 24.94 | 25.00 | 25.00 | 25.19 | 25.03 | 25.20 | 26.45 | 26.23 | 26.40 |
| 77.50 | 22.46 | 22.50 | 22.51 | 22.77 | 22.57 | 22.77 | 24.28 | 24.14 | 24.28 |
| 80.00 | 20.00 | 20.00 | 20.01 | 20.39 | 20.14 | 20.37 | 22.22 | 22.13 | 22.23 |
| 82.50 | 17.50 | 17.50 | 17.51 | 18.05 | 17.77 | 18.02 | 20.26 | 20.21 | 20.27 |
| 85.00 | 14.99 | 15.00 | 15.01 | 15.75 | 15.48 | 15.74 | 18.40 | 18.39 | 18.41 |
| 87.50 | 12.51 | 12.50 | 12.52 | 13.54 | 13.29 | 13.54 | 16.65 | 16.68 | 16.66 |
| 90.00 | 10.08 | 10.00 | 10.04 | 11.44 | 11.24 | 11.46 | 15.01 | 15.07 | 15.02 |
| 92.50 | 7.63 | 7.52 | 7.58 | 9.50 | 9.36 | 9.53 | 13.48 | 13.57 | 13.50 |
| 95.00 | 5.19 | 5.11 | 5.19 | 7.74 | 7.66 | 7.77 | 12.08 | 12.18 | 12.09 |
| 97.50 | 3.05 | 2.98 | 3.01 | 6.20 | 6.16 | 6.23 | 10.78 | 10.89 | 10.81 |
| 100.00 | 1.52 | 1.41 | 1.37 | 4.88 | 4.87 | 4.90 | 9.60 | 9.71 | 9.65 |
| 102.50 | 0.65 | 0.51 | 0.53 | 3.78 | 3.78 | 3.80 | 8.52 | 8.63 | 8.59 |
| 105.00 | 0.26 | 0.14 | 0.20 | 2.88 | 2.88 | 2.90 | 7.54 | 7.65 | 7.63 |
| 107.50 | 0.10 | 0.03 | 0.08 | 2.18 | 2.16 | 2.19 | 6.65 | 6.76 | 6.77 |
| 110.00 | 0.04 | 0.00 | 0.04 | 1.63 | 1.59 | 1.64 | 5.86 | 5.96 | 5.99 |
| 112.50 | 0.02 | 0.00 | 0.02 | 1.22 | 1.15 | 1.23 | 5.16 | 5.24 | 5.30 |
| 115.00 | 0.01 | 0.00 | 0.01 | 0.91 | 0.82 | 0.92 | 4.53 | 4.60 | 4.68 |
| 117.50 | 0.00 | 0.00 | 0.01 | 0.68 | 0.57 | 0.69 | 3.98 | 4.02 | 4.12 |
| 120.00 | 0.00 | 0.00 | 0.00 | 0.52 | 0.40 | 0.52 | 3.48 | 3.51 | 3.63 |
| 122.50 | 0.00 | 0.00 | 0.00 | 0.40 | 0.27 | 0.39 | 3.04 | 3.06 | 3.19 |
| 125.00 | 0.00 | 0.00 | 0.00 | 0.31 | 0.18 | 0.30 | 2.66 | 2.66 | 2.80 |

strikes. The data presented are directly comparable to those for the polynomial model that appear in Table 7–4.

When time left is 5 bars and volatility is 20%, the neural model tends to overestimate fair premium for strikes at or below \$87.50, underestimates it for strikes between \$90 and \$107.50, and estimates it perfectly for strikes at or greater than \$110. The error is sizable for options struck between \$100 and \$102.50, where the curve that describes the true relationship between strike and premium turns most sharply. Black-Scholes underestimates fair premium, more or less, for all strikes other than \$100 (exactly at-the-money); the greatest underestimation occurs for options that are moderately in- or out-of-the-money (strikes from \$90 to \$95 and from \$105 to \$110).

The neural model produces estimates of fair premium that match the target values exceedingly well for options having 15 bars of time remaining and trading on stocks with estimated future volatility of 50%. For this combination of volatility and time, no error is greater than \$0.03. Black-Scholes, on the other hand, underestimates fair value, falling particularly short for options that are not at- or near-the-money.

When there are 15 bars left and volatility is 100%, the neural model does all right up to a strike of \$97.50. Black-Scholes does quite poorly over this range of strikes. For strikes greater than \$97.50, the neural network significantly underestimates fair value. Black-Scholes makes errors in both directions over this range of strikes: first, overestimating value, and then underestimating it.

The error statistics for Table 7–7 reveal an interesting pattern: errors for the neural model vary widely with the region of input space in which volatility and time appear. The root-mean-square error for the neural model and for Black-Scholes, for each of the three combinations of time and volatility, are \$0.051 and \$0.033, \$0.017 and \$0.151, and \$0.091 and \$0.088, respectively. Average error has corresponding values of \$0.032 and \$0.025, \$0.014 and \$0.180, and \$0.070 and \$0.076. Finally, the maximum absolute error figures are \$0.15 and \$0.08, \$0.03 and \$0.26, and \$0.15 and \$0.17.

As easily seen from these error statistics, the neural model performs exquisitely on the middle combination of time and volatility, leaving Black-Scholes in the proverbial dust. It performs on par with Black-Scholes (better, in the authors' view, because of good results for strikes less than \$97.50) on the third combination of time and volatility. The neural model's performance is worst for the first time and volatility combination, a combination in which both time and volatility are at the lower limits of their ranges in the training fact set.

Table 7–8 contains premiums estimated with the neural model (NN), Black-Scholes (BS), and from stock return data (CD).

**TABLE 7–8**

Theoretical Call Premiums Computed with a Neural Network, Black-Scholes, and Conditional Distributions as a Function of Volatility for Three Strikes

| | *TIME = 10 STRIKE = 80* | | | *TIME = 10 STRIKE = 100* | | | *TIME = 10 STRIKE = 120* | | |
|---|---|---|---|---|---|---|---|---|---|
| **VLTY** | **NN** | **BS** | **CD** | **NN** | **BS** | **CD** | **NN** | **BS** | **CD** |
| 0.20 | 20.02 | 20.00 | 20.02 | 1.67 | 1.59 | 1.71 | 0.01 | 0.00 | 0.01 |
| 0.25 | 20.04 | 20.00 | 20.02 | 1.97 | 1.99 | 1.98 | 0.01 | 0.00 | 0.02 |
| 0.30 | 20.05 | 20.00 | 20.04 | 2.32 | 2.38 | 2.35 | 0.02 | 0.00 | 0.03 |
| 0.35 | 20.07 | 20.00 | 20.06 | 2.72 | 2.78 | 2.74 | 0.05 | 0.01 | 0.05 |
| 0.40 | 20.09 | 20.01 | 20.09 | 3.14 | 3.18 | 3.15 | 0.09 | 0.03 | 0.10 |
| 0.45 | 20.13 | 20.02 | 20.14 | 3.57 | 3.58 | 3.58 | 0.16 | 0.08 | 0.16 |
| 0.50 | 20.17 | 20.04 | 20.19 | 4.00 | 3.97 | 3.98 | 0.26 | 0.14 | 0.25 |
| 0.55 | 20.23 | 20.08 | 20.28 | 4.42 | 4.37 | 4.39 | 0.40 | 0.24 | 0.38 |
| 0.60 | 20.31 | 20.13 | 20.35 | 4.84 | 4.77 | 4.84 | 0.56 | 0.36 | 0.55 |
| 0.65 | 20.40 | 20.20 | 20.42 | 5.24 | 5.16 | 5.21 | 0.74 | 0.51 | 0.72 |
| 0.70 | 20.50 | 20.29 | 20.51 | 5.63 | 5.56 | 5.58 | 0.93 | 0.68 | 0.89 |
| 0.75 | 20.61 | 20.40 | 20.62 | 6.01 | 5.96 | 6.00 | 1.13 | 0.88 | 1.12 |
| 0.80 | 20.73 | 20.52 | 20.73 | 6.38 | 6.35 | 6.35 | 1.33 | 1.10 | 1.31 |
| 0.85 | 20.86 | 20.66 | 20.86 | 6.75 | 6.75 | 6.71 | 1.54 | 1.33 | 1.50 |
| 0.90 | 20.99 | 20.82 | 20.97 | 7.11 | 7.15 | 7.11 | 1.76 | 1.58 | 1.78 |
| 0.95 | 21.13 | 20.99 | 21.12 | 7.46 | 7.54 | 7.41 | 1.99 | 1.85 | 1.98 |
| 1.00 | 21.28 | 21.17 | 21.27 | 7.82 | 7.94 | 7.86 | 2.23 | 2.13 | 2.32 |
| 1.05 | 21.43 | 21.37 | 21.51 | 8.17 | 8.33 | 8.27 | 2.47 | 2.42 | 2.54 |
| 1.10 | 21.60 | 21.57 | 21.60 | 8.53 | 8.73 | 8.47 | 2.73 | 2.72 | 2.68 |
| 1.15 | 21.79 | 21.79 | 21.83 | 8.90 | 9.12 | 8.96 | 2.99 | 3.03 | 3.05 |
| 1.20 | 21.98 | 22.01 | 21.95 | 9.27 | 9.52 | 9.17 | 3.26 | 3.35 | 3.11 |
| 1.25 | 22.19 | 22.24 | 22.22 | 9.65 | 9.91 | 9.60 | 3.54 | 3.68 | 3.47 |
| 1.30 | 22.41 | 22.48 | 22.46 | 10.02 | 10.31 | 10.04 | 3.84 | 4.01 | 3.81 |
| 1.35 | 22.63 | 22.73 | 22.63 | 10.41 | 10.70 | 10.56 | 4.15 | 4.35 | 4.26 |
| 1.40 | 22.88 | 22.98 | 22.94 | 10.82 | 11.09 | 10.89 | 4.53 | 4.70 | 4.53 |
| 1.45 | 23.24 | 23.24 | 23.17 | 11.34 | 11.49 | 11.28 | 5.01 | 5.05 | 4.88 |
| 1.50 | 23.81 | 23.51 | 23.77 | 12.04 | 11.88 | 12.15 | 5.60 | 5.40 | 5.79 |
| 1.55 | 24.01 | 23.77 | 24.11 | 12.10 | 12.27 | 12.07 | 5.54 | 5.76 | 5.44 |
| 1.60 | 23.34 | 24.05 | 23.42 | 11.37 | 12.67 | 11.49 | 4.98 | 6.12 | 5.14 |
| 1.65 | 23.86 | 24.32 | 23.82 | 12.15 | 13.06 | 12.15 | 5.73 | 6.49 | 5.79 |
| 1.70 | 24.52 | 24.60 | 24.53 | 13.06 | 13.45 | 13.06 | 6.36 | 6.86 | 6.41 |
| 1.75 | 25.43 | 24.89 | 25.41 | 14.38 | 13.84 | 14.41 | 7.18 | 7.23 | 7.27 |

**TABLE 7–8**

*(Continued)*

| | TIME = 10 STRIKE = 80 | | | TIME = 10 STRIKE = 100 | | | TIME = 10 STRIKE = 120 | | |
|---|---|---|---|---|---|---|---|---|---|
| VLTY | NN | BS | CD | NN | BS | CD | NN | BS | CD |
| 1.80 | 25.30 | 25.17 | 25.22 | 14.45 | 14.23 | 14.32 | 7.56 | 7.61 | 7.49 |
| 1.85 | 25.26 | 25.46 | 25.29 | 14.48 | 14.62 | 14.44 | 7.68 | 7.99 | 7.71 |
| 1.90 | 25.48 | 25.76 | 25.87 | 14.95 | 15.01 | 15.43 | 8.28 | 8.37 | 8.52 |
| 1.95 | 25.40 | 26.05 | 25.07 | 14.52 | 15.41 | 14.29 | 7.85 | 8.75 | 7.25 |
| 2.00 | 26.03 | 26.35 | 26.07 | 14.93 | 15.80 | 15.37 | 7.64 | 9.13 | 8.06 |

The premiums are broken down by moneyness and volatility. In-the-money calls are represented by those with a strike of $80, at-the-money by those struck at $100, and out-of-the-money by those having a strike price of $120. Note that the strikes were chosen to fall well within the range of strikes found in the fact set on which the neural model was trained. In all cases, the time left was 10 bars. As with the strikes, the time left was chosen to avoid the edges of the input space, defined by the training facts, where neural networks often have trouble.

The differences in Table 7–8 between premiums derived from the neural model and those constituting the target do not appear to represent systematic pricing errors; rather, they appear to reflect random variation, "noise," in the empirically determined target. Consistent with the noise hypothesis, the errors become larger as volatility increases and the samples of returns on which the target premiums are based become smaller. This contrasts to Black-Scholes, which evidences a systematic pattern of error. At high levels of volatility, Black-Scholes overprices options; this is true for all strikes or levels of moneyness. Black-Scholes underestimates fair premium at low-to-moderate volatility for both in-the-money and out-of-the-money options, but not for at-the-money options.

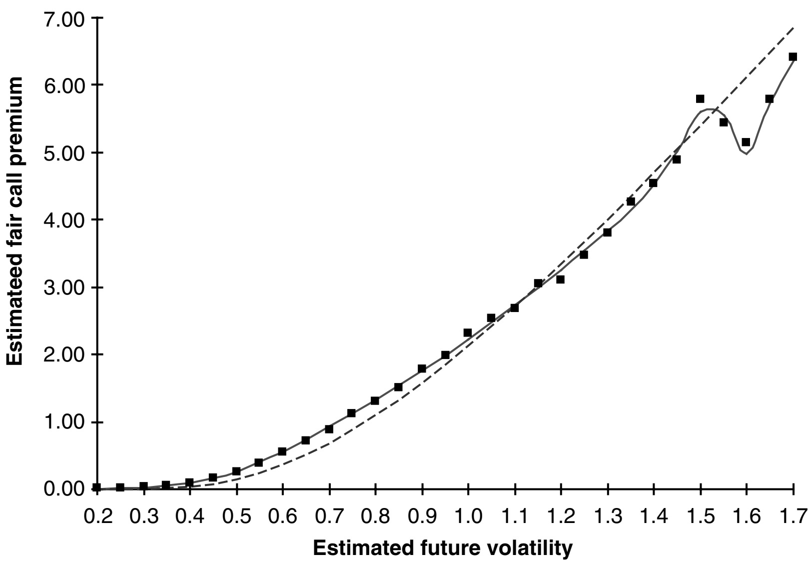
The error statistics associated with the numbers in Table 7–8 demonstrate the superiority of the neural model over Black-Scholes across all three strikes. The root-mean-square errors are \$0.093 versus \$0.269, \$0.128 versus \$0.366, and \$0.144 versus \$0.405, with the neural model and Black-Scholes figures presented side-by-side for each of the three strikes examined in the table. The corresponding average absolute errors are \$0.046 versus \$0.185, \$0.073 versus \$0.228, and \$0.081 versus \$0.257. The largest absolute errors were \$0.40 versus \$0.98, \$0.47 versus \$1.18, and \$0.60 versus \$1.50. Regardless of the measurement used, the error for the neural model is less than half the error experienced with Black-Scholes.

What does the fit of the neural model to the data actually look like? Figure 7–4 provides the answer.

In Figure 7–4, premium estimates (the $y$-axis) derived from the neural network (solid line), Black-Scholes (broken line), and real stock returns (rectangular markers) are plotted against estimated future volatility (the $x$-axis). In all cases, the premiums are for an out-of-the-money call struck at \$120. As is easily seen in Figure 7–4, Black-Scholes underestimates fair premium for

**FIGURE 7–4**

Call Prices Estimated by a Neural Model, Black-Scholes, and Terminal Expectation, as a Function of Volatility



volatility levels less than 110% and overestimates it for volatility levels above that threshold. The incorrect estimation of fair value is likely due to the application of a model that assumes a log-normal distribution of returns to a market that exhibits a leptokurtic distribution. The neural model, on the other hand, fits the data quite well.

In fact, the neural model fits the data too well. It seems that the neural network has not only modeled the true relationship that probably exists between volatility and fair premium, but it has also successfully modeled some of the random variation or noise present in the data. Observe how the neural network handles the hook near the upper-right corner of the chart shown in Figure 7–4. Then take a look at the chart in Figure 7–3 and observe how the multivariate polynomial model, developed in Study 3, responds to the same data. The neural network, unlike the polynomial model, attempts to include the hook in its estimate of fair option premium. This is evidence of an excessively curve-fit solution of the kind often associated with a disproportionate number of free parameters in the model; a tendency to excessively curve-fit the data is a common problem with neural networks.

## Discussion

The current study demonstrates that a neural network can provide a good fit to the data, especially over interior regions of the input space as it is represented in the training facts. Given training data with a sufficiently low noise level, a neural network can learn to accurately price options.

The problem is that a neural network can almost as easily learn to model any noise present in the data. There are various ways around this problem. One way to reduce the unwanted impact of noise on the model is to use the kind of biased estimation that was discussed earlier in this chapter. For example, the target premiums used to train a neural net or fit a polynomial may be intentionally biased toward Black-Scholes to the extent that their error-of-estimate is high. Data points based on large samples and having little estimation error (like those in Figure 7–4 that correspond to volatilities less than about 100%) would be biased hardly at all; those that derive from small samples (like the ones in Figure 7–4 associated with volatilities greater than 150%) would be rather strongly biased in the direction of Black-Scholes. Although not examined here, this method can be very effective; it can make the construction of a complex neural pricing model that does not excessively curve-fit the data feasible.

Another way to construct a model that better describes the true relationships required for pricing options and that is less affected by noise is to incorporate prior knowledge directly into the model. Taken to the limit, this leads to a model based purely on theory and not trained on, or fitted to, the empirical data at all. However, it is possible to go halfway, building some prior knowledge—even if only approximate—into the model, but allowing room for the model to tweak itself on the market on which it is to be used. Such models lie in the realm of hybrid entities. One such model is examined in Study 5.

## STUDY 5: PRICING OPTIONS WITH A HYBRID NEURAL MODEL

This study examines a hybrid model not unlike the one discussed earlier in this chapter and diagrammed in Figure 7–2. The hope is that such a model can provide a good fit to the data with fewer free parameters, and less sensitivity to noise in the training facts, than can a straight neural network or multivariate polynomial regression. A reduction in the number of free parameters is important for more reasons than just to reduce sensitivity to noise. For one thing, a simpler model leaves more headroom for additional inputs, like skew and kurtosis, that might improve pricing accuracy.

The model tested here has exactly the same architecture as a multilayer perceptron (i.e., an everyday, feed-forward neural network), except that the usual sigmoidal output neuron has been replaced with a different kind of neuron—one having some additional inputs and possessing rather unusual properties. What is the nature of this special output neuron? It is a hybridization of a standard neuron with Black-Scholes.

The usual output neuron receives its inputs exclusively from neurons in the previous layer. These inputs are added together. A transfer function (the logistic function is commonly used) is then applied to the resultant sum in order to obtain a number that represents the neuron's output.

By way of contrast, the special hybrid neuron receives two sets of inputs: one set derives from neurons in the previous layer exactly as with a standard neuron; the other set comes from the outside world. The inputs from the outside world are time left, strike price, stock price, and risk-free interest rate. Inside the neuron, the inputs from the previous layer are added together and passed through a logistic transfer function. The output from the logistic function, however, does not become the output from the neuron, as it normally does. Instead, the number obtained from the logistic function is scaled to an appropriate range and passed to the volatility input of a Black-Scholes subsystem. The second set of inputs to the special neuron supply the remaining inputs required by Black-Scholes, which is then used to compute a theoretical premium. The theoretical premium then becomes the hybrid neuron's output.

As a whole, the hybrid neural model, like the hybrid output neuron, has two sets of inputs. The first set (which goes to the standard input layer of the model) consists of volatility, together with any other variables that seem relevant, e.g., skew and kurtosis. The second set (which goes directly to the hybrid neuron) consists of the standard Black-Scholes variables with volatility omitted.

In the current study, the first set of inputs consisted of estimated future volatility, time left (which was also included in the second input set), and normalized moneyness (discussed below). Normalized moneyness was included so that the hybrid entity could model the effect of general market kurtosis, which increases the value of out-of-the-money and in-the-money options, especially at low levels of volatility. Time left was included in the first set of inputs because future volatility was estimated for one look-ahead period and no corrections were made to the volatility estimate for the time remaining before expiration—the intention was to let the neural component of the hybrid model perform any corrections that were necessary.

The inputs in the second set were simply taken from the fact file (time left, strike) or set to constants (stock price, risk-free interest). The target to which the hybrid model was trained was the terminal option price expectation, also taken directly from the fact file.

## Method

The analysis employed the same fact set used in Study 4. Each fact in this fact set was comprised of three independent variables and one target. The independent variables were estimated future volatility, time left (in bars), and strike. The dependent variable was the statistical expectation of the terminal price of a call computed from real stock returns for the specified levels of the independent variables. There were 12,432 facts in the fact set.

Because of the hybrid output neuron, an off-the-shelf neural network package could not be used to train the model; instead, custom software had to be written. The software, comprised of a *Neural Hybrid Options Model* library together with some glue code, was written in ISO-standard C language. Employing a form of gradient-descent as the minimization algorithm, this software made it possible to train the model, i.e., to find connection weights that minimized the sum of the squared errors over all facts in the fact set.

The inputs to the first layer of the model were estimated future volatility, time left, and normalized moneyness. The first two inputs were simply scaled to a mean of zero and a standard deviation of one, and then passed to the first layer of the neural component of the model. Such scaling is the default used by N-Train, the neural network development package employed in Study 4. The last input to the first layer of the model was

$$M = \frac{\ln(s/k)}{v\sqrt{t}} \tag{7.12}$$

where $s$ was the stock price (always \$100), $k$ was the strike price, $v$ was the estimated future volatility, and $t$ was the time left. Normalized moneyness $M$ expresses moneyness in standardized terms, analogous to a $z$-score in statistics. These first-layer inputs were considered sufficient for the neural front-end to determine the "effective" volatility, to compensate for regression to the mean, and to account for distortions in pricing attributable to the long-tailed distribution that characterizes returns from real stocks.

The additional (beyond those received from the previous layer) inputs to the special hybrid neuron were time left, strike price, stock price, and risk-free interest. Time left and strike price need no discussion; they were taken, unaltered, from each fact. Stock price and risk-free interest, however, were supplied internally (by the glue code). Stock price was set to a nominal $100 and risk-free interest was set to zero percent, consistent with the assumptions used in preparing the fact set.

The model's target was the terminal expectation of a call; it was taken directly from each fact and is the same target used in Studies 3 and 4.

As with a standard neural network, the developer must specify the number of input neurons, layers, and other parameters in the model. This test employed a simple three-layer model that had three input neurons, 12 active middle-layer neurons, and one hybrid output neuron. Note the relatively small number of free parameters ($48 = 3 * 12 + 12 * 1$) in this model, when compared to the straight neural network in Study 4 ($294 = 3 * 26 + 26 * 8 + 8 * 1$). The gradient multiplier (something like a learning rate) was set to 0.01 when carrying out the minimization of the sum of the squared errors.

Once the model architecture and parameters were specified, training began. Tens of thousands of passes over the fact set were made in the course of training or optimization, which was carried out to full convergence.

## Results

The correlation between the fully optimized hybrid model's output and the target was 0.999576, the root-mean-square error was $0.2387, and the largest absolute error was $1.3807.

Regardless of how measured, error was less for the hybrid neural model than for the polynomial model, despite the larger number of free parameters in the latter. In other words, the hybrid model gave a better overall fit to the data, yet consumed far fewer degrees of freedom.

Error figures for the hybrid model were greater than for the straight neural model. This was expected, given that the basic neural network had an abundance of free parameters, and that it scrupulously modeled the noise, along with the true pricing relationships that existed in the training set. Clearly, the hybrid model has done a better job of smoothing out or ignoring the noise; hopefully, not at the expense of failing to accurately capture the true relationships that determine the worth of an option.

Table 7–9 shows premiums derived from the hybrid neural model, Black-Scholes, and actual stock returns, for each of three combinations of time left and volatility, and over a range of strikes.

Both the hybrid model and Black-Scholes perform similarly when time and volatility are at their lowest levels. This is reminiscent of the basic neural model, which also had trouble at the edges of the input space as it was defined by the fact set. The largest absolute error for both models was $0.08, the root-mean-square error was $0.032 for the hybrid model versus $0.033 for Black-Scholes, and the average absolute error was $0.024 versus $0.025, respectively. The pattern of errors is also similar for both models.

For the middle combination of time and volatility, the hybrid model decisively out-performs Black-Scholes. Black-Scholes underprices in-the-money options quite severely, while the hybrid model prices them fairly accurately. At the other end of the moneyness spectrum, Black-Scholes again undervalues the options, while the hybrid model overvalues them (but to a lesser degree). At-the-money options are priced correctly by Black-Scholes; these options, however, are slightly overpriced by the hybrid model. The largest absolute error is $0.07 for the hybrid model versus $0.26 for Black-Scholes; the root-mean-square error is $0.053 versus $0.151; and the average error is $0.046 versus $0.130, respectively.

The combination with the highest levels of time and volatility shown in Table 7–9 has the hybrid model performing only slightly better than Black-Scholes. Again, this is reminiscent of the lack of better performance by the basic neural model when compared to Black-Scholes in Study 4. The largest absolute error

**TABLE 7–9**

Theoretical Call Premiums Computed with a Hybrid Model, Black-Scholes, and Conditional Distributions as a Function of Strike for Three Combinations of Time and Volatility

| | *TIME = 5 VLTY = 0.25* | | | *TIME = 15 VLTY = 0.50* | | | *TIME = 15 VLTY = 1.00* | | |
|---|---|---|---|---|---|---|---|---|---|
| **STRIKE** | **NH** | **BS** | **CD** | **NH** | **BS** | **CD** | **NH** | **BS** | **CD** |
| 75.00 | 25.02 | 25.00 | 25.00 | 25.19 | 25.03 | 25.20 | 26.37 | 26.23 | 26.40 |
| 77.50 | 22.51 | 22.50 | 22.50 | 22.76 | 22.57 | 22.77 | 24.26 | 24.14 | 24.27 |
| 80.00 | 20.01 | 20.00 | 20.01 | 20.37 | 20.14 | 20.37 | 22.23 | 22.13 | 22.23 |
| 82.50 | 17.50 | 17.50 | 17.51 | 18.02 | 17.77 | 18.02 | 20.29 | 20.21 | 20.27 |
| 85.00 | 15.00 | 15.00 | 15.01 | 15.74 | 15.48 | 15.74 | 18.45 | 18.39 | 18.41 |
| 87.50 | 12.50 | 12.50 | 12.52 | 13.56 | 13.29 | 13.54 | 16.71 | 16.68 | 16.66 |
| 90.00 | 10.01 | 10.00 | 10.04 | 11.49 | 11.24 | 11.46 | 15.08 | 15.07 | 15.02 |
| 92.50 | 7.52 | 7.52 | 7.58 | 9.57 | 9.36 | 9.53 | 13.55 | 13.57 | 13.50 |
| 95.00 | 5.11 | 5.11 | 5.19 | 7.83 | 7.66 | 7.77 | 12.14 | 12.18 | 12.09 |
| 97.50 | 2.97 | 2.98 | 3.01 | 6.29 | 6.16 | 6.23 | 10.83 | 10.89 | 10.81 |
| 100.00 | 1.43 | 1.41 | 1.37 | 4.97 | 4.87 | 4.90 | 9.64 | 9.71 | 9.65 |
| 102.50 | 0.57 | 0.51 | 0.53 | 3.86 | 3.78 | 3.80 | 8.56 | 8.63 | 8.59 |
| 105.00 | 0.18 | 0.14 | 0.20 | 2.97 | 2.88 | 2.90 | 7.57 | 7.65 | 7.63 |
| 107.50 | 0.04 | 0.03 | 0.08 | 2.26 | 2.16 | 2.19 | 6.69 | 6.76 | 6.77 |
| 110.00 | 0.00 | 0.00 | 0.04 | 1.71 | 1.59 | 1.64 | 5.89 | 5.96 | 5.99 |
| 112.50 | 0.00 | 0.00 | 0.02 | 1.29 | 1.15 | 1.23 | 5.19 | 5.24 | 5.30 |
| 115.00 | 0.00 | 0.00 | 0.01 | 0.98 | 0.82 | 0.92 | 4.56 | 4.60 | 4.68 |
| 117.50 | 0.00 | 0.00 | 0.01 | 0.75 | 0.57 | 0.69 | 4.00 | 4.02 | 4.12 |
| 120.00 | 0.00 | 0.00 | 0.00 | 0.58 | 0.40 | 0.51 | 3.51 | 3.51 | 3.62 |
| 122.50 | 0.00 | 0.00 | 0.00 | 0.46 | 0.27 | 0.39 | 3.08 | 3.06 | 3.19 |
| 125.00 | 0.00 | 0.00 | 0.00 | 0.36 | 0.18 | 0.30 | 2.70 | 2.66 | 2.80 |

is $0.12 for the hybrid model versus $0.17 for Black-Scholes; the root-mean-square error is $0.073 versus $0.088; and the average error is $0.062 versus $0.076, respectively. The hybrid model mildly overprices options for strikes near $90 (modestly in-the-money) and underprices them for strikes greater than $105 (out-of-the-money). Black-Scholes seriously underprices options at both extremes of moneyness and overvalues them in the middle.

Table 7–10 presents hybrid neural, Black-Scholes, and market-derived premiums broken down by strike or moneyness,

**TABLE 7–10**

Theoretical Call Premiums Computed with a Hybrid Model, Black-Scholes, and Conditional Distributions as a Function of Volatility for Three Strikes

| | TIME = 10 STRIKE = 80 | | | TIME = 10 STRIKE = 100 | | | TIME = 10 STRIKE = 120 | | |
|---|---|---|---|---|---|---|---|---|---|
| VLTY | NH | BS | CD | NH | BS | CD | NH | BS | CD |
| 0.20 | 20.09 | 20.00 | 20.02 | 1.72 | 1.59 | 1.71 | 0.00 | 0.00 | 0.01 |
| 0.25 | 20.06 | 20.00 | 20.02 | 2.05 | 1.99 | 1.98 | 0.00 | 0.00 | 0.02 |
| 0.30 | 20.06 | 20.00 | 20.04 | 2.42 | 2.38 | 2.35 | 0.01 | 0.00 | 0.03 |
| 0.35 | 20.07 | 20.00 | 20.06 | 2.81 | 2.78 | 2.74 | 0.04 | 0.01 | 0.05 |
| 0.40 | 20.09 | 20.01 | 20.09 | 3.21 | 3.18 | 3.15 | 0.10 | 0.03 | 0.10 |
| 0.45 | 20.13 | 20.02 | 20.14 | 3.63 | 3.58 | 3.58 | 0.19 | 0.08 | 0.16 |
| 0.50 | 20.18 | 20.04 | 20.19 | 4.05 | 3.97 | 3.98 | 0.30 | 0.14 | 0.25 |
| 0.55 | 20.24 | 20.08 | 20.28 | 4.46 | 4.37 | 4.39 | 0.43 | 0.24 | 0.38 |
| 0.60 | 20.32 | 20.13 | 20.35 | 4.87 | 4.77 | 4.84 | 0.59 | 0.36 | 0.55 |
| 0.65 | 20.40 | 20.20 | 20.42 | 5.27 | 5.16 | 5.21 | 0.76 | 0.51 | 0.72 |
| 0.70 | 20.50 | 20.29 | 20.51 | 5.66 | 5.56 | 5.58 | 0.94 | 0.68 | 0.89 |
| 0.75 | 20.61 | 20.40 | 20.62 | 6.05 | 5.96 | 6.00 | 1.14 | 0.88 | 1.12 |
| 0.80 | 20.72 | 20.52 | 20.73 | 6.43 | 6.35 | 6.35 | 1.35 | 1.10 | 1.31 |
| 0.85 | 20.85 | 20.66 | 20.86 | 6.81 | 6.75 | 6.71 | 1.58 | 1.33 | 1.50 |
| 0.90 | 20.99 | 20.82 | 20.97 | 7.18 | 7.15 | 7.11 | 1.81 | 1.58 | 1.78 |
| 0.95 | 21.13 | 20.99 | 21.12 | 7.54 | 7.54 | 7.41 | 2.04 | 1.85 | 1.98 |
| 1.00 | 21.28 | 21.17 | 21.27 | 7.89 | 7.94 | 7.85 | 2.28 | 2.13 | 2.32 |
| 1.05 | 21.44 | 21.37 | 21.51 | 8.23 | 8.33 | 8.27 | 2.51 | 2.42 | 2.54 |
| 1.10 | 21.62 | 21.57 | 21.60 | 8.57 | 8.73 | 8.47 | 2.75 | 2.72 | 2.68 |
| 1.15 | 21.81 | 21.79 | 21.82 | 8.92 | 9.12 | 8.96 | 2.99 | 3.03 | 3.05 |
| 1.20 | 22.03 | 22.01 | 21.95 | 9.29 | 9.52 | 9.17 | 3.25 | 3.35 | 3.10 |
| 1.25 | 22.26 | 22.24 | 22.22 | 9.68 | 9.91 | 9.60 | 3.53 | 3.68 | 3.47 |
| 1.30 | 22.50 | 22.48 | 22.46 | 10.08 | 10.31 | 10.04 | 3.84 | 4.01 | 3.81 |
| 1.35 | 22.74 | 22.73 | 22.63 | 10.49 | 10.70 | 10.56 | 4.17 | 4.35 | 4.26 |
| 1.40 | 22.95 | 22.98 | 22.94 | 10.86 | 11.09 | 10.89 | 4.50 | 4.70 | 4.53 |
| 1.45 | 23.15 | 23.24 | 23.17 | 11.20 | 11.49 | 11.28 | 4.81 | 5.05 | 4.88 |
| 1.50 | 23.34 | 23.51 | 23.77 | 11.51 | 11.88 | 12.14 | 5.09 | 5.40 | 5.79 |
| 1.55 | 23.54 | 23.77 | 24.11 | 11.80 | 12.27 | 12.07 | 5.35 | 5.76 | 5.44 |
| 1.60 | 23.78 | 24.05 | 23.42 | 12.13 | 12.67 | 11.49 | 5.62 | 6.12 | 5.14 |
| 1.65 | 24.09 | 24.32 | 23.82 | 12.52 | 13.06 | 12.15 | 5.94 | 6.49 | 5.79 |
| 1.70 | 24.49 | 24.60 | 24.53 | 13.01 | 13.45 | 13.06 | 6.34 | 6.86 | 6.41 |
| 1.75 | 24.96 | 24.89 | 25.41 | 13.60 | 13.84 | 14.41 | 6.82 | 7.23 | 7.27 |

**TABLE 7–10**

*(Continued)*

| | TIME = 10 STRIKE = 80 | | | TIME = 10 STRIKE = 100 | | | TIME = 10 STRIKE = 120 | | |
|---|---|---|---|---|---|---|---|---|---|
| VLTY | NH | BS | CD | NH | BS | CD | NH | BS | CD |
| 1.80 | 25.39 | 25.17 | 25.22 | 14.14 | 14.23 | 14.32 | 7.28 | 7.61 | 7.49 |
| 1.85 | 25.59 | 25.46 | 25.29 | 14.38 | 14.62 | 14.44 | 7.46 | 7.99 | 7.71 |
| 1.90 | 25.50 | 25.76 | 25.87 | 14.23 | 15.01 | 15.43 | 7.30 | 8.37 | 8.52 |
| 1.95 | 25.56 | 26.05 | 25.07 | 14.32 | 15.41 | 14.29 | 7.39 | 8.75 | 7.25 |
| 2.00 | 25.99 | 26.35 | 26.07 | 14.90 | 15.80 | 15.37 | 7.94 | 9.13 | 8.06 |

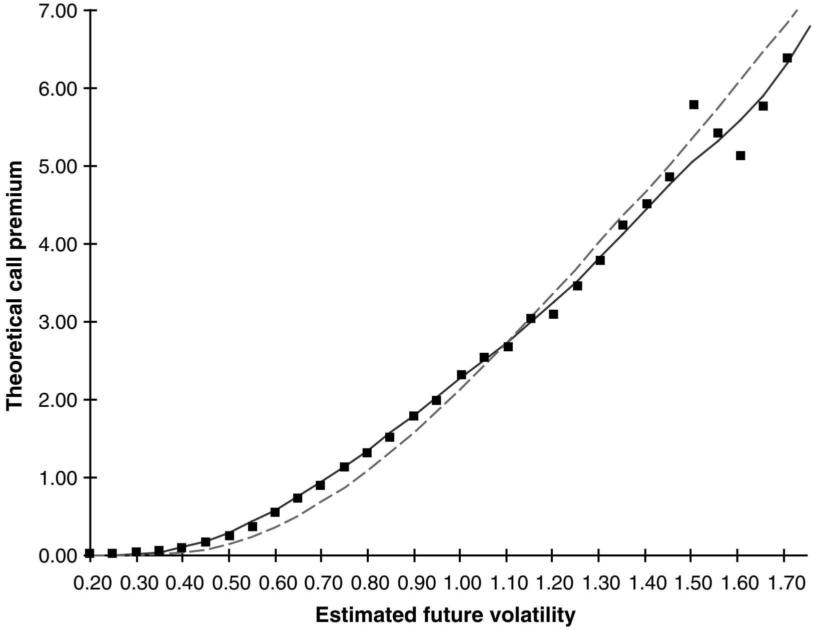
as well as by estimated future volatility. At low levels of volatility (between 30% and 100%), in-the-money calls (strike of \$80) are priced quite accurately by the hybrid model, but are very significantly underpriced by Black-Scholes. Out-of-the-money options (strike of \$120) are also priced much more accurately by the hybrid model than by Black-Scholes, which again yields theoretical premiums that are well below market-derived values. The hybrid model and Black-Scholes handle at-the-money calls (strike of \$100) about equally well; at-the-money calls are somewhat overpriced by the hybrid model, while Black-Scholes errs in both directions.

At high levels of volatility (above 140%), Black-Scholes overprices calls, regardless of strike, with the greatest error occurring with out-of-the-money and at-the-money options. The hybrid model appears to price in-the-money options correctly and tends to price the at-the-money and out-of-the-money options below their probable value. Keep in mind when examining the data in Table 7–10 that, regardless of the model being evaluated, individual data points exhibit large errors in both directions due to small samples at high levels of volatility.

Figure 7–5 displays some of the data in Table 7–10 in visual format. In this figure, the square markers represent the target premiums, the solid line represents the premiums computed with the hybrid neural model, and the broken line represents premiums derived from Black-Scholes. The data shown are for an out-of-the-money call struck at \$120.

**FIGURE 7–5**

Fair Call Premiums from Hybrid Model, Black-Scholes, and Terminal Expectation, as a Function of Estimated Volatility




As seen in Figure 7–5, Black-Scholes premiums fall below probable true value when estimated future volatility is less than 110% and above probable true value when volatility is greater than 110%. The hybrid model can be seen to do a much better job of estimating the true fair premium as reflected in the overall trend of the data points.

The astute observer will note, however, that Delta is not monotonic with respect to volatility for the hybrid model. This is revealed by the fact that the slope of the solid line does not always either increase or remain the same with increasing volatility; at some levels of volatility, the slope decreases. Given a monotonic relationship, the line should be gently turning counterclockwise, increasing its rate of ascent, or continuing on a straight path; it should never be turning clockwise, even slightly. Monotonicity of Delta (the first derivative of premium with respect to volatility) is clearly lost near the hook that appears close to the upper-right corner of the chart.

The neural network in Study 4 also described a relationship between fair value and volatility that became nonmonotonic in Delta near the hook. In that case, not only was monotonicity lost for Delta; it was also lost for the theoretical premium itself. The neural network had simply curve-fit the hook, a probable sampling artifact, as well as any more significant relationship present in the fact set. At least with the hybrid model, the noise appears to have been less aggressively curve-fit.

Why has the issue of monotonicity in the derivative (Delta) been raised? Because the true relationship between fair premium and volatility probably involves a Delta that is monotonic with respect to volatility. The slight "bend" in the solid line near the hook, therefore, represents model error—whether due to overfitting or to other factors. If such conditions as monotonicity of Delta could somehow be included in a hybrid model as prior knowledge, then better results could no doubt be achieved. The problem, of course, is how to minimize the sum of the squared errors subject to such constraints in an efficient manner; a genetic algorithm could do the job, but it would take a very long time to converge to the extent necessary to solve the option pricing problem. Nevertheless, this is something to explore down the road.

### Discussion

Overall, the hybrid neural model performed fairly well, especially considering how much more miserly it was in its consumption of degrees of freedom when compared to either the polynomial regression or basic neural network tested in Studies 3 and 4, respectively.

Although the model examined here was far from perfect, hybrid neural technology represents an interesting approach that, with further effort, can almost certainly be made to yield more accurate option appraisals.

## SUMMARY

This chapter examined the application of such nonlinear modeling technologies as neural networks and polynomial regressions to the problem of pricing options. It was demonstrated that models based on such technologies could be trained to emulate Black-Scholes. Not only could these models emulate Black-Scholes, they could also do significantly better than Black-Scholes when properly trained on real-market data.

Only the standard inputs (volatility, time, strike price, stock price, and interest rate) were used in the models investigated in this chapter; additional inputs, like skew and kurtosis, were not considered or incorporated into the models. It was first necessary to discover whether models based on the nonlinear technologies of concern would perform adequately in the simpler case and manage to account for the general characteristics of real stock prices, like the long-tailed distribution of returns. It has been clearly shown that nonlinear models indeed can be fit to data derived from actual stock prices and can outperform Black-Scholes on such data. The implication is that, once the details are fully worked out, models of the kind studied in this chapter can be extended to encompass additional input variables without great difficulty.

The different approaches studied were shown to have unique strengths and weaknesses, faults and virtues. First, consider polynomial models. Polynomial regression modeling has many strengths and virtues. A polynomial regression is simple, fast to compute, and easy to understand. In addition, this approach produces excellent results both when emulating Black-Scholes and when modeling real-market data. With real-market data, a model based on polynomial regression can apparently capture the true pricing relationships buried in noisy facts better than either a basic neural model or a hybrid one; there is less evidence of undesirable curve-fitting and of violations of reasonable monotonicity requirements. The major weakness of polynomial regression as a modeling technique is the problem of combinatorial explosion: the tendency for the number of free parameters (regression coefficients) to grow in proportion to the number of inputs taken to a power equal to the order of the polynomial required for a good fit to the data.

Neural networks also have their virtues. Like polynomial models, they can accurately mimic Black-Scholes and capture pricing relationships in empirical data. They appear better than polynomial models at coping with large numbers of inputs without requiring an astronomical number of free parameters to obtain a reasonably tight fit to the data. On the negative side, neural networks are slow to train (even on fast computers) and, more importantly, they are prone to localized over-fitting. The neural network developed in Study 4 demonstrated localized over-fitting when it modeled a sampling artifact in the data—the hook in the upper-right corner of Figure 7–4—almost as well as it modeled the sought-after pricing relationships. Finally, regardless of the number of inputs, the number of free parameters (connection weights) in a neural model that are necessary to get a good fit to the data is likely to be large.

Lastly, there are the hybrid models. The great strength of a well-designed hybrid model is that a good fit to the data can be achieved with a relatively small number of free parameters; this means the consumption of fewer degrees of freedom and the reduction in the likelihood of unwanted curve-fitting. Well-designed hybrid models that incorporate valid knowledge regarding a domain are also likely to be more robust than are more general nonlinear models. As with basic neural networks, training a hybrid model requires a massive amount of computation and is a time-consuming process. The hybrid model developed in Study 5 exemplifies some of the strengths and weaknesses of such models. Compared to the 294 degrees of freedom consumed by the basic neural model of Study 4 or the 160 degrees of freedom consumed by the polynomial regression in Study 3, only 48 degrees of freedom were required for the hybrid model to achieve roughly the same level of performance. Over-fitting was substantially less of a problem with the hybrid model than with the basic neural model. Surprisingly, localized over-fitting appeared to be more of a problem with the hybrid model than with the polynomial regression, despite the latter having twice the number of free parameters. Like the basic neural model, the hybrid model was slow to train. In addition, it required customized mathematical programming.

Overall, of the three models developed in this chapter, the current preference is for the polynomial regression. The polynomial regression model provided a reasonably good fit to the data, did not appear to curve-fit local features (such as the hook seen in Figures 7–3 through 7–5), and satisfied various monotonicity constraints.

The neural approach, basic or hybridized, needs further exploration, with an eye towards controlling excessive curve-fitting and achieving the kinds of monotonicity that probably characterize the true pricing relationships being sought, without seriously compromising the good fit to the data or the graceful handling of many inputs that are the strengths of the method. To this end, studies need to be done of regularization and biased estimation with neural-type models; except for the implicit bias in the hybrid model, no such studies appear in this book because of limitations of time and space. Additional avenues of research include polynomial hybrids, and training or fitting algorithms that allow specification of such things as monotonicity constraints. More efficient training algorithms would also be highly desirable and would make research in this area easier and faster to accomplish.

Although still in a development and research stage in the context of option pricing applications, nonlinear modeling technologies have been shown to hold great promise as tools for developing advanced pricing models and are worth exploring in greater depth.

## SUGGESTED READING

Two classic texts on the subject of linear regression are *Introduction to Linear Regression Analysis* (Montgomery and Peck, 1982) and *Classical and Modern Regression with Applications* (Myers, 1986). These well-written texts cover the basic theory of linear regression (in both univariate and multivariate forms) and discuss a variety of issues involved in its use, including the problem and diagnosis of collinearity. Polynomial and nonlinear regression models are covered in *Nonlinear Regression Analysis and Its Applications* (Bates and Watts, 1988) and in *Estimation with Applications to Tracking and Navigation* (Bar-Shalom et al., 2001). A discussion of neural networks and some of the important issues (such as curve-fitting and degrees of freedom) that must be addressed before they can be used effectively appears in *Virtual Trading* (Lederman and Klein eds., 1995). A general overview of neural network theory can be found in *Advanced Methods in Neural Computing* (Wasserman, 1993). The use of neural networks for noise reduction and pattern detection in time series data—remember, volatility and stock prices are time series—are discussed in *Neural, Novel and Hybrid Algorithms for Time Series Prediction* (Masters, 1995) and in *Neural Networks for RF and Microwave Design* (Zhang and Gupta, 2000). Finally, *The Nature of Mathematical Modeling* (Gershanfeld, 1998) covers problems of regularization and of nonlinear model-building generally.

CHAPTER 8

# Volatility Revisited

In the context of option pricing, volatility is unquestionably the variable most deserving of in-depth study. First, it is a variable that has a major influence on the worth of options. Second, the critical volatility is that which occurs in the future, not that which has occurred in the past. Third, volatility is thus problematic and must be estimated or predicted; it cannot be easily evaluated or merely specified, as can the other well-known variables that are important determinants of an option's value.

This chapter continues the exploration of volatility in the spirit of the earlier one (Chapter 5) on that subject. It presents studies designed to answer questions raised in the course of preparing Chapters 6 and 7, as well as to break entirely new ground. Among the subjects explored in this chapter are the influence of historical kurtosis and skew on future volatility, the usefulness of technical indicators (such as moving averages and oscillators) in the prediction of volatility, as well as the relationship of predicted volatility to implied volatility and, therefore, to actual option prices. The results of these studies should be useful to anyone who needs to price options, regardless of the particular pricing model employed.



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

## DATA AND SOFTWARE

The data used in the studies that follow came from the Worden Brothers TC-2000 database (www.worden.com) and from zipfiles downloaded from www.stricknet.com. The data were cleaned and saved in two binary databases in a compressed format designed for speed and convenience. The first of the two binary databases contained the stock data; the second contained the options data. A third binary database was also constructed; it contained implied volatility and open interest figures for both puts and calls calculated from the options data in the second binary database. Full details concerning the extraction, cleanup, and formatting of the raw data, as well as the calculation of implied volatilities, can be found in Chapters 4 and 5.

The software used to prepare the databases and to analyze the data was custom written in ISO-standard C, compatible with the GNU, Symantec, and other standards-compliant C/C++ compilers. Visualization and preparation of the results for presentation were carried out in a Microsoft Excel spreadsheet.

## STUDY 1: VOLATILITY AND HISTORICAL KURTOSIS

When studying conditional distributions, fair value was found to vary significantly with historical kurtosis. More specifically, positive kurtosis was associated with increased value for all options (puts and calls, all levels of moneyness) at lower levels of volatility, and with decreased value for options on stocks having higher levels of volatility. Was the effect of historical kurtosis on fair premium mediated exclusively by differences in the shape of the distribution of future stock returns, or was the effect at least partly due to an influence exerted by historical kurtosis on future volatility? Is historical kurtosis a useful input to a model designed to estimate future volatility, the kind of volatility crucial for pricing options? The current study attempts to answer such questions.

### Method

The data were analyzed both by bin statistics and by a second-order bivariate polynomial regression.

First, a stock was chosen from the 2,246 stocks available. The stock's data were retrieved from the database. Data series for 30-bar standard historical volatility, 100-bar historical kurtosis (based on 100 successive 1-bar returns), and 10-bar standard future volatility, were calculated using vectorized procedures. Then, a valid reference bar was selected from the 1,834 bars available for every stock. A reference bar was considered valid if (1) the stock was active (alive) over the preceding 100 bars and following 10 bars, and (2) the stock had no original (not split-corrected) price less than $2 over the past 30 bars.

Two array indices were then calculated. The first index (the *row index*) was calculated from volatility as

```
ix1 = (int)floor(0.5 + (nlvx-1)*(x1-bvxmn)/(bvxmx-bvxmn))
```

where `nlvx` was the number of levels or bins into which volatility was to be categorized for purposes of analysis, `x1` was the historical volatility at the reference bar, `bvxmn` was the center of the lowest volatility category, and `bvxmx` was the center of the highest volatility category. The second index (the *column index*) was calculated as

```
ix2 = (int)floor(0.5 + (nlad-1)*(x2-badmn)/(badmx-badmn))
```

where `nlad` was the number of levels into which kurtosis was to be categorized, `x2` was the historical kurtosis at the reference bar, `badmn` was the center of the lowest kurtosis category or level, and `badmx` was the center of the highest bin or level.

In calculating the first index (`ix1`), `nlvx` was 37, `bvxmn` was 0.20, and `bvxmx` was 2.00. The volatility levels corresponding to the centers of the bins, therefore, ranged from 20% to 200%, in increments of 5%. In the formula for the second index (`ix2`), `nlad` was 6, `badmn` was –1.5, and `badmx` was 9.0, which implied kurtosis categories centered at –1.5, 0.6, 2.7, 4.8, 6.9, and 9.0. Although not symmetric with respect to zero, the kurtosis categories encompass the vast majority of observed kurtosis figures; the asymmetry simply reflects the fact that kurtosis in stock returns is, on average, strongly positive.

The two indices were used to address elements in three arrays. In the first array, the value of the element referenced was incremented by 1; this array accumulated the bin or event counts. The future volatility at the reference bar was added to the value of the referenced element in the second array. In the third array, the squared future volatility was added to the element addressed by the index pair. The latter two arrays accumulated the data required to compute the bin statistics (the means and standard deviations of future volatility for each combination of historical volatility and historical kurtosis).

Next, data for the polynomial regression were passed to a regression object to be accumulated so that a regression could later be computed. There were six items passed to the regression object: (1) the intercept variable, which was just the constant 1; (2) the historical volatility at the reference bar; (3) the historical kurtosis; (4) the cross-product of volatility with kurtosis; (5) the square of volatility; and (6) the square of kurtosis. With a second-order polynomial regression calculated using double precision arithmetic, there was no need to use Chebyshev Polynomials, or even to center and normalize the data variables.

The accumulation of data required for computing the bin statistics and the polynomial regression only took place when historical volatility and historical kurtosis were within specified ranges; specifically, the ranges set by the bins or categories defined earlier.

At this point, the next valid reference bar was selected. When all valid reference bars for the chosen stock had been analyzed, another stock was chosen, its data retrieved from the database, and the sequence repeated. This continued until all stocks and reference bars had been processed.

The bin statistics and the regression results were then determined. Bin statistics were computed from the sums accumulated in the three arrays. The regression weights and related statistics were computed from data accumulated inside the regression object. Bin statistics consisted of the bin count, the mean of future volatility, and the standard deviation of future volatility. These three statistics were available for each possible combination of discretized historical volatility and kurtosis. In instances where the bin count was less than eight, all three statistics were given a value of zero. Regression statistics included the regression weights, the variance inflation factors, the *t*-statistics for the regressors, the multiple correlation (both raw and corrected for shrinkage), and more. Both the bin statistics and regression results were written to a standard text file and then loaded into an Excel spreadsheet for further analysis, charting, and presentation.

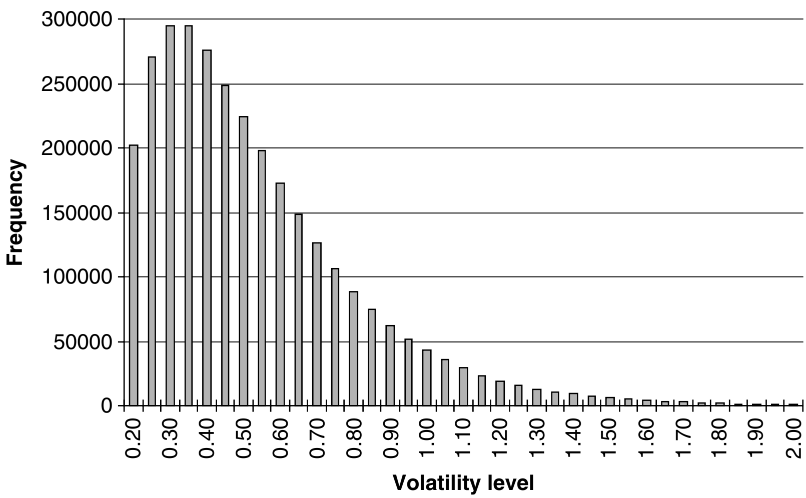
## Results

There was a total of 2,932,525 cases or data points in the sample. A histogram of the marginal frequency distribution of standard 30-bar historical volatility is presented in Figure 8–1.

In many places throughout this work, reference was made to the greater estimation error or noise caused by smaller samples (fewer data points) at higher levels of volatility. Charts like the one in Figure 8–1 were the basis for these references. The greatest number of data points fall in the bin centered at a volatility of 35%; as volatility levels rise above or fall below 35% (the mode), the number of data points in any bin (from which future volatility or any other variable may be estimated) steadily declines. Based on the data used in the current study, standard

**FIGURE 8–1**

Frequency Distribution for Volatility in the Data Sample



30-bar historical volatility has a mean of 52.6% and a standard deviation of 27.3%.

Some statisticians might argue that a variable with a skewed distribution like that of historical volatility should be transformed to improve its statistical properties. A log transform would, indeed, bring the observed distribution closer to normal. However, although this might improve certain statistical properties of historical volatility, it would also introduce serious distortion when the goal is to estimate the expectation of future volatility. Why? Option premium, $P(v)$, is roughly linear in its relationship to volatility, $v$, for at-the-money options, and approaches rough linearity for other options as volatility grows sufficiently large; in other words, for at-the-money options, and for other options as volatility grows large, fair value expressed as $E(P(v))$ is approximately equal to $P(E(v))$. However, $\exp(E(\ln(v)))$ does not equal $E(\exp(\ln(v)))$, which is nothing more than $E(v)$, and so $P(\exp(E(\ln(v))))$ does not necessarily approximate $P(E(v))$ or $E(P(v))$, even when high volatility levels or at-the-money options are involved.

The marginal frequencies for kurtosis were (in thousands) 59, 1,714, 734, 249, 111, and 63, for categories centered at –1.5, 0.6, 2.7, 4.8, 6.9, and 9.0, respectively. No separate frequency distributions were determined for future volatility, except those that appear in Table 8–1. Future volatility had a mean of 51.2% and a standard deviation of 34.1%. The standard deviation was larger for future volatility than for historical volatility because the future measure was based on a smaller number of bars (10) than the historical measure (30) and, therefore, contained more noise variance.

The regression yielded a multiple correlation of 0.598 between predicted and measured future volatility; the root-mean-square error was 0.273. In this context, the root-mean-square error is a better measure since some of the predictor variables explored in this chapter selectively restrict the sample, thereby altering the sample standard deviation and, consequently, the sample correlation; the root-mean-square error, however, continues to measure what is of interest: how well, in some absolute and comparable sense, the given predictor variable estimates future volatility.

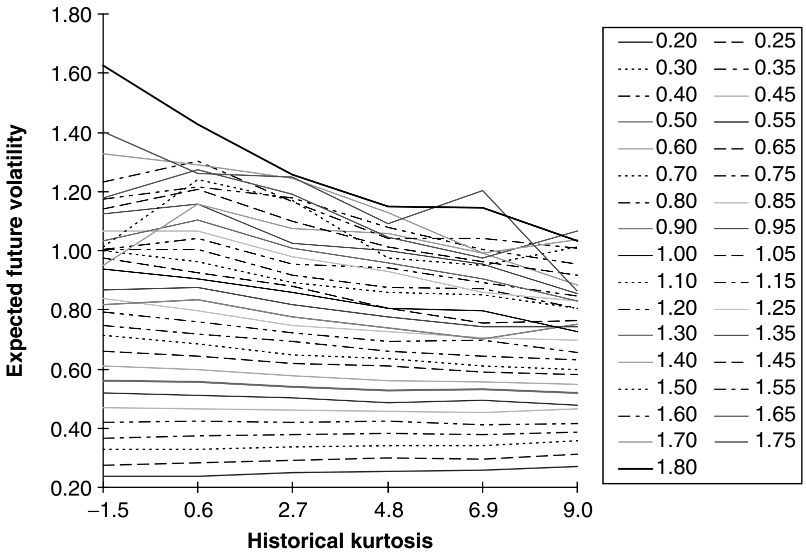
The regression weights (and associated *t*-statistics) were 0.02946 (43.3) for the intercept, 1.05765 (522.3) for historical volatility, –0.10832 (–122.5) for its square, 0.00910 (37.6) for historical kurtosis, 0.00050 (17.8) for its square, and –0.03348 (–123.9) for the cross-product or interaction term.

Although all terms in the regression were astronomically significant when considered in a statistical sense (thanks to a sample consisting of almost three million data points), in a practical sense the most important terms in the regression were the two powers (first and second) of historical volatility and the interaction term. The first two terms defined a relationship in which future volatility, the dependent variable, increased at a deaccelerating rate with historical volatility—the so-called "capping effect" described in Chapter 5. The interaction term defined a relationship in which kurtosis had a lowering effect on volatility in the future to the extent that historical volatility was high.

Figure 8–2 shows the mean observed future volatility associated with each combination of historical volatility and kurtosis.

**FIGURE 8–2**

Future Volatility as a Function of Historical Kurtosis for 33 Levels of Historical Volatility



Each curve in Figure 8–2 represents the relationship between kurtosis and expected (mean) future volatility for a particular level of historical volatility. Curves that correspond to low levels of historical volatility appear at or near the bottom of the figure, while those that correspond to high levels appear near the top. A legend has been placed alongside the chart as an aid to the reader in identifying the volatility level to which each curve corresponds.

As is fairly evident in Figure 8–2, up to an historical volatility level of 35%, future volatility tended to increase with increasing kurtosis. With few exceptions, future volatility decreased with increasing kurtosis at historical volatility levels above 45%. The exceptions resulted from estimation error, especially evident at higher levels of volatility. Estimation errors, i.e., noise, caused the curves near the top of the figure to be more irregular than those near the bottom.

## Discussion

The results suggest that historical kurtosis does have an effect on future volatility and, therefore, that it can be useful as an input to an option pricing or volatility model. Some of the effect kurtosis has on option value, as assessed by the conditional distribution methodology in Chapter 6, may be a result of its influence on future volatility and not merely on the shape of the distribution of returns. Specifically, for options trading on stocks having high levels of historical volatility, the reduction in the fair premiums observed with high levels of kurtosis may be explained in terms of the impact of kurtosis on future volatility. The greater premiums observed for both out-of-the money and in-the-money options at lower levels of historical volatility, however, cannot be fully explained in such terms; they must be explained on the basis of the shape (independent of volatility) of the future distribution of returns.

How significant, in a practical sense, is the influence of historical kurtosis on future volatility? Take a stock with an historical volatility of 100%. Given an historical kurtosis of –1.5, the expected future volatility of the stock is near 94%; it is near 73% when historical kurtosis is 9.0. Is the difference of practical importance? Most options traders would think so. Moreover, the effect becomes stronger at higher levels of historical volatility.

With historical volatility held constant, why should higher historical kurtosis be followed by lower future volatility? One possible explanation is that, for any given stock, periods of exceptional volatility (such as around earnings surprises) are marked by higher levels of kurtosis than are periods of more normal volatility. High kurtosis would then be an indicator of unusual volatility of a kind likely to return to normal in the near future (for instance, once the earnings data have been widely disseminated); it would act as a signal that mean reversion would be exaggerated in the near future.

## STUDY 2: VOLATILITY AND HISTORICAL SKEW

Like kurtosis, skew was found to affect fair premium (see Chapter 6). At high levels of volatility, positive skew was associated with higher fair premiums for out-of-the-money calls than was negative skew. This study attempts to determine whether historical skew affects future volatility and, if it does, whether its effect on volatility could be responsible for the observed influence on fair premium.

### Method

The analysis involved the same basic steps as in the previous study. It differed primarily in the variable used as the second predictor: historical kurtosis in the previous investigation and historical skew in the current one. The change in the second predictor variable required a change in the scaling parameters (`badmn` and `badmx`) used when calculating the second array index (`ix2`). In the current study, `badmn` was set to –3.5 and `badmx` was set to 3.5; the values were chosen after examining the frequency histogram associated with historical skew. The bin centers for skew were, therefore, at –3.5, –2.1, –0.7, 0.7, 2.1, and 3.5; the marginal frequencies corresponding to the bins centered at those values were 39,651, 107,131, 1,053,800, 1,776,124, 94,129, and 10,507, respectively.

## Results

A total of 3,081,342 cases were analyzed. This deviates slightly from the 2,932,525 cases analyzed previously as a result of differences in the number of cases in which the second predictor fell outside the specified range, i.e., the range covered by the bins.

The regression model yielded a multiple correlation of 0.585 between estimated and actual future volatility; the root-mean-square error was 0.278. The regression weights and the $t$-statistics (in parentheses) were 0.04551 (72.0) for the intercept, 1.03994 (534.4) for historical volatility, –0.22565 (–176.9) for its second power, –0.01104 (–28.0) for historical skew, –0.006 (–62.4) for its second power, and 0.02818 (51.0) for the cross-product of historical volatility and historical skew (the interaction term).

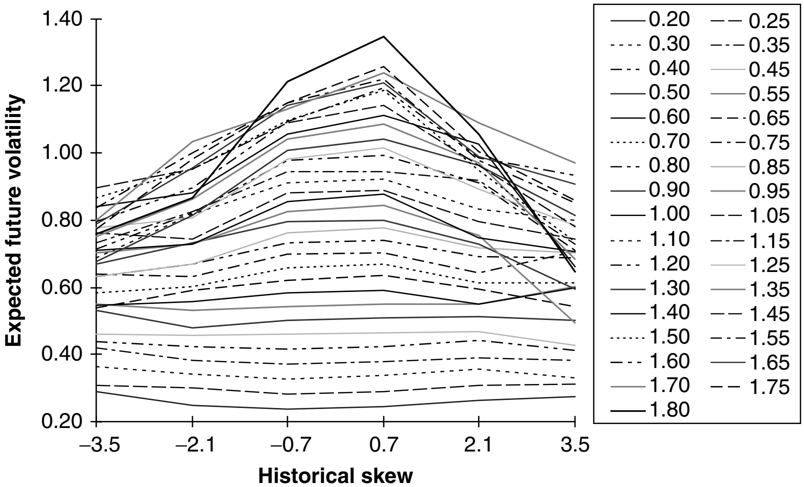
A visual comparison was made between the actual levels of future volatility and the estimates generated by the regression when plotted as a function of historical volatility and historical skew. The comparison indicated that a second-order polynomial regression was inadequate to provide a good fit to the data.

The data derived from the bin statistics, which appear in Figure 8–3, were unaffected by the regression and, therefore, correctly reflect the true relationships between the three variables. In Figure 8–3, historical skew is represented by the $x$-axis and future volatility by the $y$-axis. Each curve in the figure corresponds to a given level of historical volatility, ranging from 0.20 (20%) to 1.80 (180%), in increments of 0.05 (5%). The curves appear mostly in order of historical volatility (when observed near the midpoint of historical skew), from 20% at the bottom to 180% near the top. A legend is provided in the figure to aid in the identification of curves that fall close together.

For historical volatility levels below 40%, the curves trace out a smile-like shape: future volatility was at a minimum when historical skew was just below zero and rose as historical skew moved away from this point in either direction. Historical skew had little effect on future volatility when historical volatility levels were between 45% and 60%. Above 60% historical volatility, the curves tend to trace out an inverted smile: future volatility reached its peak when historical skew was slightly positive, and rapidly declined as skew deviated in either

**FIGURE 8–3**

Future Volatility as a Function of Historical Skew for 33 Levels of Historical Volatility



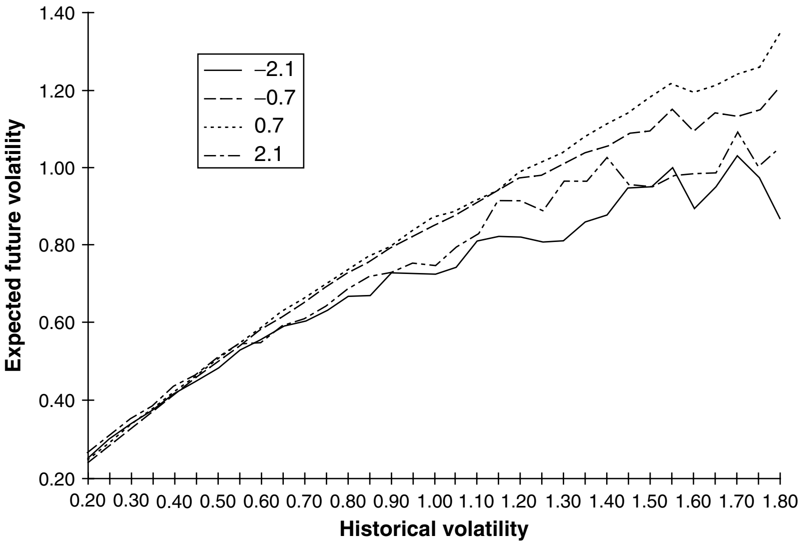

direction. The sharpness of the peak and the extent of the decline became more pronounced as higher levels of historical volatility were reached. As usual, the curves are less smooth, revealing greater noise and more estimation error, at higher levels of historical volatility.

Figure 8–4 shows the influence of historical measures of skew and volatility on future volatility from another perspective. In this figure, historical volatility appears on the *x*-axis, while different levels of historical skew are represented by distinct curves. Only four skew categories are shown; more than that would lead to a lack of clarity in the figure. A legend is provided for identification of the curves.

The uppermost curve (dotted line) is for a historical skew of 0.70, and depicts the strongest (and most linear) positive relationship between future and historical volatility. The weakest and most nonlinear—or "capped"—relationship between future and historical volatility was observed with a skew of –2.10; it is described by the lowermost curve (solid line). As well as in the relationships they describe, the curves also vary in noise. The variation in noise level resulted from a variation in the number

**FIGURE 8–4**

Future Volatility as a Function of Historical Volatility for Four Levels of Historical Skew




of cases in the different bins. When skew was 0.70, there were many more cases at all levels of volatility than there were when skew was –2.10; this resulted in smoother curves and less noise in the former instance and more ragged, noisier curves in the latter.

## Discussion

Does historical skew have a meaningful influence on future volatility? If it does, is the influence strong enough to be of practical relevance to the trader or hedger? Should historical skew be included in a volatility estimation model? The answer to all three questions is "yes."

Historical skew has an easily measurable influence on future volatility. The influence is not merely of academic interest, but is sizable enough to have practical, bottom-line relevance. Consider a stock with an historical volatility of 100%. With an historical skew that falls in the bin centered at 0.70, the expected future volatility is 87.8%. Compare this to expected future volatilities of 71.1% or 70.5% for historical skews that fall in the bins centered at –3.5 or 3.5, respectively. Moreover, the effect of historical skew on future volatility becomes even more pronounced at higher historical volatility levels. There is no question that historical skew should be included in a volatility forecasting model.

Another question asked at the beginning of this study was whether the influence of historical skew on future volatility could account for the differences in premium found in Chapter 6 using the conditional distribution methodology. That method of estimating fair premium automatically corrects for volatility errors and distributional shape and thus provides accurate estimates of fair value; however, it does not reveal how that fair value is causally determined—i.e., whether shape is the only relevant element or whether volatility is a mediating variable.

The results obtained in this study suggest that the effect of historical skew on fair premium found in Chapter 6 cannot be explained in terms of the influence of historical skew on future volatility. The influence of skew on future volatility appears to be roughly symmetric, with both positive and negative skew being associated with reduced volatility in the near future. If volatility was the mediating variable (rather than the shape of the distribution of returns), then option premiums should be about the same for both positive and negative historical skew and reduced when compared to zero-skew premiums. However, positive skew is associated with larger premiums (especially for calls) at high levels of volatility.

## STUDY 3: STOCHASTIC OSCILLATOR AND VOLATILITY

In this study, *Lane's Stochastic* is examined. The stochastic oscillator is a so-called "overbought/oversold indicator" that is popular with directional traders and technical analysts. It basically locates where, in the range of recent prices, the current price resides. One way the stochastic oscillator is used is to consider a crossing from oversold to neutral territory as a signal to enter a long position in the stock or future, and a crossing from overbought to neutral territory as a signal to enter a short position. Traditionally, oversold conditions are indicated by oscillator values below 20%, while overbought conditions are marked by oscillator values that exceed 80%. Another way to use the stochastic oscillator is to look for divergence between the oscillator and prices. When the oscillator makes a higher low despite prices tracing out a lower low, a reversal of price trend is anticipated and a long position might be established. Likewise, when the oscillator forms a lower high as the stock or future makes a higher high, a short position might be established to profit from the anticipated reversal of trend. A detailed investigation of Lane's Stochastic appears in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000); the results of that study suggest that, if used in one of the aforementioned ways, the stochastic oscillator will not provide a happy trading experience.

The fact that an indicator is of little value for directional trading does not imply an equal lack of value when trading or forecasting volatility; indeed, Study 6 of Chapter 6 hints at the possible value of the stochastic oscillator in the latter application. It was found that a crossover of the stochastic oscillator from below to above 20% was associated with higher fair premiums, especially for put options; a crossover from above to below 80% was associated with lower fair premiums. Perhaps the differences in fair premiums associated with the two crossover patterns is mediated by volatility—if not fully, at least partially—implying an association between that variable and oscillator behavior. It is the relationship between the behavior of the stochastic oscillator and future volatility that is investigated in the current study.

## Method

The present study, unlike the previous two, encompasses two separate, but related, analyses. The first analysis began, as usual, with the choice of a stock from the 2,246 available in the database. Data series were then retrieved for the chosen stock. Historical volatility (standard, 30-bar), future volatility (standard, 10-bar), and the stochastic oscillator were computed for all 1,834 bars from these series using fast, vectorized routines. The stochastic oscillator used in the study was the *slow stochastic* (also referred to as *Slow D%*) with a period of 14 bars for the oscillator look-back and 9 bars for the moving average.

Next, a valid reference bar was selected. To be valid, a reference bar was required to satisfy three conditions: (1) there had to be a sufficient number of bars before the reference to have valid figures for historical volatility and for the oscillator, (2) there had to be a sufficient number of bars after the reference to have a valid figure for future volatility; and (3) the stock had to have an unadjusted (for splits) price greater than $2 on each of the 30 bars immediately preceding the reference.

A determination was then made as to the direction in which the oscillator was moving. If the oscillator was on its way up, as evidenced by its value on the reference bar exceeding its value on the previous bar, then two array indices were computed; otherwise, they were not.

The first array index (the *row index*) was calculated as

```
ix1 = (int)floor(0.5 + (nlvx-1)*(x1-bvxmn)/(bvxmx-bvxmn))
```

where `nlvx` was 37, `x1` was the historical volatility, `bvxmn` was 0.20, and `bvxmx` was 2.00. The row index (`ix1`) took on a value of 0 for an historical volatility that fell in a bin centered at 0.20 (20%), a value of 1 for a volatility that fell in a bin centered at 0.25, all the way up to a value of 36 for an historical volatility that fell in a bin centered at 2.00 (200%); this is exactly as in the previous two studies.

The second array index (the *column index*) was calculated as

```
ix2 = (int)floor(0.5 + (nlad-1)*(x2-badmn)/(badmx-badmn))
```

where `nlad` was 8.0, `x2` was the value of the slow stochastic at the reference bar, `badmn` was 0.20, and `badmx` was 0.80. This index (`ix2`) took on values of 0, 1, ...7 for slow stochastic figures that fell in bins centered at 0.200, 0.286, ... 0.800, respectively.

If both indices were computed, and both were within their specified ranges, bin statistics were accumulated in three arrays and regression statistics were accumulated in a regression object, exactly as in the previous two studies.

At this point, the next valid reference bar was selected and all steps described above were again performed. When no more reference bars were left for the given stock, another stock was chosen and the entire process repeated. This continued until all bars of all stocks had been analyzed.

The accumulated data were then transformed into usable results. From the accumulated bin data, the mean future volatility was determined for each indexed combination of historical volatility and slow stochastic reading. From the data accumulated in the regression object, regression statistics (including weights and *t*-statistics) were calculated. The results were written to ordinary text files, which were then loaded into an Excel spreadsheet. Spreadsheet programs like Excel are useful when there is a need to view data in graphical form.

The second analysis was carried out in exactly the same way except for one minor change: rather than computing the array indices and accumulating the statistics only when the stochastic oscillator was on its way up, the indices were computed and statistics accumulated only when the oscillator was going down.

## Results

There were 1,407,176 data points associated with a rising slow stochastic and 1,426,966 data points associated with a falling slow stochastic. The sum of these numbers was 2,834,142, approximately the number of cases analyzed in each of the previous studies.

The regression for the rising slow stochastic yielded a multiple correlation of 0.590 and a root-mean-square error of 0.271. The regression weights were 0.12522, 1.11611, –0.25686, –0.32440, 0.25238, and –0.13321. The *t*-statistics associated with these weights were 55.0, 312.7, –148.2, –43.5, 37.4, and –32.9. The independent variables to which these statistics corresponded were the intercept, historical volatility, its square, the slow stochastic value, its square, and the cross-product term.

For the falling slow stochastic, the polynomial regression produced a multiple correlation of 0.588 and a root-mean-square error of 0.282. The regression weights were 0.12695, 1.17613, –0.25993, –0.32854, 0.26164, and –0.19597. The corresponding *t*-statistics were 58.7, 324.0, –139.7, –44.2, 37.5, and –45.3. The variables to which these numbers corresponded were again the intercept, historical volatility, its second power, the slow stochastic, its second power, and the interaction term.

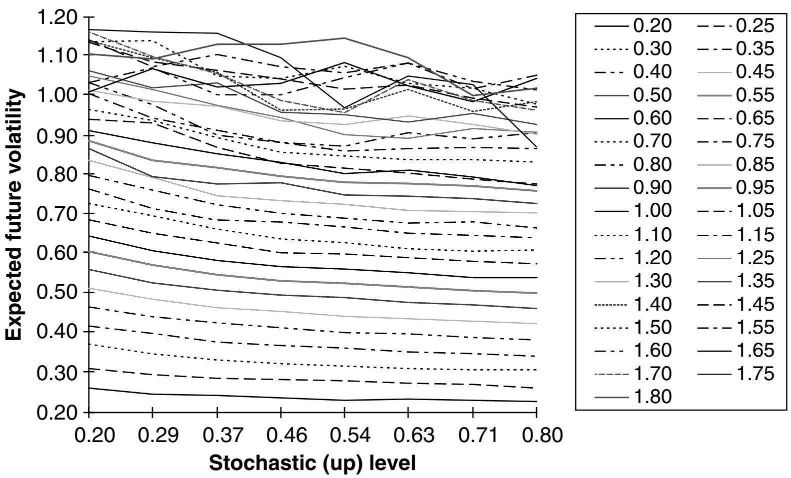
The regression weights and *t*-statistics were similar across both regressions. Judging by the *t*-statistics, historical volatility and its square are the most important variables in the model. The slow stochastic, its square, and the cross-product term, also contribute significantly to the model, albeit to a lesser degree.

Figure 8–5 shows expected future volatility (*y*-axis) plotted as a function of the level of the rising slow stochastic (*x*-axis) and the historical volatility (the 33 curves). The data shown in the figure were derived from the bin statistics.

As is clearly visible in Figure 8–5, the higher the level of the rising slow stochastic, the lower the expected future volatility.

**FIGURE 8–5**

Future Volatility as a Function of Stochastic (up) Threshold for 33 Levels of Historical Volatility



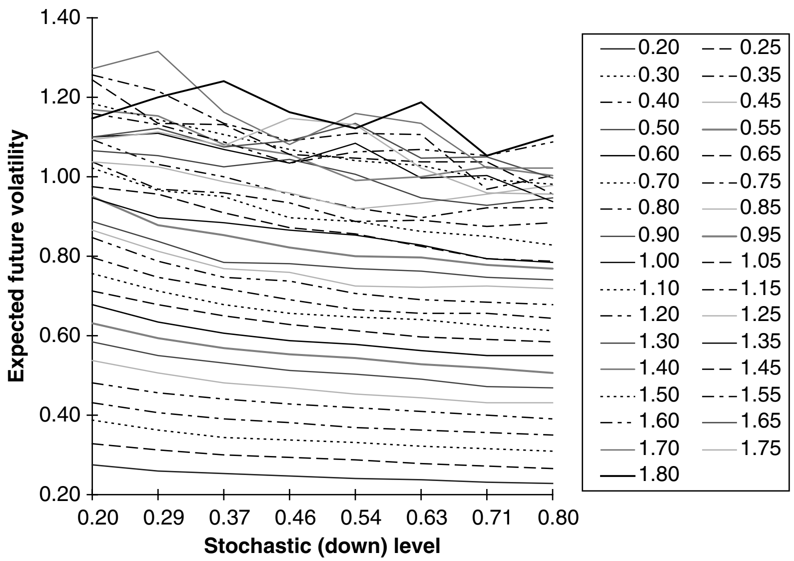
For any level of historical volatility, expected future volatility is highest at the leftmost edge of the chart, which corresponds to a stochastic of 20%. Expected future volatility initially descends at a rapid pace. The rate slows as the stochastic approaches 45%, but the decline continues all the way to the rightmost edge of Figure 8–5, at a stochastic level of 80%. The same pattern seems to exist at all levels of historical volatility, although the true underlying relationship is obscured by noise when volatility is at its highest levels, as represented by the curves that appear near the top of the chart.

Figure 8–6 is similar to Figure 8–5, except that it was based on the bin statistics for the falling slow stochastic, rather than for the rising one.

Essentially the same relationship appears to exist between expected future volatility and the level of the falling slow stochastic as was observed between expected future volatility and the level of the rising slow stochastic. Except where noise

**FIGURE 8–6**

Future Volatility as a Function of Stochastic (down) Threshold for 33 Levels of Historical Volatility



intrudes at the highest levels of historical volatility, expected future volatility is greatest when the falling slow stochastic is at 20%, declines rapidly as the slow stochastic approaches 45%, and continues to descend at a slower rate all the way to a falling slow stochastic of 80%.

The most important difference between the rising and falling stochastic is that the mean future volatility tends to be larger for the falling stochastic than for the rising stochastic at all levels of historical volatility. For example, at an historical volatility of 30%, a rising stochastic at the 0.20 level is associated with an expected future volatility of 37.4%, while a falling stochastic is associated with a mean future volatility of 38.7%; the corresponding figures for the rising and falling stochastics at the 0.80 level are 30.8% and 31.5%, respectively. When historical volatility is 100%, rising and falling stochastics at 0.20 are associated with future volatilities of 91.0% and 94.7%, respectively; rising and falling stochastics at the 0.80 level are associated with mean future volatilities of 77.0% and 78.5%.

One possible explanation for the higher expected future volatility observed with the falling stochastic has to do with price momentum or trend follow-through. Perhaps the downward spiraling stock prices that lie behind a falling stochastic have momentum, continue in their downward motion, and result in lower stochastic oscillator levels in the bars immediately following the reference. Lower stochastic levels are associated with higher levels of expected volatility, thus accounting for the higher future volatility levels seen with the falling stochastic. A similar argument works for explaining the association of lower expected future volatility with a rising stochastic.

## Discussion

When it comes to the estimation of future volatility, where the stochastic oscillator, a popular technical indicator, is coming from, and where it is, clearly has relevance. Imagine a stock with an historical volatility of 120%—a volatile tech stock, perhaps. Further imagine that the slow stochastic has just crossed the traditional 20% threshold, going from a reading of 17% on the previous bar to 21% on the current bar. Given this information, it could be concluded that the expected future volatility is about 104%. On the other hand, if the slow stochastic has just dropped below the traditional 80% threshold, landing at 71%, a future volatility of only 92% would be anticipated. Notice how both future volatility levels (104% and 92%) are well below the commonly used historical measurement (130%), and how the future volatility associated with the previously overbought stochastic (92%) is lower than that associated with the previously oversold stochastic (104%). Needless-to-say, technical analysts who trade options may find their indicators more useful than they previously thought.

What could be responsible for the higher future volatility levels observed with lower readings on the stochastic oscillator? Stochastic readings are usually high when a stock's prices are rising and near the top of their recent range; readings are low when prices are falling and near the bottom of their range. Trading lore suggests that falling prices and bottoms are usually marked by volatility, while steadily rising prices are often associated with quieter times in the market. If trading lore is correct, higher future volatility when stochastic readings are low, and vice versa, makes perfect sense. Naturally, this is only one—and not necessarily the best—of many possible explanations for the phenomenon.

Regardless of the explanation, Study 3 demonstrates that at least one technical indicator, Lane's Slow Stochastic, although of little use for directional trading, may have some real value when volatility, rather than direction, is at stake.

## STUDY 4: MOVING AVERAGE DEVIATION AND VOLATILITY

This study makes use of another tool that is popular with technical analysts: the *moving average*. Moving averages are used to smooth data and reveal trends. Many trading systems are built around moving averages. One such system, the *moving average crossover*, takes a long position in the market when prices are above the moving average and takes a short position when they are below. The idea is that, when prices are above the moving average, the trend is up; when they are below, it is down. Moving average systems such as this one were investigated in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000) and found to be ineffective for directional trading. Such systems are not predictive of the direction of future price movements.

But what about volatility? Is volatility, especially near-future volatility, affected by where prices are relative to the moving average? Study 4 investigates how the position of current price, vis-à-vis a standard moving average, affects expected future volatility and, consequently, fair option premium.

## Method

The analysis involved basically the same procedure as in the previous studies; the only differences were in the calculation and scaling of the second predictor or independent variable. In the current study, the second independent variable was calculated as

```
x2 = 100.0 * (cls[ibar]/ma[ibar] - 1.0)
```

where `cls[ibar]` was the closing (last) price at the reference bar and where `ma[ibar]` was the moving average at the reference bar. A standard 50-bar simple moving average (the kind most popular with stock traders) was employed in the calculation. The associated array index was computed as

```
ix2 = (int)floor(0.5 + (nlad-1)*(x2-badmn)/(badmx-badmn))
```

where `nlad` was 7 (the number of bins or categories), `x2` was as defined immediately above, `badmn` was –25 (the center of the first bin), and `badmx` was 25 (the center of the last category or bin). As before, the first independent variable, `x1`, was standard 30-bar historical volatility; the dependent variable, $y$, was standard 10-bar future volatility.

What is being investigated, therefore, is the difference between the closing price and the moving average (measured as a percentage of the moving average), and its impact on future volatility in the context of a given level of historical volatility.

## Results

The regression produced a multiple correlation of 0.576 and a root-mean-square error of 0.261. The root-mean-square error was smaller in this study than in any of the previous studies in this chapter. Large *t*-statistics were not only associated with terms involving historical volatility, but also with those involving the deviation of price from the moving average (hereafter referred to as *moving average deviation*). The *t*-statistics were 93.2, 469.3, –178.8, –138.9, 196.8, and 47.5 for the intercept, historical volatility, its second power, the moving average deviation, its second power, and the interaction term, respectively. Regression weights for these terms were 0.05849, 0.96750, –0.25129, –0.00477, 0.00019, and 0.00233, in the same order as the *t*-statistics. The regression weights for terms involving the moving average deviation may appear relatively small; this is not because these terms have little influence, but rather because they have a large standard deviation—0.259 for historical volatility versus 11.05 for the moving average deviation, or 122.13 for its square. The relatively small *t*-statistics for the interaction term suggests that the two variables, historical volatility and moving average deviation, exert a somewhat independent effect on future volatility.

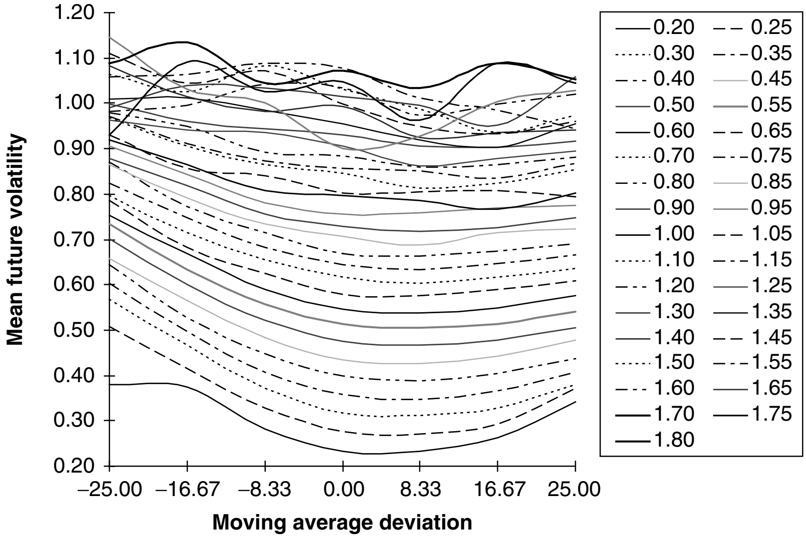
The statistical expectation of future volatility (*y*-axis) is shown as a function of percent deviation for the moving average (*x*-axis) and standard 30-bar historical volatility (set of curves) in Figure 8–7. At all levels of historical volatility less than 100%, expected future volatility is at a minimum when prices are just above the moving average. Expected future volatility increases rapidly as prices fall below the moving average; it rises more gradually as prices move further above the moving average. A similar pattern appears at higher levels of historical volatility, except that there is less of a gain in future volatility with rising prices and the data are much noisier.

## Discussion

When prices are lower than they have been in the recent past, volatility tends to be higher. Volatility is also higher, but not quite so much, when prices are higher than they have recently been. The lowest volatility occurs when a stock's price is at or just above the moving average, when there is a gentle upward drift in prices.

**FIGURE 8–7**

Future Volatility as a Function of Moving Average Deviation for 33 Levels of Historical Volatility




One possible explanation is that, for some period after a sharp drop in a stock's price, the absolute day-to-day variation in price remains as it was prior to the decline. Since volatility measures variation in price—not on an absolute basis, but relative to the prevailing price level—volatility appears higher to the extent that prices are lower.

Another possible reason for the greater volatility observed when prices deviate greatly from the moving average is the tendency of "overbought" and "oversold" conditions—unusually high or low prices—to bring out aggressive buying and selling. The reader surely recalls seeing the frantic and volatile activity at a "blow-off top" or "panic bottom." During such situations, stops get triggered, trend followers engage in panic exits (selling into dropping prices or buying into rising ones), and countertrend traders buy the dips or sell the peaks. The result is a price that swings rapidly up and down—i.e., volatility—as buyers and sellers step up to the plate.

What is odd is that one might have expected a volatility bump right at the moving average, caused by traders taking action when prices cross over or bounce against the moving average line. However, no such bump in volatility appears in the data.

Regardless of the explanation, the affect on near-future volatility of price, vis-á-vis the moving average, is quite large. For example, at 30% historical volatility, future volatility is 31% when the current price lies near the moving average, but reaches 57% (almost twice as great) when the current price of the stock is 25% below the moving average (a fairly common occurrence). An effect this large should be of great interest to traders and hedgers.

## STUDY 5: VOLATILITY AND MOVING AVERAGE SLOPE

The slope of a moving average can almost be considered a direct—albeit delayed—expression of trend. In the previous study, it was found that another noisier, but less delayed, measure of trend was strongly related to future volatility. What about the slope of the moving average? Is it also strongly related to future volatility? Is the relationship similar to one between future volatility and the deviation of prices from the moving average? These are the questions answered in Study 5.

### Method

The analytic procedure in the current study differed from that in the previous study only in the second independent variable and its scaling.

In the current study, the second predictor was calculated as

```
x2 = 100.0 * (ma[ibar]/ma[ibar-1] - 1.0)
```

where `ma[ibar]` was the 50-bar simple moving average at the reference bar and `ma[ibar-1]` was the same moving average at the bar preceding the reference. The variable `x2`, therefore, represents the slope of the 50-bar moving average of closing price, expressed as the percentage of change over one bar. The associated array index was determined as

```
ix2 = (int)floor(0.5 + (nlad−1)*(x2−badmn)/(badmx−badmn))
```

where `nlad` (the number of bins) was 7, `badmn` (the center value of the lowest bin) was –2.0, and `badmx` (the center of the highest bin) was 2.0.

As in Study 4, bin statistics and a simple polynomial regression were determined and written to a standard text file. The data in that file was loaded into an Excel spreadsheet where tables and charts could be prepared. The bin statistics, especially, lent themselves to graphical representation.

## Results

The two-dimensional bins that were identified with some combinations of historical volatility and moving average slope had either no data points falling in them or too few data points (less than eight) to calculate stable bin statistics. Because of this, some of the curves in Figure 8–8 appear to be cut short; they do not extend to the right and left sides of the chart.

**FIGURE 8–8**

Expected Future Volatility as a Function of Historical Volatility and Moving Average Slope

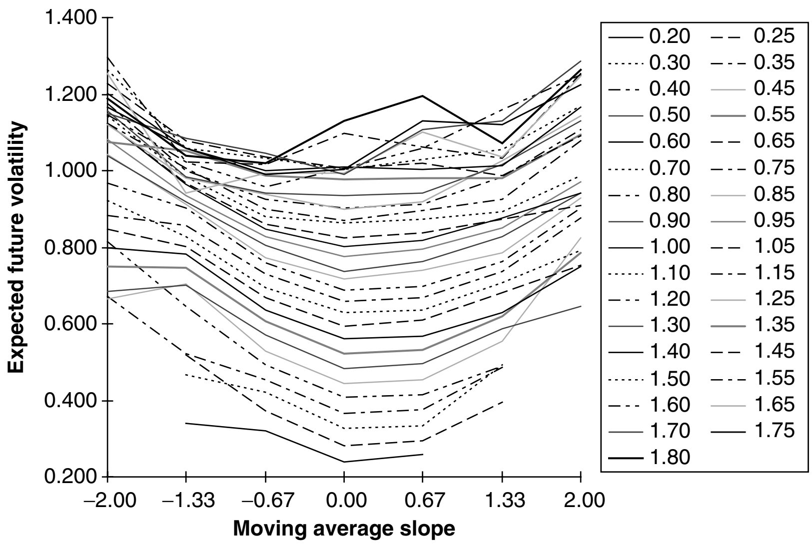


Figure 8–8 shows the mean future volatility ($y$-axis), as determined by the slope of the moving average ($x$-axis), for each of 33 levels of standard, 30-bar historical volatility (the 33 curves).

In Figure 8–8, the relationship depicted between moving average slope and historical volatility, on the one hand, and expected future volatility, on the other, is very similar in flavor to the relationship found between moving average deviation, historical volatility, and mean future volatility in Study 4. In both instances, near-zero values of the second predictor (moving average deviation or moving average slope) are followed by the lowest levels of future volatility. As the second independent variable (here, the moving average slope) moves away from the center of its range, an increase is observed in future volatility, the increase being greater for movement to the left (toward more negative or downward slope) than for movement to the right (toward more positive or upward slope).

Compared to the moving average deviation examined in Study 4, the rise in mean future volatility as one moves away from the center of the chart appears more extreme, especially at higher levels of historical volatility. The greater apparent effect on mean future volatility may be at least partly a consequence of the range of slope examined; had a somewhat smaller range—say, from –1.33% to 1.33%—been examined, the increases in future volatility with increasingly positive or negative slope might have been more modest and, therefore, more in line with those observed with the moving average deviation.

Another difference between the results for the moving average slope and the moving average deviation is that, with the former, the response of future volatility to increasing positive slope does not flatten out at higher levels of historical volatility as was the case with the latter.

As in all earlier studies in this chapter, higher levels of historical volatility are associated with more noise in the estimates of expected future volatility, noise which manifests as more jagged curves in the chart.

Finally, the data shown in Figure 8–8 suggest that the effects of the two predictors (historical volatility and moving average slope) on mean future volatility are fairly independent. This is indicated by the fact that the curves identified with different levels of historical volatility have roughly the same shape.

The regression analysis supported this observation: ignoring the intercept, the interaction term has the smallest $t$-statistic (73.80). The first and second powers of historical volatility have $t$-statistics of 532.6 and –223.0, respectively, while the first and second powers of moving average slope have $t$-statistics of –122.0 and 217.60, respectively. The regression weights were 0.04595 for the intercept, 1.03756 and –0.28663 for the first and second powers of historical volatility, –0.09374 and 0.07741 for the first and second powers of moving average slope, and 0.06712 for the interaction term. The weights for the two powers of moving average slope define a U-shaped parabola, consistent with the shape of the curves in Figure 8–8 with respect to the $x$-axis. The weights for the two powers of historical volatility describe an inverted parabola (with the maximum near the highest level of volatility) that describes the tendency of future volatility to level off as historical volatility increases. The regression yielded a multiple correlation of 0.587 and a root-mean-square error of 0.277.

## Discussion

Overall, the findings regarding the influence of moving average slope on future volatility are similar to those for the influence of price deviations from the moving average on the same variable.

Even when historical volatility is held constant, stocks that are in a strong downtrend or a strong uptrend (as measured by the slope of the moving average) have a higher mean future volatility than do stocks that exhibit little directional movement or trend.

How great is the influence of moving average slope (trend) on expected future volatility? At an historical volatility of 25%, future volatility is 28% when the moving average is going sideways (slope of zero); it is a much higher 52% when the moving average is going down (slope of –1.33%). At a 50% historical volatility, a “flat” moving average is associated with an expected future volatility of 48%, while a falling moving average (slope of –1.33%) is associated with a future volatility of 70%, and a rising average (slope of +2.00%) is associated with a future volatility of 65%. These differences in expected future volatility are large enough to imply substantial differences in fair premium for options trading on the relevant stocks.

## STUDY 6: RANGE PERCENT AND VOLATILITY

How is mean future volatility affected by the position of the current price in the range defined by prices over the past 50 bars? If historical volatility is held constant, is future volatility higher when current prices are down near support or up near resistance (with the potential for breakouts to new lows or highs) than when current prices are in the middle? When the estimation of future volatility is at issue, what is the interaction, if any, between historical volatility and the relative position of current prices with respect to their recent historical range? These are the questions that Study 6 attempts to answer.

### Method

The measure of the position of current price vis-á-vis the recent historical range was termed *range percentile*. It was defined as

$$\text{range percentile} = \frac{C - L}{H - L} \tag{8.1}$$

where $C$ represents the current closing price of the stock (i.e., its closing price at the reference bar), $L$ represents the lowest low price to occur over the past $m$ bars, and $H$ represents the highest high price to occur over the same $m$ bars. The variable $m$, of course, represents the period of the range percent measure or indicator; it was set to 50 in the current study.

The reader may notice the similarity of the range percentile measure to the stochastic oscillator. Indeed, range percentile (as defined here) is exactly that—the stochastic—but in a raw form, stripped of the usual smoothing, and computed over a much longer period than is typical for the standard oscillator.

In the current study, the second predictor variable, `x2`, was the range percentile, as defined above. The array index, `ix2`, was calculated from `x2` in the usual manner, with `nlad` set to

11, `badmn` set to 0.05, and `badmx` set to 0.95. All other variables, calculations, and aspects of the analytic procedure were unchanged from the previous study.

## Results

A multiple correlation of 0.592 and a root-mean-square error of 0.278 were obtained from the second-order polynomial regression. Ignoring the intercept, the largest $t$-statistics were associated with the first and second powers of historical volatility (487.4 and –225.7, respectively). Next in size were the $t$-statistics for range percentile and its second power (–159.5 and 149.9). The $t$-statistics for the intercept (132.8) and for the interaction term (–35.9) were the smallest. Regression weights for historical volatility and its square, range percentile and its square, the interaction or cross-product, and the intercept were 1.12038 and –0.28088, –0.42243 and 0.33673, –0.07068, and 0.13621, respectively.

Figure 8–9 depicts graphically the relationship between a stock's current range percentile reading ($x$-axis) and its expected future volatility ($y$-axis). The relationship is shown for each of the 33 levels of historical volatility (the curves in the figure): from 20% at the bottom to 180% at the top.

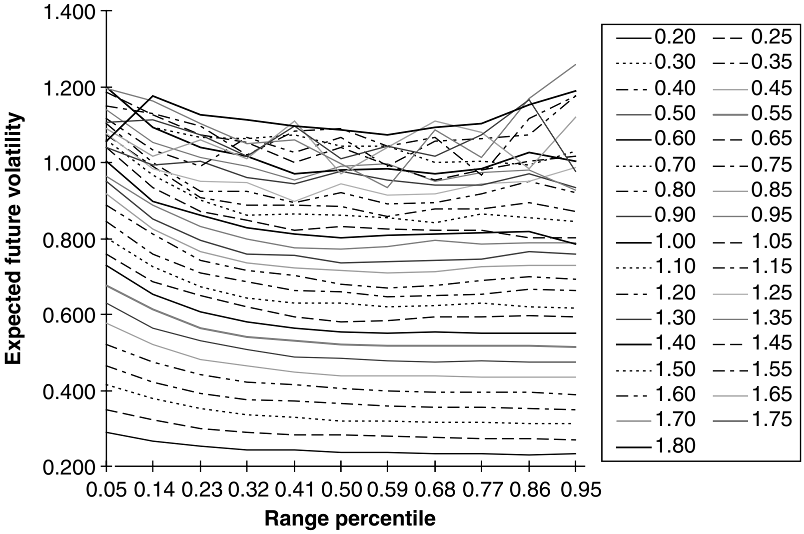
At all levels of historical volatility, as range percentile increases from 5 to 50, expected future volatility declines, at first rapidly, then more gradually. This is what appears in Figure 8–9 when attention is on the true relationships between the variables and when the noise in the data at higher levels of historical volatility is filtered out by the observer. Considered from a different perspective, future volatility increases as prices approach the lower limit of their 50-bar range.

Another observation is that, as prices approach support at the bottom of their 50-bar range, the gain in future volatility tends to be greater at higher levels of historical volatility than at lower levels. This suggests some interaction between historical volatility and range percentile.

For a range percentile reading between 50 and 100, the influence upon expected future volatility also changes with the level of historical volatility, revealing an interaction. Here, the change is

**FIGURE 8–9**

Mean Future Volatility as a Function of Range Percentile for 33 Levels of Historical Volatility




not merely one of amplitude or effect size, but also one of direction. At low levels of historical volatility (20% to 60%), future volatility, which fell rapidly as range percentile climbed from 5 to 50, continues to decline much more gradually as range percentile travels from 50% to 95%. When the level of historical volatility is moderate (65% to 155%), future volatility remains fairly constant as range percentile goes from 50% to 95%. At high levels of historical volatility, expected future volatility increases with range percentile over the 50% to 95% interval.

Finally, as was observed in all previous studies, the curves get choppier, revealing greater noise or measurement error at higher levels of historical volatility where sample sizes are smaller.

## Discussion

What does the effect of range percentile on future volatility look like in tangible numbers? Consider a case where the standard, 30-bar historical volatility is 30%. At a range percentile of 5, expected future volatility is 41%. Future volatility decreases to 32% at a range percentile of 50, and to 31% at a range percentile of 95. When historical volatility is 60%, mean future volatility is 73% for a range percentile of 5; it is 55% for a range percentile of 50, and it continues to be 55% for a range percentile of 95.

The pattern seen with range percentile is in some ways similar to that encountered with the moving average indicators. In all cases, higher future volatility is associated with low stock prices (those near the bottom of the range), even when historical volatility is held constant. Moreover, in some instances, higher future volatility is also associated with high stock prices (those near the top of the range).

What could account for the patterns observed in the data? Why was heightened future volatility found at all levels of historical volatility when stock prices were near support, but only at very high historical volatility levels when stock prices were near resistance?

It is easy to understand the phenomenon of heightened future volatility when historical volatility is high and stock prices are either low (near support) or high (near resistance) as opposed to middling. If stock prices are near support or near resistance, a high level of recent historical volatility implies that they arrived there in a sharp move. Sharply lower prices on high volatility (and volume) characterize a "panic bottom," which often marks the final sell-off before a reversal of trend and a rapid (hence volatile) recovery following the event. Likewise, sharply higher prices on high volatility (and volume) are typical of a "blow-off top," which is generally followed by a noteworthy correction.

But what about when prices merely drift lower or higher, eventually reaching support or resistance on low-to-moderate volatility? One possible explanation for the increased mean future volatility near support has to do with news and corporate events. A gradual decline in prices is often seen when a company fades from the spotlight. Perhaps earnings have lost their luster. Perhaps the company's industry group is no longer perceived as "hot." At some point, however, an earnings surprise brings the company back into the spotlight. Or, perhaps, prices drift sufficiently low to make the company a tempting target for a takeover bid. Or, then again, prices start to drift up from the bottom, triggering a short squeeze. Either way, the result is a strong surge in price and volatility.

The authors have seen this happen again and again, especially among ignored, low-cap stocks that are marked by sparse volume, low prices, and wide ask-bid spreads. Of course, some of these downtrodden stocks just continue their decline, which may even accelerate on bad news. Again, the result is high volatility.

On the other hand, observations of the market suggest that a calm, upward drift in prices can be sustained for a long time, without either a sudden correction or a buying frenzy and the concomitant volatility. Behind the gently rising prices often lies a company with modest, but stable, earnings growth. The relatively high prices make the company less tempting than a company with a depressed stock, and a strongly positive earnings surprise is harder to come by given the already good earnings growth. Naturally, a negative surprise is possible; however, it is more likely to occur when earnings are shooting to the moon, the company is a high-flyer, and volatility is already quite high.

Of course, these are just some possible explanations for why volatility should always be higher near support, but only sometimes higher near resistance; undoubtedly, the reader can think of many other equally compelling ones. Regardless of the reasons behind the observed effects, they are large enough to have a serious impact on the bottom line.

## STUDY 7: MONTH AND VOLATILITY

In Chapter 4, the effect of month-of-year on volatility, as well as on other statistical moments (mean, skew, and kurtosis), was examined. In that chapter, the month-of-year effect was investigated on its own; the analysis was univariate. The current study employs a bivariate analysis; month-of-year is examined together with historical volatility. By considering both month-of-year and historical volatility together, any interaction between the two variables can be observed—something that was not possible in the earlier analysis.

Study 7 was designed to address a number of questions. Do monthly differences in volatility persist when historical volatility is held constant? Do these differences evidence the same quarterly pattern that was seen when historical volatility was not held constant, as was the case in Chapter 4? And, most importantly, is there an interaction between month-of-year and historical volatility, i.e., does the relationship between month and future volatility change with the level of historical volatility?

### Method

The analysis was conducted in the same manner as in all earlier studies in this chapter. The only difference was in the determination and scaling of the second independent variable which, in the current study, was month-of-year. As previously, standard 30-bar historical volatility was the first predictor and standard 10-bar future volatility was the dependent variable or target.

For month-of-year, the scaling parameters `nlad` (the number of bins), `badmn` (the center value of the lowest bin), and `badmx` (the center value of the highest bin) were set to 12, 1, and 12, respectively. The scaling parameter `nlvx` remained set at 37, `bvxmn` at 0.20, and `bvxmx` at 2.00, for historical volatility.

### Results

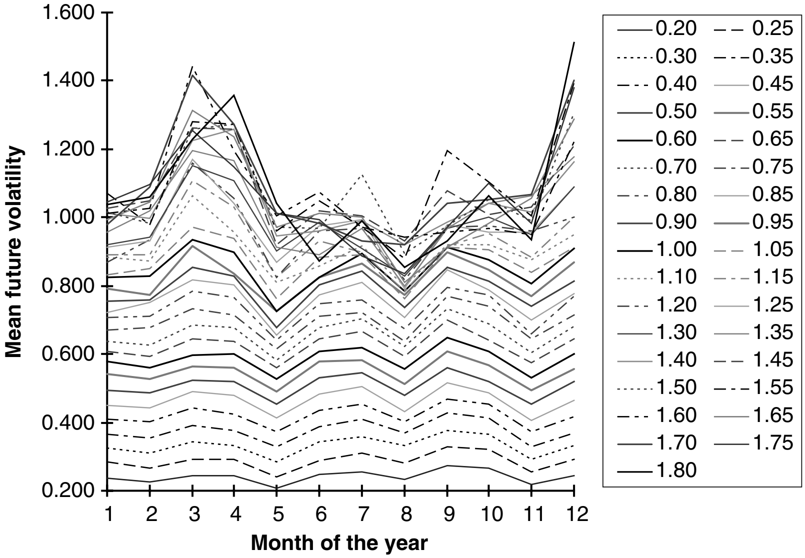
Figure 8–10 shows mean future volatility plotted against month-of-year for each of the 33 levels of historical volatility.

As clearly seen in Figure 8–10, expected future volatility varies dramatically with the month of the year, even when historical volatility is held constant. Lows in expected future volatility occur every three months—in February, May, August, and November—regardless of the level of historical volatility. Highs in mean future volatility occur either one or two months earlier, with the exact timing dependent on historical volatility and time of the year, suggesting an interaction.

An interaction between historical volatility and month-of-year is also evident in the behavior of mean future volatility in March and December. The extent to which future volatility in March and December exceeds that seen in other months is

**FIGURE 8–10**

Expected Future Volatility as a Function of Month of the Year for 33 Levels of Volatility




controlled by the level of historical volatility: the higher the historical volatility, the greater the relative future volatility in March and December. This interaction is not apparent for the months of July and September, months that are also characterized by relatively high levels of mean future volatility.

Overall, May is the quietest month. In May, future volatility hits its yearly low at most historical volatility levels. At low levels of historical volatility, the month with the highest mean future volatility is September. For stocks that exhibit high levels of historical volatility, December is the month that has the highest future volatility. Because high volatility stocks yield a large volatility payoff, these are probably the stocks that contribute most to the so-called “January effect.”

Although a second-order polynomial regression was computed, the results are not reported here. The curves in Figure 8–10 do not exhibit either linear or parabolic trends of a kind that a second-order polynomial can effectively model; rather, these data exhibit cycles that are best modeled with sines and cosines, as in Study 4 from Chapter 5. The second-order regression computed in the standard procedure could not model the obvious cyclic relationships in the data; this was clearly indicated by the negligible *t*-statistics for all terms in the regression that involved month-of-year.

## Discussion

The current study confirms the presence of a strong quarterly cycle in volatility, the kind that was discussed in Chapter 4 in the context of seasonal cycles. It also demonstrates (1) the persistence of this cycle when historical volatility is held constant and (2) the presence of a strong, easily observed interaction between a stock's historical volatility and the month of the year when these two variables are considered as predictors of future volatility.

What causes the quarterly cycle? One factor that contributes to the cycle is the quarterly release of earnings reports. Another contributing factor is the expiration of the S&P 500 Index futures, together with the options, on what are known as "triple witching days," which are notorious for their volatility.

Explaining the interaction between historical volatility and month-of-year is much more difficult than finding reasons for the quarterly cycle. Perhaps the phenomenon has something to do with actions taken in response to the annual tax cycle. Such actions have been hypothesized to be at the root of the "January Effect," and may play a role in the interaction between month and volatility. Exactly how this might work, however, is unclear, especially considering that the second peak in the relative future volatility occurs in March.

How much is future volatility influenced by month-of-year and by its interaction with historical volatility? With historical volatility at 30%, the mean future volatility in May is 28%; it is 36% (8% points higher) in July, and 33% (only 5 points higher) in December. The expected future volatility is 97% in May, 113% (16 points higher) in July, and 131% (a whopping 34% points higher) in December, when historical volatility is held constant at 150%. When considered on their own, these are all quite sizable differences in volatility. For example, an at-the-money option on a stock with a historical volatility of 30% purchased in May would be worth roughly 28% more than an identical option purchased in December. The interaction effects are also large, as is easily seen when these differences are considered in relation to one another. When historical volatility is low (30%), the difference in future volatility between May and October is 8 points; when historical volatility is high (150%), the difference is 34 points.

## STUDY 8: REAL OPTIONS AND VOLATILITY

Study 8 differs from the earlier studies in this chapter in that it examines the relationships of historical variables, not only to future volatility, but also to the volatility implied by the prices at which options actually trade.

Do option-implied volatilities track standard historical volatility? Or, do they track better estimates of future volatility, such as the one developed in Chapter 5, Study 4? And what about kurtosis? Does it drive up real option prices and, therefore, implied volatility, when these are considered relative to a good estimator of future volatility? These are some of the questions addressed below.

If implied volatility closely tracks standard historical volatility, then the decision-making process of most traders and hedgers (including the large institutional ones) probably utilizes standard historical volatility and pricing models that do not implicitly correct for biases in the volatility measures used as model inputs. On the other hand, if implied volatility tracks a better estimator of future volatility, then perhaps the bigger players in the options arena use good future volatility estimates and better pricing models than the standard ones; or, perhaps, they use deep market experience to make subjective corrections to the theoretical premiums produced by the standard models with raw historical volatility.

Traders and hedgers who act primarily based on the implied volatility tend only to cause options to stay in line with one another; they do not generally influence overall implied volatility or option price levels, either up or down, by their trading activities. This is because implied volatility can only be used to make relative appraisals of one option against another.

## Method

The procedure involved the (by now, familiar) accumulation of bin statistics, although it differs in the details from the procedure used for the earlier studies in this chapter.

The first step was to choose the volatility measure or estimate that was to serve as the independent variable in the analysis. There were four possible choices. The first choice was standard 30-bar historical volatility. Average range historical volatility, also computed over 30 bars, was the second choice. The third choice was again 30-bar standard historical volatility, but this time corrected for nonlinearity and regression to the mean. A good multivariate estimate of future volatility (see Study 4, Chapter 5) was the fourth choice.

Once the volatility measure or estimate was chosen, a stock was selected. Data were retrieved for the stock from the first of the two binary database files. This database was the same one employed in all earlier studies in this chapter; its construction is fully described in the section on raw data in Chapter 4. Implied volatility data for options trading on the selected stock were retrieved from the second of the two binary database files. The construction of the implied volatility database from the raw options data are discussed in the section on the calculation of implied volatility found in Chapter 5.

A valid reference bar was then designated for analysis. Validity of the reference bar was judged based on three criteria: first, the stock had to be active over the last 30 bars prior to the reference bar; second, the lowest price encountered over those 30 bars had to be greater than $2; and third, a valid figure for overall implied volatility had to exist at the reference bar.

The overall implied volatility was calculated as the average of the put-implied and call-implied volatilities that were obtained from the second binary database. The put-implied and call-implied volatilities stored in the database were weighted averages of the implied volatilities for several near-term puts and several near-term calls, respectively; they were not the implied volatilities of individual options. If the overall implied volatility was greater than 0.05 and less than 3.50, and the absolute difference between the put-implied volatility and the call-implied volatility was less than 0.20, then the overall implied volatility figure was considered to be valid. Assuming that a valid reference bar had been found, the overall implied volatility was taken as the dependent variable in the analysis.

At this point, the chosen volatility measure or estimate that was to serve as the independent variable in the analysis was computed. This volatility estimate was based only on the stock's past (pre-reference) behavior and possibly on the time of year; it did not involve either future stock behavior (as does future volatility), or real-market option premiums (as does implied volatility).

Given the independent and dependent variables as defined above, the bin statistics were accumulated. The volatility measurement or estimate that was to serve as the independent variable was used to compute an array index. In C language, the calculation was

```
ix = (int)floor(0.5 + (nlvx-1)*(x-bvxmn)/(bvxmx-bvxmn))
```

where `ix` was the desired array index, `nlvx` was the number of bins or levels, `x` was the value of the independent variable (the chosen volatility estimate), `bvxmn` was the center for the lowest bin, and `bvxmx` was the center for the highest bin. In the current study, `nlvx` was 37, `bvxmn` was 0.20, `bvxmx` was 2.00, and `ix` was in the range of 0 to 36. The index, `ix`, was used to address two arrays. The first array served to accumulate the number of cases falling in each bin and thus associated with each level of the independent variable; the accumulation was accomplished by incrementing the value of the array element addressed by the index, `ix`, by 1. The second array was used to accumulate the corresponding sums of the dependent variable; the sums were accumulated by adding the value of the dependent variable (implied volatility) to the value of the element addressed by `ix`.

Once these data were accumulated, the next valid reference bar was designated. When no more valid reference bars were available for the selected stock, another stock was selected.

The sequence was continued until all valid reference bars for all the stocks had been processed.

At this point, the mean of the dependent variable was determined for each level of the independent variable. This was accomplished by dividing the sum accumulated in each element of the second array by the case count in the corresponding element of the first array. Then, the array index (`ix`) was stepped from 0 to 37 in increments of 1; the bin center that corresponded to the index (the level of the independent variable) was written to a standard text file, followed by the mean that was computed from the pair of array elements addressed by the index.

Finally, the next of the four possible independent variables was chosen. The arrays used to accumulate data for the bin statistics were then cleared and the whole analysis was repeated. This went on until analyses had been carried out with each of the four measures or estimates of volatility.

The results of all these calculations were four sets of data. Each data set consisted of two columns. The first column in each set contained the levels of the measured historical, or predicted future, volatility corresponding to the bin centers. The second column in each data set contained the associated mean implied volatility figures. Each data set, therefore, revealed the relationship between one of the four kinds of measured historical or estimated future volatility, on the one hand, and option-derived implied volatility, on the other.

## Results

The data generated by the analyses appear in Table 8–1. In the table, only one column, the first (BINCTR), was used for the bin centers that defined the discretized values of the independent variable; this was possible because the bin centers were identical across all four sets of data. Each of the remaining four columns in Table 8–1 was simply the second column found in each of the four data sets generated by the analyses.

Each row of the second column (STDHVX) in Table 8–1 contains the mean implied volatility that was observed when the standard 30-bar historical volatility fell in the bin centered at the value found in the same row of the first column. The third column

**TABLE 8–1**

Option-Implied Volatility versus Several Measures of Historical Volatility and Predicted Future Volatility

| BINCTR | STDHVX | ARHVX | STDADJ | PREST |
|---|---|---|---|---|
| 0.200 | 0.288 | 0.273 | 0.244 | 0.238 |
| 0.250 | 0.337 | 0.314 | 0.294 | 0.271 |
| 0.300 | 0.386 | 0.359 | 0.346 | 0.319 |
| 0.350 | 0.436 | 0.408 | 0.401 | 0.372 |
| 0.400 | 0.475 | 0.453 | 0.452 | 0.430 |
| 0.450 | 0.519 | 0.496 | 0.501 | 0.487 |
| 0.500 | 0.561 | 0.538 | 0.553 | 0.549 |
| 0.550 | 0.605 | 0.585 | 0.608 | 0.604 |
| 0.600 | 0.647 | 0.629 | 0.660 | 0.662 |
| 0.650 | 0.683 | 0.671 | 0.713 | 0.722 |
| 0.700 | 0.722 | 0.715 | 0.768 | 0.776 |
| 0.750 | 0.762 | 0.757 | 0.824 | 0.830 |
| 0.800 | 0.793 | 0.796 | 0.864 | 0.875 |
| 0.850 | 0.835 | 0.837 | 0.899 | 0.922 |
| 0.900 | 0.854 | 0.873 | 0.946 | 0.961 |
| 0.950 | 0.878 | 0.907 | 0.992 | 0.985 |
| 1.000 | 0.902 | 0.935 | 1.001 | 1.045 |
| 1.050 | 0.910 | 0.967 | 1.034 | 1.078 |
| 1.100 | 0.943 | 0.988 | 1.127 | 1.124 |
| 1.150 | 0.983 | 1.016 | 1.116 | 1.187 |
| 1.200 | 0.989 | 1.059 | | 1.226 |
| 1.250 | 1.006 | 1.092 | | 1.250 |
| 1.300 | 0.982 | 1.111 | | 1.297 |
| 1.350 | 1.016 | 1.127 | | 1.326 |
| 1.400 | 1.053 | 1.169 | | 1.391 |
| 1.450 | 1.042 | 1.238 | | 1.398 |
| 1.500 | 0.994 | 1.260 | | 1.435 |

(ARHVX) contains the mean implied volatility for each level of 30-bar average range historical volatility. The fourth column (STDADJ) lists the expected implied volatility for each level of adjusted standard historical volatility. The mean implied volatility associated with each level (BINCTR) of regression-estimated future volatility appears in the fifth column (PREST). The table only contains data for levels of the independent variable up to 1.50 (150%).

When volatility is low, implied volatility is much higher than standard 30-bar historical volatility; when volatility is high, implied volatility is much lower than the standard historical measure. As an example, a 25% standard historical volatility is associated with a mean implied volatility of 34%; a 120% standard volatility is associated with a 100% implied volatility. The pattern suggests regression to the mean, a phenomenon that was observed in Chapter 5 when standard historical volatility was used, uncorrected, as an estimator of volatility in the near future. The fact that there is little gain in implied volatility as standard historical volatility rises beyond 125% (at 150% standard volatility, mean implied volatility is only 99%) also implies the presence of the capping effect observed in Chapter 5, when the standard historical measure was used as a proxy for future volatility.

Regression to the mean is less, and capping much less, when average range historical volatility is examined as a predictor of implied volatility. With average range historical volatility at 25%, a mean implied volatility of 31% is observed; with average range volatility at 120%, the mean implied volatility is 106%; and with average range at 150%, implied volatility has a mean of 126%. Average range historical volatility was also found to exhibit less capping and mean reversion when predicting future volatility.

Corrected for regression to the mean and for nonlinearities such as capping when used as a predictor of future volatility, adjusted standard historical volatility does much better as a predictor of the volatility implied by option prices. With adjusted standard volatility as the independent variable, mean reversion and capping are gone. All that remains is a tendency for expected implied volatility to hover about 5% above the adjusted standard measure over much of the volatility spectrum. Because this measure only took on values up to 115%, no comparisons can be made at higher levels of volatility, as they were for the other two measures.

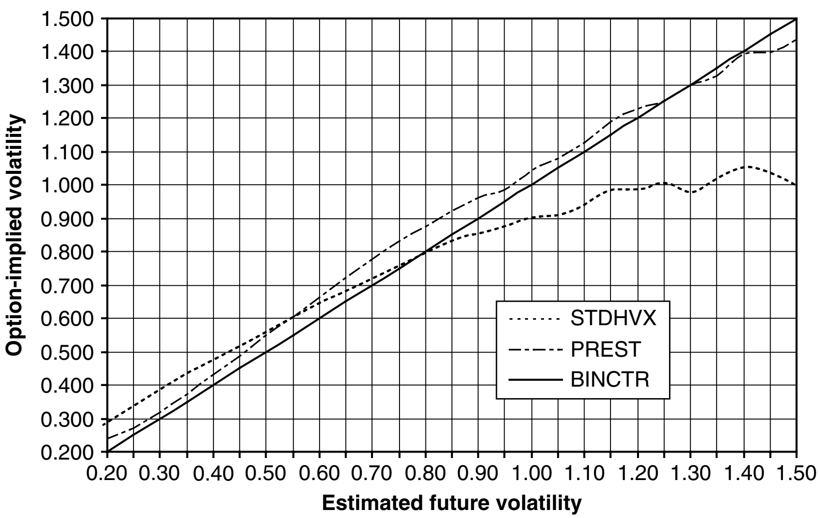
There is no question that the volatility measure that best estimated mean implied volatility was the multivariate regression estimate of future volatility (developed in Study 4 of Chapter 5) that was based on two measures of historical volatility together with several seasonal harmonics. There is little evidence of either mean reversion or capping with this estimate. And, as was also the case when it was used to predict future volatility, the multivariate estimator is capable of predicting fairly high levels of volatility; its range is not restricted as is the range of the adjusted standard volatility measure.

When the estimated future volatility (using the multivariate regression estimator) is at 25%, expected implied volatility is 27%; when estimated future volatility is at 120%, implied is 123%; and when estimated future volatility is at 150%, the mean implied volatility is 144%. There is still some tendency for implied volatility to be greater than the estimate of future volatility, although the disparity is much smaller with the good regression estimator of future volatility than with the cruder ones discussed earlier.

Figure 8–11 shows how implied volatility behaves in response to standard historical volatility and to the better regression-based estimator of future volatility. The dotted line

**FIGURE 8–11**

Implied Volatility Plotted against Two Very Different Estimates of Expected Future Volatility



(STDHVX) in the figure depicts the relationship of implied volatility (on the *y*-axis) to standard historical volatility (on the *x*-axis). The broken dash-dot line (PREST) shows how expected implied volatility (again, the *y*-axis) responds to the regression-based estimate of future volatility (on the *x*-axis). The straight, solid line (BINCTR) in Figure 8–11 represents the relationship that would exist between implied volatility (*y*-axis) and a perfect estimator thereof (*x*-axis); it was placed in the figure to serve as a baseline against which the other lines and the relationships they represent could be compared.

When estimated future volatility is low, the dotted line falls above the solid line; at levels less than about 80%, standard historical volatility (taken here as the estimate of future volatility) underestimates implied volatility. At about 80% estimated future volatility, the dotted line drops below the solid line and, as estimated future volatility continues to increase, the gap between the solid line and the dotted one widens; above 80% standard historical volatility, implied volatility is overestimated and the overestimation becomes progressively worse as standard historical volatility continues to rise.

The pattern of over- and underestimation just described suggests regression to the mean, as well as a blunted response of implied volatility to increases in standard historical volatility at higher levels—the capping effect. Regression to the mean and capping were also observed in Study 1 of Chapter 5, when standard historical volatility was examined as a predictor for future volatility. Overall, standard historical volatility is a poor estimator of future volatility, as well as of implied volatility, at low levels and, even more so, at high levels.

The broken dash-dot line in Figure 8–11 traces a path that lies much closer to the solid line than does the dotted line. This attests to the much closer fit between implied volatility and the regression estimate of future volatility than between implied volatility and the future volatility estimate based on raw standard historical volatility. In fact, the fit between implied volatility and the polynomial regression estimate of future volatility is quite good, except for the slight bulge above the baseline centered at an estimated future volatility (*x*-axis) of around 70%.

The bulge reflects the tendency of implied volatility to be underestimated by (i.e., greater than) the future volatility predicted by the regression model from Study 4 in Chapter 5. The higher relative implied volatility indicated by the bulge may be a consequence of the use (by traders) of a pricing model (Black-Scholes) that does not factor in the general presence of kurtosis in the returns from stocks; kurtosis leads to higher option prices (if they are efficiently priced by the market) and, in turn, to higher implied Black-Scholes volatilities for in-the-money and out-of-the-money options at moderate levels of volatility.

## Discussion

Overall, the best estimators of future volatility are also the best estimators of implied volatility. Both the raw standard and the average range historical measures underestimate implied volatility, just as they underestimate future volatility, when volatility levels are low; they overestimate implied and future volatility when levels are high. Adjusted standard volatility and the regression-based estimator of future volatility do much better. These estimators provide less distorted estimates of future volatility over their applicable ranges and fairly consistent, but slightly low, estimates of implied volatility.

Because of results like this—in which implied volatility appears to behave just like future volatility with respect to various historical measures and predictive estimates—much of the discussion has implicitly assumed that implied volatility reflects correct or "efficient" option premiums. Tacit in the discussion of the results found in Table 8–1 and Figure 8–11 was the assumption that option prices are fairly representative (at least on average) of actual fair value (as might be computed from an empirical distribution) and, therefore, that implied volatilities are efficient, except for errors in the option pricing model used to compute them (like the failure of Black-Scholes to take into account the general kurtosis in stock returns).

The results suggest that options are, indeed, priced efficiently in the sense that implied volatility (assuming that a good pricing model is used to compute it) comes close to what good prediction models yield as the expected future volatility. This is in no way meant to suggest that implied volatility is influenced by actual future volatility (in Chapter 5, a path analysis demonstrated that implied volatility was mostly determined by historical volatility, not future volatility); rather, it suggests that enough traders and hedgers, who set option prices by their activities in the market, are using reasonably good estimates of future volatility, as well as good pricing models, to support their option-related activities. Given the resources of many market makers and option houses, it is only natural to expect that influential hedgers and traders are using good future volatility estimates and pricing models. And, even though many arbitrageurs and traders still use Black-Scholes and standard historical volatility, they know how to make “discretionary” or subjective corrections to what may appear on their computer screens.

On average, implied volatility is best estimated by good predictors of future volatility. Standard historical volatility does not predict implied volatility very well, suggesting that, for the most part, it is not being used with Black-Scholes or other models as the volatility input by those who set option prices—unless they are doing a good job of subjectively correcting assorted important variables. In other words, actual option prices are more consistent with proper estimates of future volatility than with raw measures of historical volatility. Of course, all of this applies to averages—i.e., to expectation. There are many option prices that are consistent with raw historical volatility and this is not necessarily an indication that future price movement demands a different option premium; in other words, such options may simply be mispriced, for any number of reasons. Whether such options are truly mispriced and can be traded for a profit is a question which needs further study.

There is one consistent pattern in the data in Figure 8–11 and Table 8–1 that has not been discussed above. The pattern is rather curious and calls for an explanation. The two better estimates of future volatility—corrected standard volatility and the regression estimate—tend to underestimate implied volatility by a small amount, except at levels of volatility greater than 100%. Standard and average range volatility, uncorrected for regression to the mean and for capping, cross implied volatility between 75% and 80%. Given that mean volatility tends to be around

50%, the observed crossover point is rather high, which suggests that a similar tendency towards underestimation would be apparent with these measures of historical volatility even were it not for the distortions introduced by mean reversion.

One hypothesis is that the apparently higher-than-predicted implied volatility is due to distortions introduced into the statistics by the decidedly nonnormal distribution of volatility (see Figure 8–1), whether historical or implied. To test this hypothesis, a normalizing transformation (the natural logarithm is effective in this case) was applied, the statistics were re-calculated, and the logarithmically scaled expectations were transformed back to the original scale. The results of the re-analysis were extremely similar to those obtained from the original analysis. Apparently, the extended upper tail of the distribution of volatility was not the explanation.

Another hypothesis momentarily considered was that the underestimation of implied volatility by the various measures and estimates was caused by a different mean level of measured, estimated, or implied volatility in the current data sample, as compared to the original sample on which the estimates were based. The mean of the independent variable (the measured historical or predicted future volatility) was 50.8% and the mean of the dependent variable (implied volatility) was 55.2%. Although the volatility levels are in the same range as those found in samples used in earlier studies that did not examine implied volatility, there is a clear difference between the mean level of standard volatility and the mean of implied volatility. A similar difference also exists between the other volatility measures and implied volatility. No doubt, the fact that mean implied volatility is greater than the other volatility measures explains the observed pattern—even the size of the difference is in the correct ballpark.

The question then becomes, "Why should implied volatility be systematically higher than the other measures of volatility?" One possible answer has to do with the influence of kurtosis on the measure of implied volatility used in the study.

Implied volatility was calculated as a weighted average of the individual implied volatilities of all options with a valid price greater than zero that had between 3 and 48 days of life remaining. The weighting scheme gave greatest weight to at-the-money options that were flush with time value; however, out-of-the-money and in-the-money options also contributed significantly to the weighted average, i.e., to the overall implied volatility. As has been repeatedly pointed out in a variety of contexts, out-of-the-money and in-the-money options are the ones with premiums that are driven up by positive kurtosis. Since Black-Scholes (a model that ignores kurtosis) was used in the calculation of the individual implied volatility figures, the increase in option premiums caused by kurtosis is reflected in higher implied volatilities. The higher relative implied volatilities of the out-of-the-money and in-the-money options (the volatility smile) raises the weighted average, which is taken over all options trading on a given stock at a given time or bar. The weighted average, naturally, is the implied volatility used in the current study. Given what has just been said, the weighted average—implied volatility—would be expected to have a somewhat greater value than either historical volatility or estimated future volatility; indeed, it does.

The most important finding in this study, however, is that implied volatility (and the real option premiums doing the implying) is more consistent with good estimates of expected future volatility and less consistent with poor estimates like uncorrected historical volatility, whether standard or average range. Enough market participants are either using good future volatility estimates in their pricing models, or are "intuitively" adjusting (based on observation and experience) for the distortions that go with poor estimates, so as to cause options to be fairly efficiently priced. The options player who does not correct for volatility estimation error or use a good volatility-forecasting model will thus be at a severe disadvantage in the options arena.

## SUMMARY

The reader may have expected coverage of ARCH, GARCH, and other well-known approaches to modeling and forecasting volatility. Such approaches were not covered here for several reasons. One reason was that these approaches have been extensively studied and discussed by many practitioners, both academicians working in universities and professional quants operating within large financial institutions and options boutiques. And, because of the way scientific—and, no doubt, financial—paradigms operate in a social context, it is a near certainty that many more investigations within the standard paradigm will be conducted and published in the literature.

This is not to say that GARCH models, for instance, do not have a certain elegance and appeal; they clearly do: such models are rather miserly in their consumption of degrees of freedom, will correct for regression to the mean, and even embody a hypothesis regarding the process that underlies the signal (in this context, the price movements) under study. However, given their extremely wide dissemination and, presumably, use in the financial community, these models are unlikely to provide a trader or hedger with a significant edge; in the markets, one gains an edge, an advantage over the other participants, not by following the crowd (they usually lose), but by being creatively different.

The second reason for not covering models forged in the standard paradigm is because the intention in this work was to investigate techniques that are creatively different, that are not widely covered in the literature, and, hopefully, that are not being extensively employed. As short-term, electronic options traders, the goal was to find an edge—something that might reveal inefficiencies that could be exploited for a quick profit, even by off-floor traders. It seemed that to gain a practical advantage in the marketplace, a more empirical and flexible (even if less theoretically elegant) approach was required. Such an approach would make it easier to investigate uncommon variables and to discover ones that might contribute to better estimates of future volatility. An estimation model that incorporates variables not in common use as predictors of volatility (but which, nevertheless, have predictive value) is the kind of model that is most likely to produce estimates of future volatility with respect to which the market is not yet very efficient—estimates of volatility that might provide an off-floor trader or hedger with a tangible advantage in the options game. The methodology was, therefore, kept simple and the uncommon variables with the potential for predicting volatility were the focus of the investigations.

The third reason for not covering the standard paradigm is that the basic models are neither nonlinear (in the specific way required here) nor multivariate. It should be noted that there are multivariate generalizations of ARCH/GARCH and related models; these more complex models can handle multiple time series and, therefore, with additional adjustments for nonlinearity, might be used to study the kinds of relationships investigated in this chapter. However, at this stage of the investigation, why bring in all the excess complexity and theoretical baggage? It seemed that the well-known KISS philosophy—Keep It Simple, Stupid—was the more appropriate one in the current context.

What was learned in this chapter? It was found that historical skew and kurtosis have an influence on future volatility that is independent of the influence exerted by historical volatility. Future volatility is also affected by the location of a stock's price relative to its moving average, by the slope of the moving average, by the level and direction of Lane's Stochastic, by where the current price lies with respect to recent highs and lows, and by the month of the year. The influence of these items on future volatility often involves some statistical interaction with historical volatility; i.e., at different levels of historical volatility, these variables tend to have different effects on future volatility. The implication of the findings is that variables like those studied should be part of a good volatility forecasting model; the fact is that they can add unique information and predictive value—with respect to which the market may not yet be efficient—to a model designed to estimate future volatility.

Also studied in this chapter was the relationship between various predictors of future volatility, on the one hand, and the volatility implied by actual option prices, on the other. Poorer predictors, like raw historical volatility, evidenced regression to the mean when used to estimate implied volatility, just as they do when used to estimate future price movements. The better predictors, like corrected historical volatility, tended to systematically underestimate implied volatility because of the latter's higher average level. The higher average level of implied volatility, vis-á-vis both actual and estimated future volatility, appeared to be a result of higher prices for out-of-the-money and in-the-money options which, in turn, are an expected consequence of the leptokurtic distribution of returns.

Undoubtedly, a model could be developed specifically to predict implied volatility and, in turn, actual option prices in the marketplace—prices that are not necessarily always efficient or theoretically fair. Clearly, the relationship between historical, estimated future, actual future, and implied volatilities are of great importance and are worthy of further study.

## SUGGESTED READING

A good introduction to GARCH models appears in "Generalized Autoregressive Conditional Heteroskedasticity" (Bollerslev, 1986). Coverage of standard models for the prediction of volatility can also be found in *Modelling Stock Market Volatility* (Rossi, 1996) and in *Forecasting Volatility in the Financial Markets* (Knight and Satchell, 2002). *Estimation with Applications to Tracking and Navigation* (Bar-Shalom et al., 2001) and *Nonlinear Regression Analysis and Its Applications* (Bates and Watts, 1988) contain material on the least-squares fitting of polynomials and other nonlinear functions to real-world data. The subject of technical analysis using moving averages, Lane's Stochastic, and other indicators is covered in *The Encyclopedia of Trading Strategies* (Katz and McCormick, 2000) and in *The Encyclopedia of Technical Market Indicators* (Colby and Meyers, 1988). Finally, an excellent source on technical analysis using rigorously defined chart patterns is *The Encyclopedia of Chart Patterns* (Bulkowski, 2000).

CHAPTER 9

# Option Prices in the Marketplace

For the most part, previous chapters presented investigations concerned either with the price behavior of stocks (the underlying securities) or with the *theoretical fair premiums* of options. Theoretical option premiums are the premiums that can be considered fair given the behavior of the underlying securities; they are not the same as the premiums at which options are actually trading, which may be above or below what is theoretically fair. In a few instances earlier in this book, the focus of attention was on *implied volatility*, which is a measure of volatility derived from the premiums at which options are being bid or offered. In no case, however, was attention directly paid to real-market option premiums. That changes with the current chapter, which focuses on the premiums at which options are actually trading in the marketplace, rather than merely on theoretical fair premiums or stock price behaviors.

Do real options trade at prices consistent with theoretical premiums, such as those calculated with the conditional distribution methodology? How do actual option premiums compare to terminal price expectations? Can option pricing approaches like those investigated in this book provide an edge to the options trader? Can such models be used to generate profitable trades? These are among the questions the study in this chapter begins to address by comparing theoretical fair value with observed market premiums.



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

## DATA AND SOFTWARE

Two databases were used in the studies in this chapter. The first database contained data (extracted from Worden Brothers TC-2000 at www.worden.com) for 2,246 stocks and 1,834 bars. The data ran from January 2, 1996 through April 14, 2003. Database fields, for every bar and every stock, included *Date*, *Open*, *High*, *Low*, *Close*, *Volume*, and *Split Factor*. The data were back-adjusted for splits, but the original (uncorrected) prices and volumes were recoverable, if needed, thanks to the inclusion of a split factor among the database fields. The data were stored in a binary format for speedy access. For full details, consult the section on raw data in Chapter 4.

The second database used in the following studies contained raw options data. The data were extracted from files downloaded from www.stricknet.com, and spanned the period from March 1, 2002 through March 27, 2003. Each database record contained the following fields: *Quote Date*, *Underlying Symbol*, *Stock High*, *Stock Low*, *Stock Close*, *Stock Volume*, *Strike Price*, *Bid*, *Ask*, *Volume*, *Open Interest*, *Expiration Year*, *Expiration Month*, and *Option Type* (put or call). Database records were arranged in chains so that data for all options trading on a specified stock on a given day (or bar) could be retrieved as single unit with one efficient procedure call. There were, on average, 2,100 options chains collected per day on roughly 2,200 stocks. Each chain averaged about 56 options (including LEAPS). The total number of option quotes in the database was over 30 million. As with the stock database, the options database was maintained in a highly compressed, variable record size binary format designed for maximum speed and storage density.

The databases were maintained and accessed using libraries of routines written in ISO-standard C/C++. Standard C/C++ was also used for most of the analytic calculations. Some final analyses were performed in Excel, which was also employed for graphics and presentation-quality tables.

## STUDY 1: STANDARD VOLATILITY, NO DETRENDING

This study compares the premiums at which options are actually trading with theoretical fair values derived both from

Black-Scholes and from terminal option price expectation. Both actual market premiums and theoretical premiums are examined as a function of strike and standard historical volatility. The need to examine stock price is avoided, and the interpretation of the results is simplified, by the rescaling of all initial stock prices to a nominal $100, and the rescaling of all option premiums to be consistent with the rescaled stock prices.

Where do real option bid and offer prices fall with respect to Black-Scholes and to a good (even if somewhat noisy) theoretical estimate of fair value? Can money be made off pricing errors? Are real option premiums affected by the volatility payoff phenomenon? These are the questions addressed by the current study.

## Method

Because of the need to analyze both stock data and options data, the analytic procedure was more complex in the current study than in most of the studies found in earlier chapters.

The analysis began with the selection of a reference bar from one of the 1,834 bars in the binary stock database. A bar qualified for selection as a reference only if (1) the binary options database contained data for options trading on the date associated with that bar, and (2) there were `mtlcal` calendar days left until option expiration. In the current investigation, `mtlcal`, the time left in calendar days, was set to 14.

Next, a stock was selected from the 2,246 stocks in the binary stock database. A stock qualified for selection only if the following criteria were met: (1) the stock was alive over the 30 bars immediately preceding the reference bar (required for calculation of historical volatility), (2) the unadjusted (not corrected for splits) stock price on the reference bar was greater than $5, (3) options trading on the stock could be found in the binary options database, and (4) a valid options chain existed for the stock on the currently selected bar or date.

An options chain was retrieved for the selected bar of the selected stock. The chain consisted of all options (including LEAPS) trading (actively or not) on the selected stock at the specified bar. The retrieved chain was checked to verify that the options data matched the stock data; any data mismatch would indicate a database error and was programmed to trigger a so-called "fatal error" and terminate the analysis program. It should be noted that no such terminations were encountered.

Each option in the chain was then examined. First, the expiration date was checked. If the expiration date was `mtlcal` calendar days beyond the reference bar, then processing continued; otherwise attention shifted to the next option in the chain. Second, the option premiums were checked for validity. If the ask price was less than the bid price, or if the ask was more than $0.25 greater than the bid, the option premiums were considered invalid. A bid greater than the ask suggested either a data error or a crossed market; an ask too much greater than the bid indicated an illiquid market and the possible abdication by market makers. If the premiums appeared valid, the next step was performed; otherwise, attention again shifted to the next option in the chain. Third, the option's type was examined. If calls were being analyzed and the option was a put, attention moved to the next option in the chain. If the option was of the correct type, the next step was performed.

The next step was to rescale the price of the stock at the reference bar to a nominal $100 and to adjust the option's strike, bid, and ask prices to be consistent with the rescaled stock price. The terminal stock price, i.e., the price of the stock at option expiration, was also adjusted to be consistent with the rescaled stock price at the reference bar. These rescalings and adjustments simplified the analysis as well as the interpretation of the results by eliminating stock price (at the reference bar) as a variable in need of consideration.

Array indices were then calculated. The first array index, the row index, was calculated as

```
ix1 = (int)floor(0.5 + (nlsk-1)*(x1-bskmn)/(bskmx-bskmn))
```

where `nlsk` was the number of strike levels or categories, `x1` was the option's rescaled strike price, `bskmn` was the center of the lowest strike category or bin, and `bskmx` was the center of the highest strike level.

The second array index, the column index, was determined as

```
ix2 = (int)floor(0.5 + (nlvx-1)*(x2-bvxmn)/(bvxmx-bvxmn))
```

where nlvx was the number of volatility levels, x2 was the standard 30-bar historical volatility, bvxmn was the center of the lowest volatility category, and bvxmx was the center of the highest volatility level.

If the array indices were within range, i.e., $0 \leq$ ix1 $\leq$ nlsk and $0 \leq$ ix2 $\leq$ nlvx, then data for computing the bin statistics were accumulated in eight arrays. The first array was used to accumulate the event or bin counts; bin counts were accumulated by incrementing the value of the element addressed by the indices by one. The second array accumulated the sums of the first independent variable, the strike price, i.e., the strike price was added to the value of the element addressed by the indices. The third array accumulated the sums of the second independent variable, the historical volatility, in the same manner. The fourth array was used to accumulate the sums of the terminal option prices, which were computed from the terminal stock prices and strikes. The fifth array was used to accumulate the option bid premiums, while the sixth array was used to accumulate the ask premiums. Black-Scholes premiums—computed based on the historical volatility (x2), strike (x1), time left (mtlcal), and interest rate (at the reference bar)—were accumulated in the seventh array. Finally, the eighth array was used to accumulate the terminal stock prices for each combination of strike and volatility.

Once the arrays were updated, the next stock was selected. After all stocks were processed at the current bar, the next reference bar was chosen. Processing continued in this fashion until, for every valid stock and bar, options trading on that stock and bar had been examined.

The next step in the analysis involved taking the data accumulated in the eight arrays and using it to compute various bin statistics. For each combination of strike level (ix1) and volatility level (ix2), the following statistics were computed: (1) the mean strike, (2) the mean volatility, (3) the mean or expected terminal option price, (4) the mean option bid price, (5) the mean option ask price, (6) the mean Black-Scholes price, and (7) the mean terminal stock price. These seven statistics were calculated by dividing the values in arrays two to eight by the value in array one, the case count, on an element-by-element basis. If the case count for a given combination of strike and volatility was zero, all bin statistics were set to zero to indicate their noncomputability.

Finally, the case counts in the first array, as well as the additional statistics in arrays two to eight, were written to a standard text file as a set of tables, one table for each level of volatility. The entire analysis was performed for both puts and calls. The resultant text files containing the bin statistics were then loaded into an Excel spreadsheet so that further analysis could be performed, and tables and charts prepared.

## Results

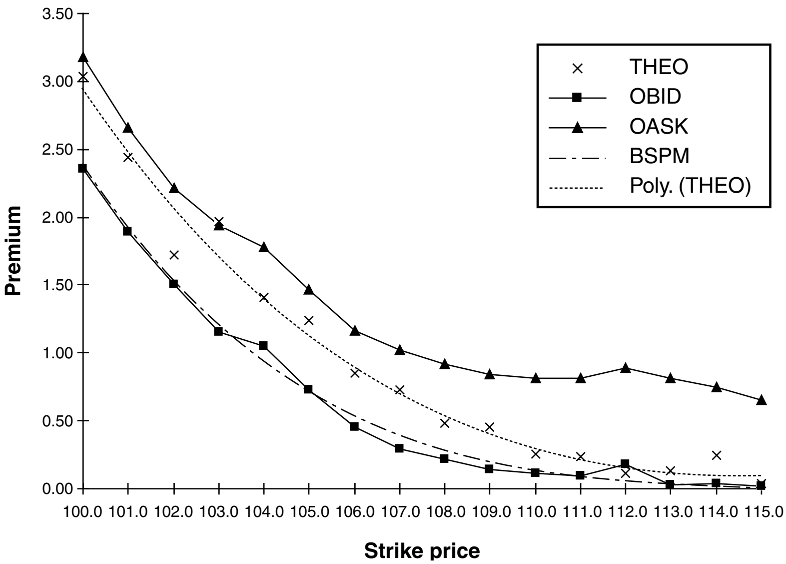
The results consisted of several large arrays of figures. Only four subsets of these figures are presented as graphs and discussed. Each subset corresponds to a few columns of data for a particular combination of option type and historical volatility and is discussed in its own section. The first section covers call options trading on stocks with historical volatilities centered at 30%, the second section covers put options on stocks with 30% volatility, the third section covers call options on stocks with 90% volatility, and the fourth section covers puts trading on stocks with historical volatilities centered at 90%.

### *Calls on Stocks with 30% Historical Volatility*

Figures 9–1 and 9–2 show mean premiums for call options trading on stocks having 30% volatility plotted as a function of strike. The solid line with small triangular markers (OASK) represents the mean ask price for calls trading on such stocks. The solid line with square markers (OBID) represents the mean bid price. The dotted line (Poly. THEO) represents a simple smoothing polynomial fitted to theoretical premiums based on the expected terminal prices of the calls, as computed from the terminal stock prices; the small x markers (THEO) are the raw (unsmoothed) theoretical premiums. Finally, the broken dash-dot line (BSPM) represents premiums computed with Black-Scholes. Figures 9–1 and 9–2 differ only in that Figure 9–1 shows data for calls that are out-of-the-money (strike greater than the nominal $100 stock price), while Figure 9–2 shows data for calls that are in-the-money (strike less than $100).

**FIGURE 9–1**

Average Out-of-the-Money Call Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 30% Volatility

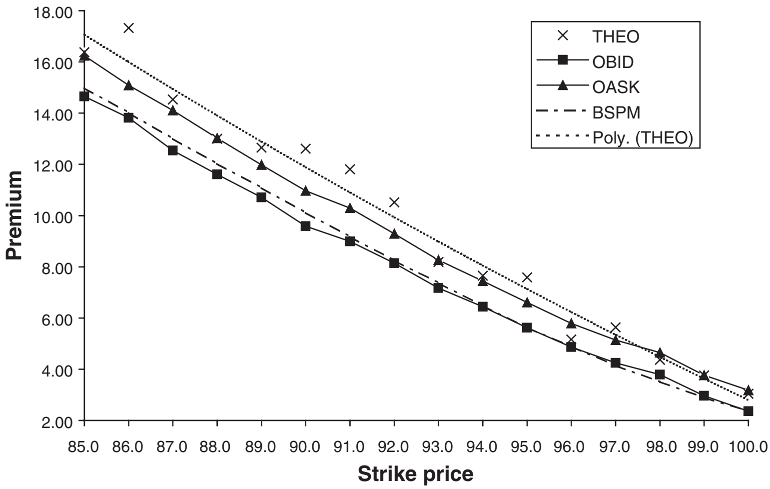



The Black-Scholes fair premiums are amazingly close to the posted bids, but well below the much higher offers (or asking prices), for all out-of-the-money calls trading on stocks having 30% historical volatility. The fair premiums based on terminal price expectation start off a little below the posted offers for calls that are at-the-money or just slightly out-of-the-money (left side of Figure 9–1); they end up just above the bids for deeply out-of-the-money calls (right side of the figure). Although all premiums except the offer are quite small for deeply out-of-the-money calls, the expectation-based fair values are several times larger than either the posted bids or the Black-Scholes price estimates.

As can be seen in Figure 9–2, the Black-Scholes premiums lie slightly above the bids for deeply in-the-money calls; for all other in-the-money calls, they closely hug the bids, just as was the case with out-of-the-money calls. For deeply in-the-money calls, the theoretical fair values derived from the terminal

**FIGURE 9–2**

Average In-the-Money Call Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 30% Volatility




expectations fall significantly above the posted offers or asking prices.

Why do Black-Scholes premiums computed using standard 30-bar historical volatility tend to hug the means of the posted bids? There are two possible (and not mutually exclusive) explanations. The first explanation is that the Black-Scholes estimates are low because of the regression to the mean to which raw historical volatility is subject. Since historical volatility was 30%, well below the mean of 52%, Black-Scholes can be expected to underprice the options; instead of the Black-Scholes dash-dot line appearing midway between the ask and the bid, it appears lower down in the chart, hovering around the bid. The second explanation is a much simpler one: traders place their bids at the Black-Scholes price because they believe that this is what the options are worth.

Why do deeply out-of-the-money calls have a theoretical fair value, determined by terminal price expectation, that is several times the fair value calculated with Black-Scholes? There are probably two reasons. The first is the tendency of Black-Scholes (but not terminal price expectation) to underprice options in low volatility situations. The second reason is kurtosis; out-of-the-money options have greater value than Black-Scholes indicates when the distribution of stock returns has longer tails than the log-normal distribution.

Finally, why are the theoretical expectation-based fair premiums greater than the posted asks for deeply in-the-money calls? Deeply in-the-money calls often seem to trade below fair value. Perhaps people do not like to buy “expensive” options that appear to be more risky and to provide less leverage. Less buying—less demand—translates into lower prices. Also, as options rise in price (as they go from out-of-the-money to in-the money), they get sold—either to take a profit or to cover a losing short position. And, there is the volatility payoff, which should, theoretically, have its greatest effect on high Delta, i.e., in-the-money options. The volatility payoff clearly affects terminal price expectation for options, but may not be reflected in the prices at which actual options are being offered.

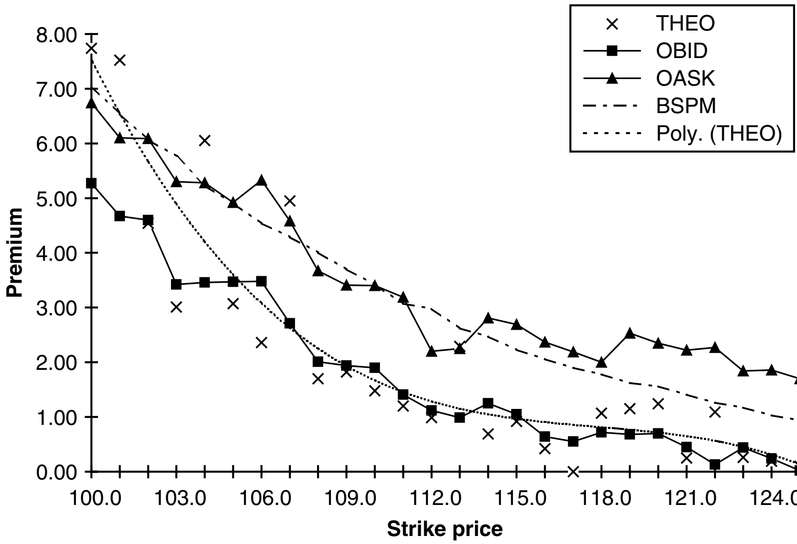
### *Calls on Stocks with 90% Historical Volatility*

Figures 9–3 and 9–4 show premiums for out-of-the-money and in-the-money calls on stocks having 90% volatility. Mean bid and ask prices at each strike are represented by the solid lines marked with squares and triangles, respectively. Theoretical fair premiums derived form Black-Scholes are represented by the dash-dot line, while those derived from mean terminal option price are represented by the x markers and by the dotted line (a smoothing polynomial fit to the data).

Black-Scholes fair premium tracks the posted offers for moderately out-of-the-money calls; it falls midway between the bids and the offers for deeply out-of-the-money calls and for most in-the-money calls. The theoretical fair value (based on terminal option price expectation) tends to hug the bids for out-of-the-money calls, rises above the offers for at-the-money options, falls between the bids and offers for moderately in-the-money calls, and goes back up to the ask (or even above) for deeply in-the-money calls.

**FIGURE 9–3**

Average Out-of-the-Money Call Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 90% Volatility



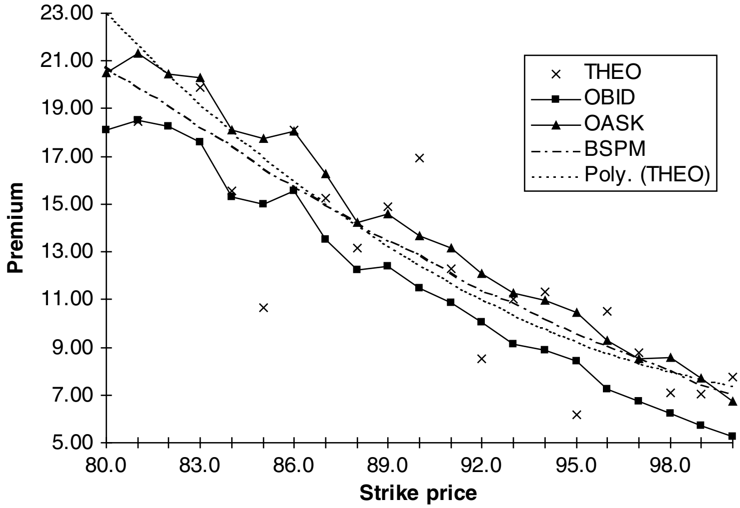

The raw (unsmoothed) theoretical premiums (depicted in Figures 9–3 and 9–4 by the x markers) are far more scattered with calls on the 90% volatility stocks than they were for calls on the stocks having 30% standard 30-bar historical volatility.

The last observation, that the premiums based on terminal option price expectation are more scattered for options on stocks with 90% volatility than for those on stocks with 30% volatility, is easy to explain; there are far fewer data points at the 90% volatility level than there are at the 30% volatility level and, consequently, the standard error of estimate, the "noise," is much greater in the former case.

The fact that Black-Scholes now tracks the offers (or falls midway between them and the bids) rather than tracking the bids (as it did for calls on stocks having 30% volatility) can be explained in terms of mean reversion. At 90%, volatility is high—well above the mean to which it will revert. This implies that Black-Scholes will tend to overprice the calls; in the earlier case

**FIGURE 9–4**

Average In-the-Money Call Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 90% Volatility




of calls on stocks with 30% volatility, Black-Scholes tended to underprice them. Another explanation is that, when volatility is high, many traders want to sell calls (either naked or as covered writes) and place their offers at what they believe is a fair price—the price indicated by Black-Scholes.

Why, then, are the true values (based on terminal option price expectation) near the bids for out-of-the-money calls? In other words, why are the actual call premiums (as represented by the average of the bid and ask prices) high relative to their true worth? Perhaps speculators (the naive ones) like to buy low-priced, out-of-the-money options on volatile stocks with the hope of striking a home run and achieving a windfall profit and thus drive up the prices of these options. At-the-money options may have premiums that are below what is, theoretically, fair because these options are the prime targets of covered writers looking to profit from the steady decay of time value. At-the-money options on volatile stocks are flush with time value, making them a good choice for covered writers intending to generate large returns.

Finally, it should be noted that there is often a wide gap between fair value as determined by Black-Scholes and fair value as determined by the presumably more valid (albeit noisier) terminal option price expectation. How wide can this gap become? As an example, consider calls struck at $112 trading on a $100 stock that has a volatility of 90%. In this case, the Black-Scholes price is around $3, while the expectation-derived fair value is just over $1, a difference that is quite noteworthy.

***Puts on Stocks with 30% Historical Volatility*** Figures 9–5 and 9–6 plot mean premiums against strike price for out-of-the-money and in-the-money puts trading on stocks with 30% standard historical volatility, respectively. The premiums plotted are the offers (solid line, triangular markers), the bids (solid line, square markers), the Black-Scholes numbers (broken, dash-dot line, no marker), and the fair premiums based on the mean terminal option prices (dotted line and x markers).

**FIGURE 9–5**

Average Out-of-the-Money Put Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 30% Volatility

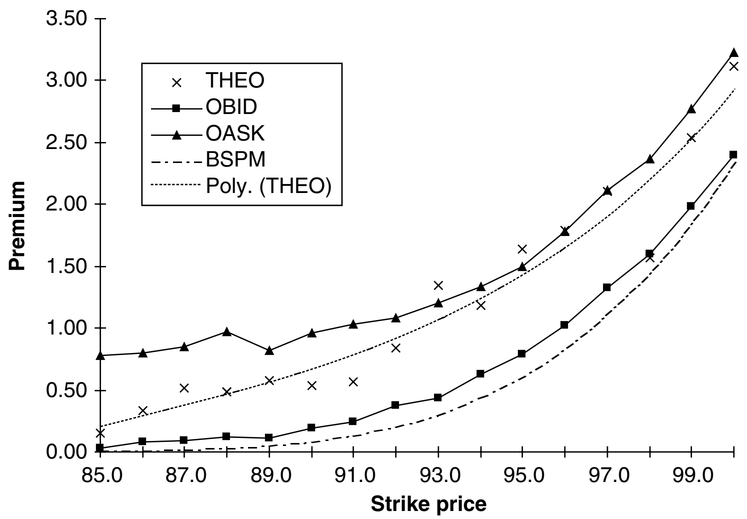
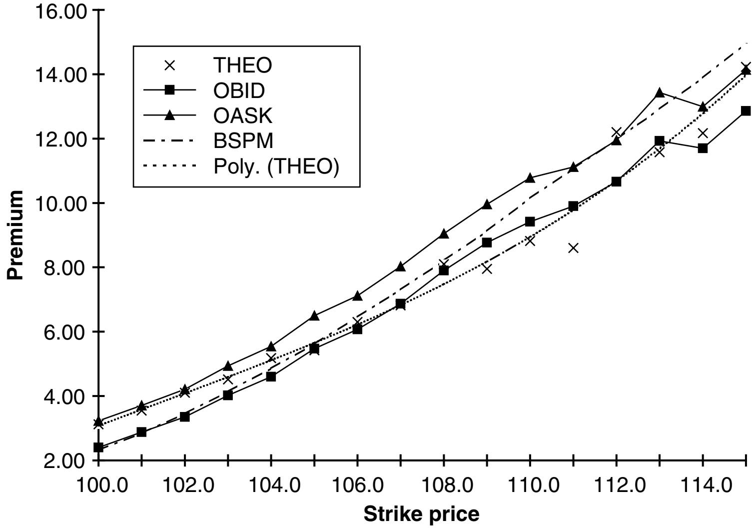


**FIGURE 9–6**

Average In-the-Money Put Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 30% Volatility




Black-Scholes premiums are lower than the posted bids for out-of-the-money and at-the-money puts, in-between the bids and offers for in-the-money puts, and at or above the offers for deeply in-the-money puts. Theoretical fair values based on terminal option price expectations are just below the offers for out-of-the-money and at-the-money puts, in-between the bids and offers for slightly in-the-money puts, and at or below the bids for moderately to deeply in-the-money puts (except for the last two data points where the theoretical fair value is back at the offers, probably due to sampling error in the latter).

Why do out-of-the-money puts appear expensive relative to Black-Scholes, but reasonably priced when future price expectation is used as the gauge? One possibility is the use of out-of-the-money puts as insurance. Hedgers are willing to pay slightly more than the Black-Scholes fair premium for the protection puts offer against sudden downdrafts and corporate scandals in uncertain times. Also, Black-Scholes underestimates value at low (e.g., 30%) volatility levels. The higher levels of the expectation-based value estimates can be explained by the fact that these estimates are unaffected by mean reversion in the volatility figures. Smart sellers might also place their offers well above Black-Scholes, closer to the actual values, in the hope of getting good prices.

Moderately in-the-money puts also appear expensive. Black-Scholes hovers just above the bid and expectation-based values hover at or below the bid. Perhaps this is due to a high demand for out-of-the-money calls, which drives up their prices. Higher prices for out-of-the-money calls translate (via put-call parity) to higher prices for in-the-money puts.

### *Puts on Stocks with 90% Historical Volatility*

Mean premiums for puts on stocks with 90% standard 30-bar historical volatility are plotted in Figures 9–7 and 9–8. Figure 9–7 contains premiums for out-of-the-money puts, while Figure 9–8 contains premiums for in-the-money puts. The premiums plotted are the asks or offers (solid line with triangular markers), the bids (solid line with square markers), the Black-Scholes

**FIGURE 9–7**

Average Out-of-the-Money Put Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 90% Volatility

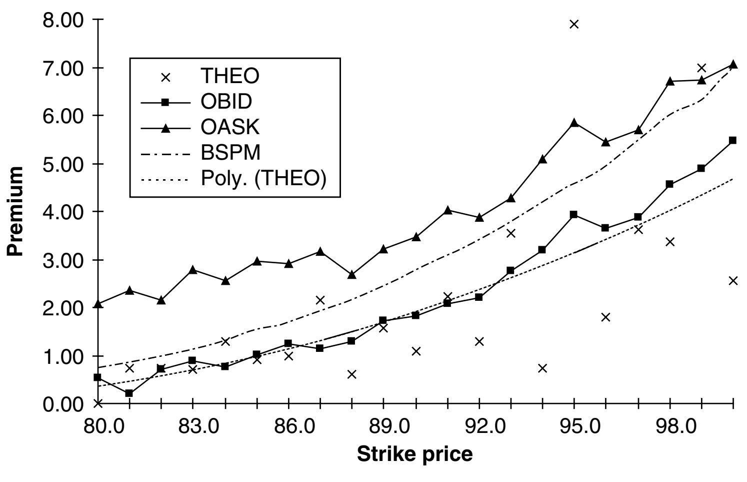
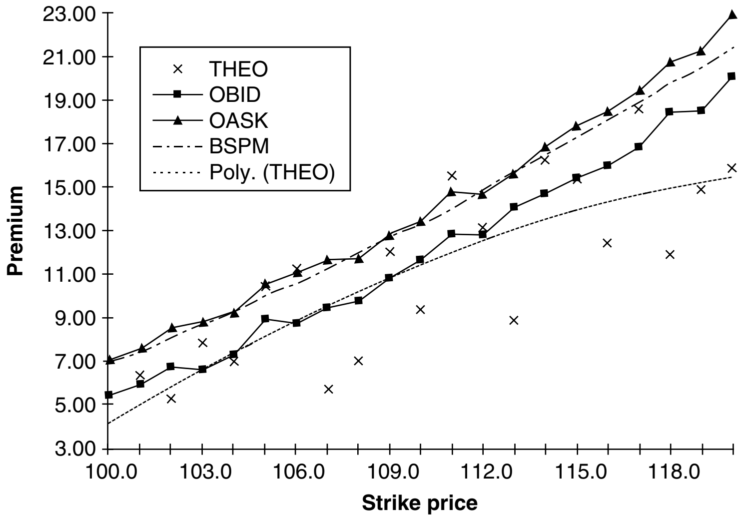


FIGURE 9–8

Average In-the-Money Put Premiums Derived from Terminal Stock Prices, Black-Scholes, Actual Bids, and Actual Asks, Plotted against Strike for Stocks with 90% Volatility




figures (dash-dot line with no markers), and the expectation-based fair premiums (dotted line, x markers).

For out-of-the-money puts on stocks with 90% historical volatility, Black-Scholes premiums fall between the bids and the asks; for in-the-money puts, Black-Scholes premiums tend to track the offers, and are far above the bids. Fair premiums based on terminal option price expectation are at or below the bids for both out-of-the-money and in-the-money puts; they fall well below the bids for near-the-money and deeply in-the-money options.

The higher Black-Scholes readings relative to the bids and offers of the puts on the stocks with 90% volatility are most likely an effect of mean reversion: at 30% volatility, Black-Scholes premiums were lower than they should be; now, at 90%, they are higher. The bids are greater than or equal to the expectation-based estimates of fair premium because puts are often bought for protection, even if the price is high, and are only occasionally sold (though they are considered equivalent, naked put selling is considered much riskier than covered writes using calls). At-the-money and deeply in-the-money puts have bids and offers that are much higher than the expectation-based estimates of fair value. This may be due to out-of-the-money calls on high volatility stocks being in high demand, together with the action of put-call parity. Another possibility is that traders are placing their bids and offers based on Black-Scholes, rather than on better estimates of fair value. Most likely of all, the puts are trading above their true value due to volatility-related trends that neither market participants nor Black-Scholes take into account; in other words, the volatility payoffs in the underlying stocks. Positive trends in the underlying stocks associated with the volatility payoff phenomenon would make puts—especially in-the-money puts with high Deltas—worth less than the value placed on them by Black-Scholes (or by market participants who accept the risk-neutral hypothesis), just as such trends cause in-the-money calls to be worth more than standard models indicate.

## SUMMARY

This chapter presented a brief investigation of real option prices in relation to two measures of theoretical fair value. Obviously, a lot more work of this kind needs to be done. Below is a list some of the avenues of investigation that may be pursued by those who feel so inclined.

The examination of the effect of historical skew and kurtosis on actual option prices is included in the "to do" list. Another topic for study is the usefulness of technical or other indicators in the prediction of real option prices, as well as in the prediction of the relationship between such prices and future price expectation.

An even more important subject to pursue is the interaction of current market premiums with volatility, strike, and other variables in the determination of terminal option price expectation (empirical fair value). Examining such interactions can reveal whether there are any pricing inefficiencies that can actually be exploited for a profit. This subject is so important that it is worth considering in further detail. For example, assume a set of \$100 stocks with historical volatilities of 90%. Further assume that \$110 calls with 14 calendar days remaining are trading on those stocks, that the bids have a mean premium of \$1 and a standard deviation of \$0.20, and the terminal option price expectation is \$1.10 (as was approximately the case in Figure 9–3). What would happen if a trader bought all the calls that have bids of \$0.60 or less (i.e., two standard deviations below the mean) and sold them later, at or near expiration? Would that trader make a profit? The answer depends on the terminal option price expectation: If it is unchanged in the subset of data points that correspond to calls with bids of \$0.60 or less, then the trader should be able to close out the positions at just over \$1, thus taking a hefty profit from the trade. However, perhaps there is an interaction between volatility and strike, on the one hand, and option price, on the other; perhaps the subset of data points has a terminal option price expectation of \$0.60. If that is the case, the trader will not even break even; he or she will incur a loss due to transaction costs. The more likely scenario is that the terminal expectation will fall somewhere between \$1.10 and \$0.60, making some profit-taking possible. Determining which outcome is likely can be accomplished through empirical studies of the kind that appear throughout this book and especially in this chapter; such studies are just the kind that are needed to locate exploitable pricing inefficiencies.

A final suggestion for further study concerns the large spreads between the bids and offers that appear in Figures 9–1 through 9–8. When trading, what is the best way to deal with these spreads? Where should one place limit orders (i.e., bids or offers)? In our experience as traders, that is an important question that needs a clear answer if one hopes to succeed in the options game.

*This page intentionally left blank.*

# CONCLUSION

The overall purpose of this book was to work toward discovering the factors that would contribute to a better option pricing model. We first examined the very element that option pricing models are designed to predict—fair value—and the influences that affect it. We then took a critical look at some of the standard option pricing models to determine how they work, what assumptions they make, and how they fare when put to the test on real-market data. We learned from the weaknesses we discovered and took the next step: exploring alternative ways of finding accurate estimates of fair value. We systematically investigated some potential nonstandard approaches that are unlike those that have already been heavily discussed as means to price options. We found that these nonstandard solutions seem to hold the promise of yielding excellent results in a difficult arena.

In this final chapter, we present you with the highlights of our findings and suggestions for further investigation.

## DEFINING FAIR VALUE

As a foundation to the study of option pricing, it was necessary to precisely define fair value in a useful manner. Beginning with the intuitive notion, which involves the concept of something being overpriced or underpriced, we moved on to a more exact



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

and operational definition of the term. Fair value was defined in terms of its relationship to arbitrage and to the mean future price of a security. More specifically, a security or option may be regarded as fairly priced when any potential to profit from either arbitrage, or from the difference between the current price and the mean (or statistically expected) future price, is nonexistent. Fair value is, therefore, closely related to the efficient market hypothesis. *An option or other security is fairly priced when it is efficiently priced, i.e., when its price discounts all current knowledge regarding the security's future price or arbitrage potential.* This definition of fair value is the one that underlies all the investigations in this book.

The operational definition of fair value makes reference to mean or expected future price. While the future and its possibilities are unknown until they occur, they can be characterized in terms of probability distributions. The mean (or expected) future price of an option, or other security, must be understood on the basis of probability distributions. A very simple Monte Carlo experiment demonstrated how the fair value of an option might be determined from certain assumptions regarding how stock price movements behave and are generated. Our attention then turned to popular models that attempt to determine fair value.

## POPULAR MODELS AND THEIR ASSUMPTIONS

### The Assumptions Themselves

One of the two most popular pricing models is Cox-Ross-Rubinstein. The primary assumption made by the standard version of this model is that stock prices are generated by a sequence of proportional, independent, random shocks, each of which is drawn from the same distribution. A complete derivation of the model was performed. The standard Cox-Ross-Rubinstein model, as demonstrated by a variety of techniques, bases its estimate of fair option price on the calculated mean terminal price of an option under the assumption that the terminal price of the stock (its price at option expiration) has a binomial distribution with a known mean and standard deviation or volatility. In the limit, as the number of time steps over which the model is computed (and over which the price shocks occur) approaches infinity, the option price estimate generated by the model approaches the mean terminal price of an option on a stock with a terminal price distribution that is log-normal. The terminal distribution, in fact, approaches log-normal very rapidly; even for a small number of time steps, the approximation to log-normal is very close. To summarize: the model assumes that stock prices in each time step move either up or down by some specified proportion or percentage, that these movements are random and independent of one another, and (thanks to the Central Limit Theorem in its proportional form) that as the number of time steps grows large, the theoretical option price generated by the model is consistent with the assumption of a log-normal distribution of returns.

Another popular model, Black-Scholes, was originally derived from the perspective of the elimination of arbitrage opportunity. An analysis using various techniques demonstrated that the Black-Scholes formula is actually a closed-form solution for the expected terminal price of an option under the assumption that the underlying stock has a log-normal distribution of returns.

During the investigation of these models, it was found that both the standard Cox-Ross-Rubinstein (computed with a large number of time steps) and Black-Scholes involve essentially similar assumptions and yield virtually identical theoretical option prices. It was further demonstrated that a log-normal terminal price distribution (or distribution of returns), which both models implicitly assume, is exactly what would be expected if returns are the result of a series of random, proportional price changes.

### Strengths and Weaknesses of Popular Models

The study of the Cox-Ross-Rubinstein and the Black-Scholes models also revealed their strengths and weaknesses. The strength of the Cox-Ross-Rubinstein (or binomial) model is its flexibility. For example, it can easily be adapted to price American-style options, to handle volatility or other influential factors that vary over time, and even to price options under different distributional assumptions. In addition, this model is transparent and easily understood. The main weakness (if it really is one in today's world of high-powered computing) is that this model is computationally intensive.

Unlike Cox-Ross-Rubinstein, the Black-Scholes model is not very adaptable. For example, it cannot be readily modified to price American-style options. Also, the Black-Scholes model is unable to price options under different assumptions regarding either distributions or the constancy of such variables as volatility and interest rates. The strength of this model is that it can be computed quickly and easily. Black-Scholes is also readily available: it has been programmed into most software packages that a trader, hedger, or other financial specialist might use.

## VOLATILITY PAYOFFS AND DISTRIBUTIONS

Until now, we have been discussing models that assume a log-normal distribution of returns. A log-normal distribution of returns has some interesting implications. Specifically, if a stock has a log-normal (or other positively skewed) distribution of returns, and if there is an equal probability that its price will either rise or fall over the holding period, then the average return will not be zero, as might naively be expected, but will be some positive number that increases with increasing volatility. In other words, there will be a volatility payoff. To have no volatility payoff (i.e., to have a return with an average of zero) when the returns have a positively skewed distribution, the probability of the stock falling must be greater than the probability of the stock rising. The phenomenon was shown to be more than just a mathematical construct. Experiments demonstrated that not only was it present in synthetic stock returns (prepared with a log-normal, pseudo-random number generator), but that it also occurred (to varying extents) with real stock returns taken from the NASD and from the NYSE.

## MATHEMATICAL MOMENTS

Given the importance of statistical distributions in determining the expected future prices of stocks and options, the mathematical moments of actual stock returns were examined. Moments are useful in characterizing distributions. The first two moments are the mean and the variance (or standard deviation). In the world of stocks and options, the mean represents trend and the standard deviation represents volatility. Additional important moments are skew and kurtosis.

## Moments and Holding Periods

The statistical moments were computed for returns that involved different holding periods, days of the week, times of the year, and day in relationship to option expiration.

If price movements are the result of a series of randomly independent shocks, then the volatility of returns from different holding periods should scale proportionately to the square root of time. For example, the standard deviation of returns from a one-week holding period should be half of the standard deviation of returns from a four-week holding period. In the context of holding periods, the study of returns demonstrated that this proportionality of volatility to the square root of time was approximately the case for individual stocks, but definitely not the case for such index tradeables as the QQQ and SPY. The scaling of volatility in proportion to the square root of time is assumed by such standard models as Black-Scholes. The fact that, for individual equities, volatility scales approximately with the square root of time suggests that individual price movements are at least approximately independent of one another; they are close to random in the sense that the price change in one time step cannot be easily predicted from the price changes in nearby time steps. For the QQQ and SPY index securities, the condition of independence was not satisfied. However, the statistical independence of the returns in one time step from those in another (i.e., the statistical independence of the price shocks) is an assumption made by the standard models.

## Moments and Distributions

Regardless of the holding period, the distribution of returns was found to differ significantly from log-normal. In our investigations, stocks typically exhibited negative skew across all holding periods; this could not be an artifact of the data analysis since a similar analysis of synthetic Monte Carlo stocks evidenced no skew. The distribution of returns also exhibited positive kurtosis. Kurtosis was greatest for short holding periods, but present for all that were examined. In other words, the distribution of returns had longer tails and a sharper peak than the normal distribution. Extreme gains and extreme losses, as well as prices that moved hardly at all, were far more frequent than predicted by a log-normal distribution, while moderate price movements in either direction were less frequent. Of course, Black-Scholes and similar models assume a log-normal distribution and, therefore, no kurtosis. A high level of kurtosis means that out-of-the-money and in-the-money options will have greater value than would be indicated by a model that assumes a log-normal distribution.

### Moments and Day of the Week

The moments characterizing the distribution of returns clearly differed with the day of the week when one bar (one trading day) returns were examined. First, volatility over the weekend was only marginally greater than the average volatility for all days during the week. This suggests that, when pricing options, bar time should be used—not calendar time, which is the measure most frequently employed. Calendar time treats weekends as if they had a volatility roughly equal to the square root of 3 (or $\sqrt{3}$) multiplied by the volatility of a typical weekday. Growth (or the mean return) for the average stock was negative on Monday, positive on Friday, and slightly positive on Wednesday; this is consistent with trading lore that suggests that Monday is a down day and that Friday is often strong. Volatility, the second moment, was highest on Monday, which is also generally considered a volatile day. Volatility was lowest on Friday for stocks and on Wednesday for indices. Skew was decidedly negative for every day of the week, with Monday being the most and Wednesday the least negative. For indices, skew was quite negative on Monday, but mildly positive on Wednesday and Friday. Kurtosis was strongly positive for every day of the week for both indices and stocks, with Monday having the most extreme positive kurtosis.

### Moments and Seasonality

Very strong seasonal effects were found in the first four moments of real-market returns. September was generally a month of negative growth and November through the end of June was a period of strong positive growth. Spikes of negative growth tended to appear in September and July.

Volatility varied dramatically with time of the year: well over 25% from its highest to lowest levels. Volatility was lowest in February, May, August, and November. It was highest in early January, early March, July, late September, and early October. In other words, volatility exhibited a clear, quarterly cycle that may be related (in timing) to the release of earnings reports. The seasonal impact on volatility was so large that it should be accounted for in any good volatility estimate or directly within an option pricing model.

Skew, like volatility, also differed with time of year. Positive skew was observed only in late December and early January. Extremely negative skew was observed in July through October, with the most negative skew occurring near the end of September. Modestly negative skew was observed at most other times of the year. It seems that crashes or sharp downdrafts characterize the period from July through October, while equally sudden upward thrusts in price are typical of late December and early January. When we say “a sharp downdraft” or “a sudden thrust,” the reference is to a sudden move that occurs against a backdrop of more moderate movements; we are not referring to trend. Seasonality was also evidenced in kurtosis. The lowest kurtosis was observed in March, April, and May, while the highest kurtosis was seen in December. Middle levels of kurtosis occurred over the rest of the year.

There is no doubt that the distribution of returns differs with the time of the year. At certain times of the year, lognormality is more violated than at others and volatility varies sufficiently to produce major differences in the worth of options.

### Moments and Expiration Date

The results demonstrated that the timing of returns relative to expiration date also affects their distributions. In terms of the first moment (mean return or growth), negative returns tended to occur on expiration Friday, despite the fact that, in general, Friday is a day of positive returns. Tuesday, usually a neutral day, had positive returns during expiration week. Returns were negative on the Wednesday and Thursday in the week prior to expiration, while the Friday and Wednesday that occur two weeks before expiration evidenced strong positive growth (normal for those days) and Monday showed negative growth (also normal).

Volatility, which might be expected to be greatest on expiration Friday, was not; instead, it was mildly elevated on the previous day, Thursday. The Monday of expiration week also exhibited elevated volatility, but it was well within the range of a normal Monday.

Not surprisingly, skew was most negative during expiration week, with Monday, Thursday, and Friday being the most negative days (in that order); these are typically the days of potential crashes or sharp spikes in a downward direction. The Friday prior to expiration Friday also had negative skew.

Along with negative skew, exceptional levels of positive kurtosis were found during expiration week. The highest levels were on Thursday and Monday, with Friday falling just behind. The extremely high level of kurtosis just prior to expiration was the primary finding of this study. During expiration week, there was an excess of extreme movement (up and down), and an excess of small or negligible returns, both at the expense of moderate ones, with a bias (due to the negative skew) towards the downside. The trader who sells premium near expiration will very frequently profit from a quiet market but, once in a while, will experience a stunning loss.

When thinking about pricing options, consideration must be given to the findings that concern the moments that characterize the distributions of the underlying stock prices. Even when using popular models (like Black-Scholes) with standard estimates of historical volatility, corrections should be made to compensate for the distortions that occur due to both violations of the assumptions of these models, as well as to variation with holding period, day of week, and time of year in the various moments, most notably, volatility.

## VOLATILITY

One of the moments, volatility, is central to all attempts to determine the theoretical fair value of an option. Volatility was, therefore, given a great deal of attention in our study of option pricing. Many important facts were learned.

When pricing options it should be understood that the volatility of interest is not the easily measured historical or implied volatility; rather, it is the volatility that will take place in the future, during the holding period prior to option expiration.

### Standard Historical Volatility as an Estimator of Future Volatility

One extremely important finding is that standard historical volatility exhibited regression to the mean. When historical volatility is extremely low, future volatility tends to be higher and options tend to have more value than would be suggested by a model that employs uncorrected historical volatility. At high levels of historical volatility, regression to the mean operates in the opposite direction. Future volatility—and, in turn, the actual worth of options—will be less than what a model like Black-Scholes might indicate, if the model is used with uncorrected historical volatility. If historical volatility is to be used as an input to a pricing model that does not implicitly correct for regression to the mean, corrections to historical volatility must be made before entering it as one of the inputs to the model. This is very easy to accomplish using any of several charts in Chapter 5, which show the relationship of future to historical volatility, or by using one of the simple polynomial regressions that appear in these charts.

In addition to regression to the mean, a capping effect was discovered: as historical volatility increases, future volatility reaches a plateau from which it rises no higher. It was discovered that a simple quadratic model could accurately describe the relationship between future and historical (or past) volatility.

Given the importance of volatility, and the tendency for mean reversion to distort option pricing, an attempt was made to use two measures of historical volatility (a longer- and a shorter-term one) to determine whether better predictions of future volatility could be obtained. It was demonstrated that each of the historical volatility measures contributed useful and independent information; therefore, they could be combined into a model that would yield more accurate estimates of the future volatility upon which option prices depend. A simple quadratic polynomial model that used two measures of historical volatility was able to produce distinctly better estimates of future volatility and to fully correct for regression to the mean.

### The Reliability of Different Measures of Volatility

At this point, we became concerned with the reliability of different measures of volatility. Although standard volatility (measured as the square root of the second moment of returns) is the most frequently used, there are other well-known ways of measuring volatility, including the average range method. Reliability is important because the reliability of an independent variable places a limit on its effectiveness when used in a model designed to predict some dependent variable. To determine the reliability levels of the various measures of volatility, methods were used from the world of psychometrics.

Implied volatility had the highest reliability, followed closely by average range volatility, and then by the other measures examined. Obviously, in any predictive model, it is probably worthwhile to use the most reliable measures for the variables of interest. With historical volatility measures there is a tradeoff. For measures like the standard one and the average range, reliability increases with the look-back period (in numbers of bars). However, an increase in the look-back period means a decrease in the representativeness of the measurement with respect to the current status of the market. For example, while a 500-bar volatility measurement, e.g., might be exceedingly reliable, its validity for predicting volatility during the forthcoming week will probably be less than that of a not quite so reliable 30-bar historical measure. In addition, based on its stability coefficient (the estimated correlation of perfectly reliable measures of volatility taken from nearby times), volatility appears to be a relatively enduring and predictable characteristic.

### Developing a Better Estimator of Future Volatility

The next step was to determine whether a notably better estimate of future volatility could be achieved by incorporating some of the findings from the study of moments. A regression model that involved a long- and a short-term measure of average range historical volatility (the most reliable measurement of historical volatility), together with three harmonics of a quarterly seasonal cycle, was fit to the data. This regression included linear, quadratic, and interaction terms for the historical volatility measurements, as well as terms consisting of sines and cosines for the three quarterly harmonics. The result was a dramatically improved estimator of future volatility.

Even with a standard model, the use of an estimator like the one just described will vastly improve the accuracy of the theoretical option prices that are generated. The regression estimate does not suffer from regression to the mean or the capping effect, and the relationship between actual future volatility and estimated future volatility over the range of 20% to 140% is described by a virtually perfect straight line. One can, of course, do even better by taking into account additional variables; the result being to reduce the uncertainty regarding the level of future volatility. At least, however, the model just described does not suffer from serious systematic distortions or biases.

### Implied Volatility

Implied volatility was also examined from a variety of perspectives. A disadvantage in the use of implied volatility is that it only permits relative evaluations of option worth (the worth of one option against another), rather than absolute appraisals. When used as a predictor of future volatility, implied volatility suffered from the same distortions—regression to the mean and capping—as standard historical volatility. It did not have any of the magical virtues sometimes attributed to it, like the ability to more accurately anticipate future stock price movements;

implied volatility was not a crystal ball that could be used to divine the future.

A path analysis showed that implied volatility was mostly determined by historical volatility. The correlation between implied volatility and future volatility was, to a great extent, a function of the influence of historical volatility on both. When implied volatility was treated as the target being influenced, historical volatility was shown to be the dominant influential factor; future volatility had only a minimal impact. An effect of insider knowledge (and other similar factors) on implied volatility may be present, but is not of sufficient strength to be very important from a practical standpoint. For the most part, implied volatility is set by the actions of traders and hedgers who influence option prices by placing trades based on the use of pricing models that employ various measures of historical volatility or estimates of future volatility computed from such measures.

## CONDITIONAL DISTRIBUTIONS

When the goal is to appraise options, the probability distribution used as the basis for determining the expectation of future stock prices does not have to be log-normal or even one of the many standard probability distributions. In fact, it is not even necessary to use just one distribution. Instead, one can employ a set of distributions—possibly determined empirically from actual stock price data—that is dependent on the status of different variables. This is the method of conditional distributions.

There are many ways to construct conditional distributions. They may be general mathematical forms that have been fit to the observed data, e.g., in terms of a set of moments or other coefficients. The distributions may also be stored in discretized form in an array, not unlike the way frequency histograms are constructed and stored. If there is a sufficient historical database, the method of conditional distributions can be an extremely effective one for pricing options.

The major problem with the use of conditional distributions is an avaricious demand for degrees of freedom. If there are too many variables in the model, there will not be enough data in even the largest data set to obtain stable results. Therefore, when using this methodology, it is imperative to simplify the option pricing problem as much as possible.

A number of investigations were performed using empirically-based conditional distributions to price options. The first investigation explored the use of distributions that were conditional upon the level of raw standard historical volatility and examined the resultant option prices. Almost by definition, conditional distribution methodology yields correct option prices, at least insofar as the data on which the distributions are based is a reasonably representative sample and there are a sufficient number of data points to keep the noise in the results to an acceptable level.

### Raw Historical Volatility: Conditional Distributions versus Black-Scholes

When compared to the prices derived from conditional distributions, Black-Scholes flagrantly overpriced options at high levels of volatility when the standard, raw measure was used as the model's volatility input. Black-Scholes underpriced options when historical volatility was low and there was little time left. The reason, of course, is that historical volatility regresses to the mean. Since Black-Scholes does not correct for this, it yields incorrect prices when raw historical volatility is used in the calculations. Since the distributions used to price the test options were specifically conditional upon raw historical volatility, there was an implicit correction of such phenomena as regression to the mean. Option prices computed using these distributions were correct, although perhaps having higher errors of estimate than would be the case if the distributions were conditional on better estimates of future volatility.

Trends in the data also seemed to affect theoretical option prices. More will be said about this later.

### Regression-Estimated Volatility: Conditional Distributions versus Black-Scholes

The next test involved regression-estimated volatility. In this test, the distributions were conditional upon a high quality estimate of future volatility. When compared against Black-Scholes, the impact of regression to the mean on the latter model disappeared. The conditional distributions produced estimates similar to those produced in the previous study, while Black-Scholes produced theoretical prices that more closely matched those from the conditional distributions. However, another phenomenon became apparent.

At low levels of volatility there was a fairly close correspondence between theoretical call prices derived from Black-Scholes and from conditional distributions. As the volatility level rose, call premiums from the conditional distributions became relatively large compared to those from Black-Scholes, while put premiums became relatively low. It turns out that the volatility payoff discussed earlier affected calls and puts, as well as the underlying stock; this flies directly in the face of the assumption of a risk-neutral world that is made by most pricing models. At very high levels of volatility, not only is the expected terminal price of the stock increased, but also the increase is reflected in the theoretical terminal prices of calls and puts trading on that stock.

### Detrended Distributions: Conditional Distributions versus Black-Scholes

A reanalysis was performed with statistically detrended distributions that were conditional upon a high quality regression estimate of future volatility. Conditional distributions are flexible in that various theoretical manipulations can be performed on them, e.g., detrending (or adjusting the first moment to zero). With trends removed, the difference between Black-Scholes and conditional distribution-based premiums at high levels of volatility were greatly reduced. The match between the premiums from the two methods became much closer. Significant differences, however, still remained between the premiums from the two models. These differences were clearly the result of the more leptokurtic distribution of returns found with real stock prices than with the log-normal distribution that underlies the Black-Scholes model. The disparity in premiums between the models was large and of practical significance. Specifically, as indicated by conditional distribution-based premiums, out-of-the-money options had greater value than Black-Scholes would suggest.

### Distributions and the Volatility Payoff

The way in which theoretical premiums responded to detrending raised an important issue. Should distributions have first moments that reflect the volatility payoff or should they not? The standard risk-neutral world stance says that they should not reflect the volatility payoff. However, a speculator buying a short-term option might want to know the expected price down the line, as it would be computed using distributions that have not had the volatility payoff effect (the positive first moment) eliminated. The findings suggest that there is a need for a re-evaluation and further analysis of how risk payoffs operate in the world of options. The risk-neutral world assumption, obviously, stems from the idea that all risk can be hedged away by using options. The means of reconciling these two observations is an important question for future study.

### Skew and Kurtosis as Conditioning Variables

Historical skew and kurtosis can be easily incorporated as conditioning variables in a pricing model based on conditional distributions. Doing so allows not only the generic skew and kurtosis of the market as a whole to be accounted for in theoretical option prices, but also the historical skew and kurtosis of individual stocks. In our investigations, we also observed complex relationships between historical skew, kurtosis, and theoretical premium. The differences in premium at different levels of historical kurtosis and skew were, at certain levels of volatility, sizable enough to be of definite practical importance. The results suggest that historical skew and, especially, historical kurtosis should be inputs to a good option pricing model. These variables are as easily computed as standard historical volatility.

### Conditional Distributions and Venue

Through the use of conditional distributions, it was demonstrated that venue also has a distinct effect on the theoretical fair value of an option. At low levels of volatility, options trading on NASD stocks had greater value than those trading on equivalent NYSE stocks. For example, take a $115 call on a $100 stock that has a raw volatility of 40%: if the stock is on the NASD, the theoretical fair value would be $0.31; on the NYSE, its fair value would be $0.20; and, if estimated by Black-Scholes, it would be $0.14. At high levels of volatility, greater value was found with options on NYSE stocks.

### Technical Indicators as Conditioning Variables

In the study of pricing by conditional distributions, at least one technical indicator was discovered to have value as a conditioning variable. Lane's Stochastic Oscillator, which is virtually useless for directional trading, is not quite so useless when it comes to pricing options. Both puts and calls had greater value when the stochastic oscillator has just crossed above 20% than when it has just crossed below 80%; this applies whether the theoretical option prices are computed with the original conditional distributions or with the detrended ones. In general, the findings suggest that technical indicators, although of little value for directional traders, may have value when it comes to estimating various aspects of the future distributions of returns (other than the first moment, i.e., trend) and, in turn, of fair option premiums.

## PRICING OPTIONS USING NONLINEAR MODELING

The main problem involved in pricing options is that of constructing a continuous, nonlinear function that maps a set of influential variables to a theoretical fair value. One approach to pricing options is to use a general, nonlinear modeling technique. For example, a neural network or polynomial regression can be trained or fit to the observed terminal option price data (assuming that the data sample on which the model is developed is sufficiently large and representative). This approach should result in an exceptionally good option pricing model. However, the techniques of general-purpose nonlinear modeling have a tendency to curve-fit the data, i.e., to capture not only the critical relationships, but also the spurious patterns of noise that are specific to the sample.

## Neural Networks versus Polynomial Regressions versus Black-Scholes

In an investigation of general nonlinear modeling techniques, it was found that neural networks can accurately emulate Black-Scholes, with the implication being that they might also be able to accurately model the true relationships found in real-market data. Multivariate polynomial regressions were also able to emulate Black-Scholes; however, polynomial models could only achieve a good fit over a more limited range of input values and so a small set of polynomial models would probably be required to cover the entire domain.

On tests involving real-market data, both neural networks and polynomial regressions worked quite well, providing results that were distinctly better than, e.g., the standard Black-Scholes model. And these were only first generation attempts. There is no doubt that, with additional effort and sophistication, even better results can be obtained.

## Strengths and Weaknesses of Nonlinear Modeling Techniques

Polynomial models are flexible, fast to train, and tend to capture the true relationships in the data necessary for pricing options while mostly ignoring the noise—in other words, they were less susceptible to curve-fitting than neural networks.

Neural networks were also found to be extremely effective in capturing the relationships in the data required for pricing options. Furthermore, a single neural network can handle the entire domain of inputs. However, neural networks have a tendency to over-fit the data.

There are several methods available for controlling or otherwise dealing with curve-fitting. One means of coping with this problem is through use of a "biased statistics" approach. Another answer may be found in the use of *regularization*, e.g., the development of a solution that not only maximizes the fit to the data, but also satisfies a set of prespecified constraints, such as that Delta be monotonic with respect to volatility.

### Hybrid Models

A third way to avoid curve-fitting, while increasing the ability of a model to capture the true relationships involved, is through hybridization. Existing knowledge regarding a subject domain can be incorporated into a hybrid neural model. Even if that information is only an approximation, the model can tweak itself to give the best possible fit to the data. The simple hybrid neural model examined during the investigation was able to achieve a reasonably good fit to the data with by far fewer free parameters than either a straight neural model or a polynomial regression; it also had much less tendency to curve-fit sampling artifacts.

The use of multivariate polynomial regressions, neural networks, and hybrid models appears to be an extremely fertile area for future investigation. There is no doubt that, with appropriate techniques, an extremely robust, as well as fast, pricing model can be developed.

## VOLATILITY REVISITED

In the course of some of the studies, certain questions regarding volatility arose. For example, does volatility mediate the effect of historical skew on option price? Can historical skew and kurtosis improve the prediction of future volatility? And, what about the whole world of so-called "technical indicators" used by technical analysts. Finally, what is the relationship between estimated future volatility and implied volatility?

### The Impact of Historical Skew, Kurtosis, and Historical Volatility on Future Volatility

Historical skew and kurtosis have a definite impact on future volatility that is independent of the level of historical volatility. In addition, the predictive relationship between historical skew, kurtosis, and historical volatility, on the one hand, and future volatility, on the other, involves interactions: the relationship between either skew or kurtosis, and future volatility differs in form at different levels of historical volatility.

### Using Technical Indicators in the Prediction of Future Volatility

It is known that most standard technical indicators are virtually useless for directional trading (for predicting the direction of price movements). However, it seems these indicators are not quite so useless when it comes to predicting future volatility, regardless of the direction of the movement involved. At each level of historical volatility examined, a relationship was found between the readings from various indicators and future volatility. Under this rubric, the location of the last price relative to its moving average, the location of the stochastic oscillator and the direction of its movement, and the location of the current price relative to its recent historical range (which reflects support and resistance), were all found to have sizable effects on future volatility. Many of these effects were large enough to be of immediate practical importance, i.e., 20% or even 30% differences in expected future volatility were observed with different readings from the indicators. Therefore, although technical indicators may be useless for directional trading, they may not be so useless to those trading volatility or pricing options—at least, not yet. Once their value becomes common knowledge, and these indicators become widely used as components in volatility forecasting models, their effectiveness may disappear.

## OPTION PRICES IN THE MARKETPLACE

It is important to study not only theoretical fair value, but also the actual prices of options in the marketplace and, especially, the relationship between the two. It is fairly straightforward to compute theoretical fair values in different ways and to make comparisons with mean option prices, both bids and offers, using the techniques developed in this book.

By performing such an analysis, a number of observations were made. It became clear that different estimators of fair value fall at different places relative to the bids and offers actually seen in the options market. For example, at low levels of volatility, Black-Scholes tended to hug the bids, while at higher levels of volatility it tended to hover near the asks; this was probably due to the effects of mean reversion when using standard volatility. Most in-the-money calls trading on stocks with 30% volatility had prices that were not too deviant from Black-Scholes (which often falls between the bid and offer), but that were significantly lower than their expectation-based fair value—a truer, but noisier, estimate of an option's real worth. Deeply in-the-money puts on stocks with 90% volatility had bids and offers that were much higher than their theoretical fair value (computed on the basis of future price expectation), but were roughly on par with Black-Scholes. The volatility payoff may have been responsible for that occurrence—the true worth of these puts may have been reduced because of the tendency for positive movement in the stocks to occur over the holding period due to volatility, while the market prices for these puts may not have been reduced because of the assumption of risk neutrality by options traders.

Another observation made in the course of the investigation is that, on average, there is a large spread between the bid and the offer.

During our studies, many questions were raised, including whether, under certain conditions involving strike and volatility, the expected terminal price of an option would remain the same given a selection from the distribution of current option prices. This last question is critical for determining whether it is possible to grab profits by exploiting mispricings on individual options. The analyses required for obtaining the answer remains on the agenda for future research.

## SUMMARY

We have done a lot of the groundwork for anyone wishing to develop better option pricing models. At the very least, you now know how critical it is to have a good pricing model which, among other things, should be consistent with the distributional quirks of the actual markets that underly the options you are interested in. If you do not have a good model, you need to be able to compensate for the flaws in a poor one and we have given some pointers on how to do that. You also now know that it is imperative to use good estimates of future volatility and that if you can find a nonstandard technique for detecting instances of changing future volatility, then you have found a very big edge that can prove quite profitable.

In the course of discovering a series of unique, workable approaches that will lead to better option pricing, quite a few issues arose concerning the way to maximize the effectiveness of these approaches. While we have identified and discussed quite a number of these issues, there is still more work to be done. Some of this work is rather tricky and requires special knowledge and ability to finesse the models. We have made a lot of progress in that realm, but we have not been able to include it because it is beyond the scope and spatial limitations of this book. For example, since it is unlikely that a unitary polynomial model could cover the full range of all the input variables that would be desired in a complete pricing model, one unexplored (herein) direction is to assemble a set of polynomial models into a mosaic that would cover the desired input domain; this would lead to coherent, fairly complete, and practically useful pricing models. Another direction to pursue is the development of trading systems based on some of the models we have discussed and which, e.g., make use of some of the more unique conditioning variables. We will be continuing our research in this realm and hope we have given you the help you need to do the same. Let us know how you are coming along!

## COMPANION SOFTWARE AVAILABLE

We invite all readers to visit our Web site at ***www.scientific-consultants.com*** or to e-mail us at ***katz@scientific-consultants.com*** (please note: we do not sell or otherwise distribute our lists). Our Web site contains updates about our research, as well as other information you may find useful. It is there that those who wish to replicate and expand on our research may obtain a copy of the data and software that were used during the course of our studies. Included is information on the following Companion Software items:

◆ **LIBRARIES FOR:** *stock database management* (routines for allocating a stock data object, reading records from the stock database, and reading records from a corresponding supplementary database that might contain implied volatilities); *options database management* (routines for creating an options database object, an options chain structure, and for retrieving option chains from the database with indexing by the date and symbol of the underlying stock; as well as additional related functions); *Black-Scholes pricing model* (routines for calculation of put and call prices, as well as put, call, and straddle implied volatilities); *mathematics and utilities* (routines for allocating memory, handling errors, as well as for conversion and calculation of Julian and other dates, calculation of historical volatility, and other general purpose supporting functions); *correlation matrix* (routines for creating a correlation matrix object, for accumulating data, and for calculating a correlation matrix from the data accumulated); *multiple regression* (routines for creating a multiple regression object, for accumulating data on a case-by-case basis, for calculating the regression based on the accumulated data, and for printing a regression report to a file); *the hybrid neural model* (contains functions for creating, loading, saving, and disposing of a hybrid neural object for setting and examining parameters such as the gradient weights, for setting inputs and retrieving outputs, for firing the model, and for causing the model to learn).

◆ **UTILITIES AND PROGRAMS FOR:** *general options data-base management* (creating database, extracting www.stricknet.com options data from zip files, saving the data in ultra-compressed format in the options database, generating index files for the database, etc.); *stock database management* (extracting data from Worden Brothers TC-2000 and placing it in standard compressed binary database format, cleaning up and editing the data to correct errors, etc.); *developing and testing neural networks*, i.e., the N-Train neural network development system.

◆ **CHAPTER-SPECIFIC SOFTWARE:** Sets of specific software code corresponding to each of the chapters. This is the code that actually implements the analyses described in the studies discussed in the chapters.

You may also contact us at the following address:

**SCIENTIFIC CONSULTANT SERVICES, INC.** **Phone & Fax:**
**20 Stagecoach Road, Selden, New York 11784** **631-696-3333**



Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

# BIBLIOGRAPHY


Aberth, Oliver. *Precise Numerical Methods Using C++*. San Diego, CA: Academic Press, 1998.

Bar-Shalom, Yaakov, Xiao-Rong Li, and Thiagalingam Kirubarajan. *Estimation with Applications to Tracking and Navigation*. New York: John Wiley-Interscience, 2001.

Bates, Douglas M., and Donald G. Watts. *Nonlinear Regression Analysis and its Applications*. New York: John Wiley and Sons, 1988.

Black, Fischer. *Business Cycles and Equilibrium*. Cambridge, MA: B. Blackwell, 1991.

Black, Fischer. *Exploring General Equilibrium*. Cambridge, MA: MIT Press, 1995.

Bollerslev, Timothy. "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*. Vol. 31, 1986, pp. 307–327.

Bulkowski, Thomas N. *Encyclopedia of Chart Patterns*. New York: John Wiley and Sons, 2000.

Bulkowski, Thomas N. *Trading Classic Chart Patterns*. New York: John Wiley and Sons, 2002.

Caplan, David L. *The New Options Advantage*. Chicago, IL: Probus Publishing, 1995.

Caplan, David L. *The New Option Secret: Volatility*. Columbia, MD: The Trader's Library, 2000.

Chriss, Neil A. *Black-Scholes and Beyond*. New York: McGraw-Hill, 1997.

Colby, Robert W., and Thomas A. Meyers. *The Encyclopedia of Technical Market Indicators*. New York: Dow Jones-Irwin, 1988.

Cox, John C., and Mark E. Rubinstein. *Options Markets*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

Crocker, Linda, and James Algina. *Introduction to Classical and Modern Test Theory*. Belmont, CA: Wadsworth Publishing, 1986.

Cronbach, Lee J. *Essentials of Psychological Testing,* 3rd ed. New York: Harper and Row, 1970.

Friedentag, Harvey C. *Stocks for Options Trading*. New York: St. Lucie Press, 2000.

Gershanfeld, Neil. *The Nature of Mathematical Modeling*. Cambridge, England: Cambridge University Press, 1998.

Harmon, Harry H. *Modern Factor Analysis*, 3rd ed. Chicago: University of Chicago Press, 1976.




Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

Hays, William L. *Statistics*. New York: Holt, Rinehart and Winston, 1963.

Heise, David R. *Causal Analysis*. New York: John Wiley and Sons, 1975.

Hirsch, Yale. *Don't Sell Stocks on Monday*. New York: Penguin, 1987.

Johnson, Mark A. *The Random Walk and Beyond*. New York: John Wiley and Sons, 1988.

Johnson, Norman L., Samuel Kotz, and N. Balakrishan. *Continuous Univariate Distributions*. New York: John Wiley-Interscience, 1994.

Jurik, Mark (ed.). *Computerized Trading*. New York: New York Institute of Finance, 1999.

Kaeppel, Jay. *The Four Biggest Mistakes in Option Trading*. Columbia, MD: The Trader's Library, 1998.

Karian, Zaven A., and Edward J. Dudewicz. *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*. Boca Raton, FL: CRC Press, 2000.

Katz, Jeffrey Owen. "Developing Neural Network Forecasters for Trading," *Technical Analysis of Stocks and Commodities*. April, 1992, pp. 58–68.

Katz, Jeffrey Owen, and Donna L. McCormick. *Calandar Effects Chart*. Selden, New York: Scientific Consultant Services Inc., 1990.

Katz, Jeffrey Owen, and Donna L. McCormick. "Vendor's Forum: The Evolution of N-TRAIN," *PCAI*. March/April, 1993, pp. 44–46.

Katz, Jeffrey Owen, and Donna L. McCormick. "Neural Networks: Some Advice to Beginners," *Trader's Catalog and Resource Guide*. Vol. II, No. 4, 1994, p. 36.

Katz, Jeffrey Owen, and Donna L. McCormick. "Neurogenetics and its Use in Trading System Development," *NeuroVe$t Journal*. July/August, 1994, pp. 8–11.

Katz, Jeffrey Owen, and Donna L. McCormick. "Introduction to Artificial Intelligence: Basics of Expert Systems, Fuzzy Logic, Neural Networks, and Genetic Algorithms," *Virtual Trading*. J. Lederman and R.A. Klein (eds.). Chicago, IL: Probus Publishing, 1995, pp. 3–34.

Katz, Jeffrey Owen, and Donna L. McCormick. "Neural Networks in Trading," *Virtual Trading.* J. Lederman and R.A. Klein (eds.). Chicago, IL: Probus Publishing, 1995, pp. 35–64.

Katz, Jeffrey Owen, and Donna L. McCormick. "On Developing Trading Systems," *Technical Analysis of Stocks and Commodities*. November, 1996, pp. 46–60.

Katz, Jeffrey Owen, and Donna L. McCormick. "A Rule-Based Approach to Trading," *Technical Analysis of Stocks and Commodities*. December, 1996, pp. 22–34.

Katz, Jeffrey Owen, and Donna L. McCormick. "Developing Systems with a Rule-Based Approach," *Technical Analysis of Stocks and Commodities*. January, 1997, pp. 38–52.

Katz, Jeffrey Owen, and Donna L. McCormick. "Genetic Algorithms and Rule-Based Systems," *Technical Analysis of Stocks and Commodities*. February, 1997, pp. 46–60.

Katz, Jeffrey Owen, and Donna L. McCormick. "Seasonality and Trading," *Technical Analysis of Stocks and Commodities*. April, 1997, pp. 50–61.

Katz, Jeffrey Owen, and Donna L. McCormick. "Cycles and Trading Systems," *Technical Analysis of Stocks and Commodities*. May, 1997, pp. 38–46.

Katz, Jeffrey Owen, and Donna L. McCormick. "Lunar Cycles and Trading," *Technical Analysis of Stocks and Commodities*. June, 1997, pp. 38–46.

Katz, Jeffrey Owen, and Donna L. McCormick. "Evaluating Trading Systems with Statistics," *Technical Analysis of Stocks and Commodities*. July, 1997, pp. 50–61.

Katz, Jeffrey Owen, and Donna L. McCormick. "Using Statistics with Trading Systems," *Technical Analysis of Stocks and Commodities*. August, 1997, pp. 32–38.

Katz, Jeffrey Owen, and Donna L. McCormick. "Sunspots and Market Activity," *Technical Analysis of Stocks and Commodities*. September, 1997, pp. 46–54.

Katz, Jeffrey Owen, and Donna L. McCormick. "Adding the Human Element to Neural Nets," *Technical Analysis of Stocks and Commodities*. November, 1997, pp. 52–64.

Katz, Jeffrey Owen, and Donna L. McCormick. "Exits, Stops and Strategy," *Technical Analysis of Stocks and Commodities*. February, 1998, pp. 32–40.

Katz, Jeffrey Owen, and Donna L. McCormick. "Testing Exit Strategies," *Technical Analysis of Stocks and Commodities*. March, 1998, pp. 35–42.

Katz, Jeffrey Owen, and Donna L. McCormick. "Using Trailing Stops in Exit Strategies," *Technical Analysis of Stocks and Commodities*. April, 1998, pp. 86–92.

Katz, Jeffrey Owen, and Donna L. McCormick. "Using Barrier Stops in Exit Strategies," *Technical Analysis of Stocks and Commodities*. May, 1998, pp. 63–89.

Katz, Jeffrey Owen, and Donna L. McCormick. "Barrier Stops and Trendlines," *Technical Analysis of Stocks and Commodities*. July, 1998, pp. 44–49.

Katz, Jeffrey Owen, and Donna L. McCormick. "Case Study: Building an Advanced Trading System," *Computerized Trading*. Mark Jurik (ed.). New York: New York Institute of Finance, 1999, pp. 317–336.

Katz, Jeffrey Owen, and Donna L. McCormick. "Trading Stocks with a Cyclical System," *Technical Analysis of Stocks and Commodities*. February, 1999, pp. 36–42.

Katz, Jeffrey Owen, and Donna L. McCormick. *The Encyclopedia of Trading Strategies*. New York: McGraw-Hill, 2000.

Katz, Jeffrey Owen, and Donna L. McCormick. *How to Start Day Trading Futures, Options, and Indices*. New York: McGraw-Hill, 2001a.

Katz, Jeffrey Owen, and Donna L. McCormick. "Market Realities and Options Pricing," *Futures*. Vol. 30, No. 6, May, 2001b, pp. 38–40.

Katz, Jeffrey Owen, and Donna L. McCormick. "More Intelligent Option Pricing," *Futures*. Vol. 30, No. 7, June, 2001c, pp. 42–45.

Katz, Jeffrey Owen, and Donna L. McCormick. "Trading Options with Finesse," *Futures*. Vol. 31, No. 3, February, 2002, pp. 42–46.

Kline, Paul. *Handbook of Psychological Testing*. London: Routledge, 2000.

Knight, John, and Stephen Satchell (eds.). *Forecasting Volatility in the Financial Markets,* 2nd ed. Oxford, UK: Butterworth-Heinemann, 2002.

Lamperti, John. *Probability*. New York: W.A. Benjamin, Inc., 1966.

Lederman, Jess, and Robert A. Klein (eds.). *Virtual Trading*. Chicago, IL: Probus Publishing, 1995.

Malkiel, Burton G. *A Random Walk Down Wall Street*, 4th ed. New York: W.W. Norton, 1985.

Masters, Timothy. *Neural, Novel & Hybrid Algorithms for Time Series Prediction*. New York: John Wiley and Sons, 1995.

McMillan, Lawrence G. *Options as a Strategic Investment*, 2nd ed. New York: New York Institute of Finance, 1993.

McMillan, Lawrence G. *McMillan on Options*. New York: John Wiley and Sons, 1996.

Merton, Robert C. *Continuous-Time Finance*. Cambridge, MA: B. Blackwell, 1990.

Minsky, Marvin, and Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press, 1969.

Montgomery, Douglas C., and Elizabeth A. Peck. *Introduction to Linear Regression Analysis*. New York: John Wiley and Sons, 1982.

Mooney, Christopher Z. *Monte Carlo Simulation*. Thousand Oaks, CA: Sage Publications, 1997.

Myers, Raymond H. *Classical and Modern Regression with Applications*. Boston, MA: Duxbury Press, 1986.

Natenberg, Sheldon. *Option Volatility and Pricing: Advanced Trading Strategies and Techniques*. Chicago, IL: Probus Publishing, 1994.

Nunnally, Jum C. *Psychometric Theory*, 2nd ed. New York: McGraw-Hill, 1978.

Press, William H., Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd ed. Cambridge, England: Cambridge University Press, 1992.

Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, England: Cambridge University Press, 1992.

Rossi, Peter E. (ed.). *Modelling Stock Market Volatility*. San Diego, CA: Academic Press, 1996.

Rozeboom, William W. *Foundations of the Theory of Prediction*. Homewood, IL: Dorsey Press, 1966.

Schiller, Jon. *The 100% Return Options Trading Strategy*. Brightwaters, NY: Windsor Books, 1998.

Shipley, Bill. *Cause and Correlation in Biology: A User's Guide to Path Analysis. Structural Equations, and Causal Inference*. Cambridge, England: Cambridge University Press, 2002.

Simons, Howard. *The Dynamic Option Selection System: Analyzing Markets and Managing Risk*. New York: John Wiley and Sons, 1999.

The Options Institute of the C.B.O.E. (ed.). *Options: Essential Concepts and Trading Strategies*, 3rd ed. New York: McGraw-Hill, 1999.

Vaga, Tonis. *Profiting from Chaos*. New York: McGraw-Hill, 1994.

Wasserman, Philip D. *Advanced Methods in Neural Computing*. New York: Van Nostrand Reinhold, 1993.

Wright, Sewall. "The Method of Path Coefficients," *Annals of Mathematical Statistics*. Vol. 5, 1934, pp. 161–215.

Zhang, Q.J., and K.C. Gupta. *Neural Networks for RF and Microwave Design*. Norwood, MA: Artech House, Inc., 2000.

# INDEX





Copyright © 2005 by Scientific Consultant Services, Inc. Click here for terms of use.

# ABOUT THE AUTHORS

***Jeffrey Owen Katz, Ph.D.,*** started out as a child prodigy in electronic engineering and rapidly advanced through biology, psychology, psychophysiology, physics, and computer science. As a graduate student in mathematics, he discovered and published a new method of factor rotation and also wrote a portion of the Numerical Taxonomy Systems Package. His multidisciplinary expertise has afforded him a unique perspective that has resulted in innovative contributions to a variety of fields. For the past 20 years, his primary interest has been in modeling and trading the markets, as well as applying artificial intelligence technology to that endeavor through his company, Scientific Consultant Services, Inc. Dr. Katz coauthored (with McCormick) *The Encyclopedia of Trading Strategies* (2000), *How to Start Day Trading Futures, Options, and Indices* (2001), and has contributed to a number of anthologies, scientific journals, and trade publications, including dozens of articles in *Technical Analysis of Stocks and Commodities* and *Futures Magazine*. In addition to providing private seminars and tutorials in the art of trading and system development, he has taught at the New York Institute of Finance, as well as at universities in both the United States and United Kingdom.

***Donna L. McCormick*** has been the vice president of Scientific Consultant Services, Inc., since 1989. Her background is in experimental psychology and psychophysics. For over 15 years, she was associated with the American Society for Psychical Research (est. 1885), where she worked as a researcher, educator, editor, and Administrative Director. She has lectured and published on parapsychology, was a contributing writer for *Technical Analysis of Stocks and Commodities*, and coauthored numerous works with Katz. She is currently involved in the study of chronic fatigue syndrome and fibromyalgia.

Readers are invited to contact the authors at: katz@scientific-consulants.com and mccormick@scientific-consultants. com.