1991

Implied Volatility and Risk Preference From Option Prices.

Kenneth Steven Bartunek

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

https://digitalcommons.lsu.edu/gradschool_disstheses/5166

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Implied volatility and risk preference from option prices

Bartunek, Kenneth Steven, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1991
Implied Volatility and Risk Preference from Option Prices

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in the
Interdepartmental Program in Business Administration

by
Kenneth S. Bartunek
B.S., University of Michigan-Flint, 1987
August 1991
ACKNOWLEDGEMENTS

I am eternally grateful to Dr. Chowdhury Mustafa, my committee chairman, whose wisdom and patience were essential to the completion of this dissertation. I also wish to thank Dr. G. Geoffrey Booth, Dr. Wai-kin Leung, Dr. Robert Martin, Dr. Gary C. Sanger, and Dr. Kwei Tang who served on my committee for their time and very helpful comments.

In addition to the guidance I received from the aforementioned, I would also like to thank my undergraduate mentors: Dr. Steven Althoen, Dr. John Kling, and Dr. Marc Nyden. I have been blessed with many fine professors who have cultivated my love of learning.

Finally, I would like to thank those without whom none of this would have been possible. I thank my parents, Steven and Frances Bartunek, and my brother, Dr. James Bartunek, for their love and support. I am also indebted to them for developing in me a love of learning. I thank God for everything, including my abilities and the opportunity to use them.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Literature Review: The Path to the Equivalent Martingale Measure</td>
<td>4</td>
</tr>
<tr>
<td>3. Estimation of Risk Aversion</td>
<td>13</td>
</tr>
<tr>
<td>A. Model Derivation</td>
<td>14</td>
</tr>
<tr>
<td>B. Simulation</td>
<td>19</td>
</tr>
<tr>
<td>C. Criterion Function</td>
<td>20</td>
</tr>
<tr>
<td>D. Data</td>
<td>25</td>
</tr>
<tr>
<td>E. Results</td>
<td>26</td>
</tr>
<tr>
<td>F. Conclusion</td>
<td>30</td>
</tr>
<tr>
<td>4. Literature Review for Estimation of Volatility</td>
<td>42</td>
</tr>
<tr>
<td>A. Historical Volatility</td>
<td>42</td>
</tr>
<tr>
<td>B. Black-Scholes Implied Volatility</td>
<td>43</td>
</tr>
<tr>
<td>C. GARCH Volatility</td>
<td>55</td>
</tr>
<tr>
<td>5. Estimation of Volatility</td>
<td>67</td>
</tr>
<tr>
<td>A. Data</td>
<td>70</td>
</tr>
<tr>
<td>B. Methodology</td>
<td></td>
</tr>
<tr>
<td>1. Ex-post Volatility Estimation</td>
<td>71</td>
</tr>
<tr>
<td>2. Implied Volatility from the Black-Scholes Formula</td>
<td>72</td>
</tr>
<tr>
<td>3. GARCH Model Estimation</td>
<td>78</td>
</tr>
<tr>
<td>4. Estimation of Historical Volatility</td>
<td>80</td>
</tr>
<tr>
<td>5. Forecast Comparisons</td>
<td>81</td>
</tr>
<tr>
<td>C. Results</td>
<td>82</td>
</tr>
<tr>
<td>D. Conclusion</td>
<td>87</td>
</tr>
<tr>
<td>6. Summary</td>
<td>103</td>
</tr>
<tr>
<td>References</td>
<td>105</td>
</tr>
<tr>
<td>Vita</td>
<td>111</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Implying $\alpha$ Using the Generalized Method of Moments with the Variance-Covariance Matrix Used as a Weighting Matrix</td>
<td>32</td>
</tr>
<tr>
<td>3-2</td>
<td>Implying $\alpha$ Using the Generalized Method of Moments with the Variance-Covariance Matrix Corrected for Autocovariance Used as a Weighting Matrix</td>
<td>33</td>
</tr>
<tr>
<td>3-3</td>
<td>Variance-Covariance Matrices for Each Set of Three Orthogonality Conditions ($\alpha$)</td>
<td>34</td>
</tr>
<tr>
<td>3-4</td>
<td>Variance-Covariance Matrices for Each Combination of Two Sets of Three Orthogonality Conditions ($\alpha$)</td>
<td>35</td>
</tr>
<tr>
<td>3-5</td>
<td>Variance-Covariance Matrix for Three Sets of Three Orthogonality Conditions ($\alpha$)</td>
<td>36</td>
</tr>
<tr>
<td>3-6</td>
<td>Implying $\alpha$ and $\alpha$ Using the Generalized Method of Moments with the Variance-Covariance Matrix Used as a Weighting Matrix</td>
<td>37</td>
</tr>
<tr>
<td>3-7</td>
<td>Implying $\alpha$ and $\alpha$ Using the Generalized Method of Moments with the Variance-Covariance Matrix Corrected for Autocovariance Used as a Weighting Matrix</td>
<td>38</td>
</tr>
<tr>
<td>3-8</td>
<td>Variance-Covariance Matrices for Each Set of Three Orthogonality Conditions ($\alpha$ and $\sigma$)</td>
<td>39</td>
</tr>
<tr>
<td>3-9</td>
<td>Variance-Covariance Matrices for Each Combination of Two Sets of Three Orthogonality Conditions ($\alpha$ and $\sigma$)</td>
<td>40</td>
</tr>
<tr>
<td>3-10</td>
<td>Variance-Covariance Matrix for Three Sets of Three Orthogonality Conditions ($\alpha$ and $\sigma$)</td>
<td>41</td>
</tr>
<tr>
<td>5-1</td>
<td>GARCH Parameters Averaged over the 506 Estimated Models</td>
<td>90</td>
</tr>
<tr>
<td>5-2</td>
<td>Regression Results for Daily Data - 20 Day Forecast Horizon</td>
<td>91</td>
</tr>
<tr>
<td>5-3</td>
<td>Regression Results for Daily Data - 40 Day Forecast Horizon</td>
<td>92</td>
</tr>
<tr>
<td>5-4</td>
<td>Regression Results for Daily Data - 80 Day Forecast Horizon</td>
<td>93</td>
</tr>
<tr>
<td>5-5</td>
<td>Regression Results for Daily Data - 120 Day Forecast Horizon</td>
<td>94</td>
</tr>
<tr>
<td>5-6</td>
<td>Mean Absolute Deviation Results for Daily Data - 20 Day Forecast Horizon</td>
<td>95</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>5-7</td>
<td>Mean Square Error Results for Daily Data - 20 Day Forecast Horizon</td>
<td>96</td>
</tr>
<tr>
<td>5-8</td>
<td>Mean Absolute Deviation Results for Daily Data - 40 Day Forecast Horizon</td>
<td>97</td>
</tr>
<tr>
<td>5-9</td>
<td>Mean Square Error Results for Daily Data - 40 Day Forecast Horizon</td>
<td>98</td>
</tr>
<tr>
<td>5-10</td>
<td>Mean Absolute Deviation Results for Daily Data - 80 Day Forecast Horizon</td>
<td>99</td>
</tr>
<tr>
<td>5-11</td>
<td>Mean Square Error Results for Daily Data - 80 Day Forecast Horizon</td>
<td>100</td>
</tr>
<tr>
<td>5-12</td>
<td>Mean Absolute Deviation Results for Daily Data - 120 Day Forecast Horizon</td>
<td>101</td>
</tr>
<tr>
<td>5-13</td>
<td>Mean Square Error Results for Daily Data - 120 Day Forecast Horizon</td>
<td>102</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-1</td>
<td>89</td>
</tr>
</tbody>
</table>

Sample Calculations of Standard Deviations
ABSTRACT

This dissertation consists of two empirical studies on options. In the first study, an estimate of the constant proportional risk aversion parameter is implied from the equivalent martingale measure framework. Within this framework, we use call options as opposed to the traditional consumption data to imply our estimate. We compare forecasts of volatility for the asset underlying an option in the second study. Three methods have been used to forecast volatility in the past. These are an historical estimate, an estimate implied from the Black-Scholes European call option pricing model, and an estimate based on generalized autoregressive conditional heteroscedasticity. The first is an unconditional estimate whereas the latter two are conditional. Previous literature has compared the forecasting ability of the unconditional estimate with one of the conditional estimates. We focus on the comparison of the two conditional volatility estimates in our second study, in addition to the unconditional versus conditional comparisons.
INTRODUCTION

This dissertation is composed of two empirical studies in the field of options. The first is motivated by the models in financial economics, such as the Lucas (1982) model, which require a risk aversion parameter. A number of studies have estimated the constant proportional risk aversion parameter, including Friend and Blume (1975), Kydland and Prescott (1982), and Hansen and Singleton (1982), using consumption data. However, there are many problems associated with measuring this consumption data as discussed in Breeden, Gibbons, and Litzenberger (1989). In our first study, we estimate the risk aversion parameter, assuming constant proportional risk aversion and using options data from the Chicago Board Options Exchange, in a valuation framework developed from the equivalent martingale measure of Harrison and Kreps (1979). Hence, we avoid the problems associated with the measurement of the consumption data.

The second study focuses on forecasts of the volatility of the returns on an underlying asset on which an option is written. With the advent of the Black and Scholes (1973) European call option pricing model, estimation of the volatility of the underlying asset has become a very popular research topic. This occurs because the model is a function of five variables, two of which are contracted and two of which are observable in the economy, leaving only the fifth, the volatility of the underlying asset, to be estimated. Three forecasts have been previously applied. Black and Scholes (1972) forecast the volatility using the past returns volatility. This is an unconditional volatility estimate. Latane and Rendleman (1976) propose implying the volatility from the Black and
Scholes formula itself. As this method uses option prices to imply the volatility, it yields a conditional volatility estimate because these option prices are based on the information in the market.¹ More recently the volatility of stock returns has been estimated using autoregressive conditional heteroscedasticity, as developed in Engle (1982), and some modifications to produce conditional volatility estimates based on historical returns.

A number of studies conducted by Latane and Rendleman (1976), Chiras and Manaster (1978), and Beckers (1981) find the Black-Scholes implied volatility is a better forecast of volatility than is historical volatility. Ak giray (1989) finds that generalized autoregressive conditional heteroscedasticity yields a better forecast of volatility than historical volatility. All of the previous studies compare a conditional volatility forecast with an unconditional forecast. In our second study, we compare the two conditional volatility forecasts, namely, the Black-Scholes implied volatility and the generalized autoregressive conditional heteroscedasticity volatility.

Chapter 2 is a brief literature review of option pricing which traces the path of the development of the equivalent martingale measure. In Chapter 3 we derive our valuation equation from the equivalent martingale measure framework, describe our estimation procedure and estimate the risk aversion parameter. Chapter 4 is a literature review for our second study which focuses on two areas. The first is the estimation of the volatility of stock returns. The second is the evidence of a

¹ This statement is based on the assumption that some form of market efficiency holds within the options market.
changing volatility. The estimation and comparison of the volatility estimates are discussed in Chapter 5.
CHAPTER 2
LITERATURE REVIEW: THE PATH TO THE EQUIVALENT MARTINGALE MEASURE

Here we provide an introduction to options and an overview of the development of option pricing theory. After defining an option and its relevant characteristics, we briefly discuss option theory previous to Black and Scholes (1973). Then we introduce the Black and Scholes European call option valuation model and its contribution to option pricing theory. Finally, we trace the development of the risk neutral valuation relationship (RNVR) to the equivalent martingale measure (EMM) developed in Harrison and Kreps (1979).

An American call (put) option is a contract giving its owner the right to buy (sell) a specified asset at a specified price through a specified date. An European option is the same except it may only be exercised on the specified date. The asset, price and date specifications are set in the contract. The specific date is termed the maturity or expiration date. The specified asset is referred to as the underlying asset. The specified price at which the call option holder may buy the underlying asset is termed the exercise or strike price. Options are often referred to under the more general headings of derivative securities or contingent claims as their value is related to the value of this underlying asset. Much work has been done trying to find a valuation formula for options of which the most prominent is Black and Scholes (1973).

Previous to Black and Scholes (BS), there is a series of papers in the mid-1960's dealing with option valuation. Sprenkle (1964) proposes an
investor will only pay the expected terminal value of the option for the option today if he is risk neutral and the interest rate is zero. Boness (1964) and Samuelson (1965) develop models taking into account a non-zero interest rate allowing for the time value of money, but do not find satisfactory explanations of what this discount factor should be. Then Black and Scholes (1973) suggest the formation of a riskless hedge from a portfolio of the underlying asset and some European call options written on this asset. They argue, that at any given moment, portfolio value changes must result from changes in asset values as the amounts of assets held do not change at a given moment. Defining the call price to be a function of the stock price and time to maturity, changes in the call price are a function of changes in these two quantities. However, at any given moment changes in the call price are a function of changes in the stock price only. Therefore, they suggest that by selecting an appropriate number of calls to hedge relative to a particular position in the stock, the portfolio can be made riskless as a change in value of one asset one way can be offset by a change in value of the other asset in the opposite direction. With the stock and option positions appropriately continuously adjusted, the return to this portfolio will be riskless. Hence, they have found a satisfactory explanation for the discount factor to be the riskless rate.

Black and Scholes (1973) make the following assumptions in their model: 1) The riskless interest rate is known and constant. 2) The stock price, \( S \), follows a stochastic process, \( \frac{dS}{S} = \mu dt + \sigma dz \), where \( \mu \) is the instantaneous expected return, \( \sigma \) is the instantaneous standard deviation of return, \( dt \) is an infinitesimal change in time, and \( dz \) is a Weiner
process, a standard unit normally distributed variable. 3) The stock pays no dividends. 4) The option can only be exercised at maturity. 5) There are no taxes or transactions costs. 6) There are no penalties for short sales. 7) The market operates continuously. The Black-Scholes European call option valuation formula is

\[
C = S N(d_1) - \exp(-rt) X N(d_2),
\]

\[
d_1 = \frac{\ln(S/X) + (r + \sigma^2/2)T}{\sigma \sqrt{T}},
\]

\[
d_2 = d_1 - \sigma \sqrt{T}, \quad \text{where}
\]

- C = European call option price,
- S = stock price,
- X = exercise price,
- r = riskless interest rate,
- T = time to maturity,
- \sigma = standard deviation of the rate of return on underlying asset S, and
- N = normal density.

This model is a significant advance as the rate of discount causing the problem in previous literature is now known and it is the riskless rate. This model is also very appealing in the sense that although the call price is a function of five variables, S, X, r, \sigma, and T, two are set in the option contract itself, X and T, and two others can be observed in the economy, S and r. Thus, only the variance of the rate of return on the underlying asset is left to estimate.

A series of papers deal with relaxing the assumptions in the BS model and many are mentioned in Smith's (1976) review of option pricing. Merton (1973) allows a stochastic interest rate, Thorpe (1973) allows for short sales constraints, and Ingersoll (1976) considers taxes. Merton
(1973) also shows that an American call on a non-dividend paying stock is equivalent to an European call and he develops an European option valuation model for a constant continuous dividend yield. Black (1975) tries to account for early exercise potential of an American option on a stock with known dividends, but only with a probability of zero or one. Roll (1977) and Geske (1979) develop a model that allows probabilities between zero and one. Whaley (1981) corrects errors in the two papers to find an option valuation formula for a stock with known dividends. The continuousness of the stochastic process is relaxed in Merton (1976) and Cox and Ross (1976) where they allow a jump in the process. Cox and Ross (1976) consider a number of alternative specifications to the stochastic process, both jump and diffusion models, from which they conclude the specification of the stochastic process for the underlying stock is crucial to the option valuation obtained. In the original Black and Scholes (1973) paper, it is assumed that the variance of the rate of return is constant. Merton (1973) shows that BS can be extended to allow the variance to change as a deterministic function of time. Cox (1973) develops the constant elasticity of variance model where the variance changes stochastically, but as a function of the stock price. Then a series of papers allow the variance to follow an independent stochastic process including Johnson and Shanno (1987), Hull and White (1987), and Scott (1987).

Cox and Ross (1976) argue that as no risk preference assumptions are made to derive the partial differential equation that yields a solution for BS, any risk preference assumption can be made to yield a solution that holds for any and all risk preferences. Since risk neutrality assigns
the riskless rate as the equilibrium rate of return to all assets, this is often chosen for simplicity. This relationship in which something can be valued as if investors are risk neutral regardless of actual risk preferences has come to be known as a risk neutral valuation relationship, RNVR. It obtains in continuous time when a riskless hedge can be formed and adjusted continuously. As the assumption of continuous trading is relaxed, a problem arises because with discrete trading the hedge cannot be adjusted continuously. A RNVR does not generally obtain with discrete trading. However, a RNVR does obtain with discrete trading under appropriate conditions.

The first discrete time model to be discussed under which a RNVR obtains is the binomial option pricing model of Cox, Ross, and Rubinstein (1979). For their model they assume the underlying stock $S$ follows a binomial process where it either ends up in one state with a probability $q$, where its value is $uS$, or ends down in the other state with a probability $1-q$, where its value is $dS$. The rates of return to the stock are $u-1$ and $d-1$, in their respective states, and $r$ which is one plus the riskless rate must be less than $u$ and greater than $d$. This requirement eliminates riskless arbitrage opportunities. They also assume the interest rate is constant, there are no limits on borrowing or lending at this rate, and there are no taxes, transactions costs or margin requirements. In their derivation they redefine the probability measure, which was $q$ and $1-q$, to

$$p = \frac{r-d}{u-d} \quad \text{and} \quad 1-p = \frac{u-r}{u-d}$$

so that the call may be valued as

$$C = \left[ pCu + (1-p)Cd \right]/r,$$
where \( C_u = \max[0,uS-X] \), or the call value if the realized state is \( u \), and 
\( C_d = \max[0,dS-X] \), or the call value if the realized state is \( d \). They show 
p is the value \( q \) would assume in equilibrium with risk neutrality. For 
risk neutrality to obtain the following equation must be satisfied as the 
expected return for all assets must be the riskless rate
\[
quS + (1-q)dS = rS.
\]
Solving for \( q \) one obtains
\[
q = \frac{(r-d)}{(u-d)}.
\]

Cox, Ross, and Rubinstein's (1979) model is derived for a two asset 
case to span the binomial process of the underlying asset in discrete time 
so that the market is complete. If the market is not complete in discrete 
time, it is necessary to make assumptions about investor preferences in 
addition to non-satiation to obtain a RNVR. Rubinstein (1976) develops a 
model in discrete time for the many asset case. Assuming a representative 
investor, a bivariate lognormal distribution of asset returns and 
aggregate wealth, and constant proportional risk aversion (CPRA), he finds 
with discrete trading that as the trading interval approaches zero the BS 
formula still holds. This is a result of the fact that the bivariate 
lognormality of the distribution in conjunction with CPRA yields a RNVR as 
shown explicitly in Brennan (1979). Brennan also shows that with the 
assumptions of a representative investor, a bivariate normal distribution 
of asset returns and aggregate wealth, and constant absolute risk aversion 
a RNVR is again achieved in discrete time.

Prompted by the arbitrage arguments used in option pricing theory, 
Harrison and Kreps (1979), working with consumption bundles, of which a 
derivative security is an example, derive the concept of an equivalent
martingale measure, EMM. They assume a discrete framework of prespecified trading dates and no arbitrage possibilities. They also require one security to be a riskless discount bond so that a normalized price system can be created by dividing all security prices at each trading date by the price of the riskless bond at these same trading dates, respectively. Then they prove that under this new price regime, if there are to be no arbitrage possibilities, a new probability measure with the martingale property must exist that is equivalent to the old. That is, if a state has non-zero probability in the old measure, it also has non-zero probability in the new measure and the price processes of the securities are martingales under this new probability measure. A price process is a martingale if its expected value given the information set today is its price today or in equation form

\[ E[S_{t+i} | I_t] = S_t, \quad \text{for all } i \geq 0, \text{ where} \]

\[ S_t = \text{security price at time } t, \]
\[ I_t = \text{information set at time } t, \text{ and} \]
\[ i = \text{a non-negative integer.} \]

This new probability measure is termed an equivalent martingale measure.

Huang and Litzenberger (1985) maximize the expected value of the sum of utilities of consumption subject to a budget constraint. That is, the consumption plan is financed by a trading strategy in which the amount reinvested in securities in each period equals the amount received from last period's investments in securities less the amount consumed. Assuming no arbitrage and normalizing the price process they show the equivalent martingale measure is the following transformation of the probability in each state:
\[
\pi^* = \left[ \frac{U'(K_T)}{U'(K_0)} \right] \left[ \frac{\pi}{B_0} \right], \text{ where}
\]

\[
\pi^* = \text{new probability},
\]

\[
\pi = \text{old probability},
\]

\[
B_0 = \text{value of riskless security at time } 0,
\]

\[
U'(K_T) = \text{marginal utility of consumption plan } K \text{ at time } T,
\]

\[
U'(K_0) = \text{marginal utility of consumption plan } K \text{ at time } 0.
\]

This equation will be referred to as the operational definition of the EMM.

With the introduction of the EMM, we can show Cox, Ross, and Rubinstein's (1979) derivation of the binomial option pricing formula is equivalent to the use of the equivalent martingale measure process. Rewriting the following equation

\[
C = \frac{pC_u + (1-p)C_d}{r}
\]

to obtain

\[
C = p\left( \frac{C_u}{r} \right) + (1-p)\left( \frac{C_d}{r} \right)
\]

the properties of the EMM are straightforward. First, the prices are normalized as \(C\) at time 0 can be thought of as divided by one, one plus the riskless rate of return at time 0, and \(C_u = \max[0,uS-X]\) and \(C_d = \max[0,dS-X]\) at time 1 are divided by one plus the riskless return at time 1, \(r\). Second, the expected value at time 1 for the normalized call value represented by the right hand side equals the normalized call value at time 0 represented by the left hand side, so that the process is a martingale. Third, observe that for this binomial process there are originally two non-zero probabilities \(q\) and \(1-q\) as shown earlier and there are now two non-zero probabilities so that the two measures are equivalent. Hence, these two new probabilities, \(p\) and \(1-p\), satisfy the
properties of an EMM. As we shall see in the next chapter, it is this EMM concept that facilitates the estimation of a risk aversion parameter from a contingent claim framework.
CHAPTER 3
ESTIMATION OF RISK AVERSION

There have been a number of studies done in which an estimate of the constant proportional risk aversion parameter is found. One such study conducted by Friend and Blume (1975) using data on household assets and liabilities produces an estimate near two. Kydland and Prescott (1982) use data on aggregate fluctuations of real output and other aggregate economic time series and find a constant proportional risk aversion parameter estimate in the range of one to two. In Hansen and Singleton (1982), estimates are found between zero and two using value weighted New York Stock Exchange and Treasury Bill monthly returns along with a corresponding consumption series. Mehra and Prescott (1985) find in their study that when the risk aversion parameter is varied within a "reasonable" range, between zero and ten, they are unable to explain the wide variation between theoretical and empirical excess returns yielding an "equity premium puzzle". More recently Kandel and Stambaugh (1990, 1991) argue against restricting the range of the risk aversion parameter and use values of 29 and 55.

All of these studies find their estimate of the risk aversion parameter using consumption data. This study proposes a new method to estimate the constant proportional risk aversion parameter. In this study, we use the equivalent martingale measure framework to develop a model from which we estimate the risk aversion parameter using call options data.
A. MODEL DERIVATION

Prior to Black and Scholes (1973) the primary problem with option pricing was the determination of the appropriate discount rate. Black and Scholes (BS) argue that with continuous trading and a position in a stock with an offsetting position in calls one can form a riskless hedge. Hence, solving the discounting problem, as the return to a riskless hedge should be the riskless rate in an equilibrium environment. Cox and Ross (1976) argue that as the option value does not depend directly on the risk preferences of investors one can make any assumption of risk preference to derive the valuation and it will produce the same valuation as any other risk preference assumption. In this environment in which the valuation is independent of risk preferences, it is often advantageous to assume risk neutrality. A situation in which the valuation is independent of risk preferences and one assumes investors act as if they are risk neutral is defined to be a risk neutral valuation relationship (RNVR).

A number of studies have dealt with the environments in which an RNVR obtains. Rubinstein (1976), in a discrete time framework with many assets, finds a RNVR obtains given a representative investor, a bivariate lognormal distribution of asset returns and aggregate wealth, and constant proportional risk aversion (CPRA). Brennan (1979) also finds given a representative investor, a bivariate normal distribution of asset returns and aggregate wealth, and constant absolute risk aversion (CARA), a RNVR obtains. In a discrete time framework, two asset, two state case, Cox, Ross, and Rubinstein (1979) find a RNVR obtains as well. Finally, Harrison and Kreps (1979), generalizing the above results, show that for every risk
preference assumption there is a corresponding distribution such that an RNVR obtains. They develop the concept of an equivalent martingale measure (EMM) which simply redistributes the original probability mass under some risk preference assumption so that the asset’s expected rate of return with this new distribution is the riskless rate. Thus, application of the EMM yields a RNVR, as regardless of their individual risk preference each investor’s expected rate of return under an EMM is the riskless rate.

It is within this EMM framework that we shall estimate a risk aversion parameter. The EMM is developed in a discrete time framework of prespecified trading dates and no arbitrage. Harrison and Kreps (1979) assume security prices are in units of the consumption good, as only relative prices are determined in equilibrium, so the spot price of consumption is one in all periods. They require one security to be a riskless bond with interest rate zero, which Huang and Litzenberger (1985) generalize to simply a riskless bond, so that a normalized price system can be created by dividing all security prices at each trading date by the corresponding price of the riskless bond at that trading date. An EMM is a new probability measure such that any state with non-zero probability in the old measure also has non-zero probability in the new measure and the price processes of the securities are martingales under the new measure. We will use an operational definition of the EMM as derived in Huang and Litzenberger’s (1985) discussion of the valuation of contingent claims by arbitrage where each probability is transformed as follows:

\[ p_i' = \frac{p_i}{B_{t+1}} \]

2 Whether this relation is one-to-one depends on market completeness. If the market is complete the relation is one-to-one.
\[ \pi^* = \left[ \frac{U'(K_T)}{U'(K_0)} \right] \frac{\pi}{B_0}, \text{ with} \]

\[ \pi^* = \text{new probability}, \]

\[ \pi = \text{old probability}, \]

\[ B_0 = \text{value of the riskless security at time 0}, \]

\[ U'(K_T) = \text{marginal utility of consumption plan } K \text{ at time } T, \text{ and} \]

\[ U'(K_0) = \text{marginal utility of consumption plan } K \text{ at time } 0. \]

They also show that under the normalized price system, consumption plans, such as \( K_0 \), have the martingale property. With this result

\[ \left( \frac{K_0}{B_0} \right) = E^* \left[ \frac{K_T}{B_T} \right], \text{ where} \]

\[ E^* = \text{expectation under EMM, and} \]

\[ B_T = \text{value of riskless security at time } T. \]

Time \( T \) is the final trading date and the riskless security is defined to have a value of one at this time \( (B_T = 1) \). ³ Substituting the operational definition of the EMM one obtains

\[ \left( \frac{K_0}{B_0} \right) = E \left[ \frac{U'(K_T)}{U'(K_0)} \right] \frac{1}{B_0} \left( \frac{K_T}{B_T} \right). \]

Using the fact that \( B_T = 1 \), the expression becomes

\[ \left( \frac{K_0}{B_0} \right) = E \left[ \frac{U'(K_T)}{U'(K_0)} \right] \left( \frac{K_T}{B_0} \right). \]

Recognizing that \( U'(K_0) \) is known at time 0 it can be taken outside of the expectations operator and moved to the other side of the equation to yield

\[ K_0 U'(K_0) = E[K_T U'(K_T)]. \]

Then specifying utility as constant proportional risk aversion

\[ U(K) = K^{1-\alpha}/(1-\alpha) \text{ with } 0 < \alpha < \infty. \]

Differentiating \( U(K) \) with respect to \( K \) we obtain the corresponding marginal utility function which is

\[ U'(K) = K^{-\alpha} \]

³ The riskless security is just a riskless discount bond.
and substituting in this result yields

\[ K_{0}^{1-\alpha} = E[K_{T}^{1-\alpha}]. \]

In applying the theoretical framework to our empirical study, we must reintroduce a discount factor as the assumption that the spot price of consumption does not change is too restrictive in reality. It is more realistic to let the spot price of consumption grow. As the appropriate discount factor for the EMM is the riskless rate, we assume the spot price of consumption grows at this same rate. Hence, we can adjust time T consumption units to time zero consumption units by simply discounting the former with the riskless rate from time zero to time T or equivalently by multiplying by \( B_{0} \). This yields

\[ K_{0}^{1-\alpha} = B_{0}E[K_{T}^{1-\alpha}]. \]

As discussed in Huang and Litzenberger (1985), a contingent claim is characterized by its payoffs in states through time. Hence, a contingent claim is just a consumption plan. As a call option is just a contingent claim, \( C_{S} \) and \( C_{T} \), representing the call values at time zero and time T respectively, can be substituted for \( K_{0} \) and \( K_{T} \), respectively. Then solving for \( C_{S} \) we obtain

\[ C_{S} = (B_{0}E[C_{T}^{1-\alpha}])^{(1/1-\alpha)}. \]

Using the value of a call at expiration, \( C_{T} = \max[0,S_{T}-X] \), derived by Merton (1973) with only dominance arguments

\[ C_{S} = (B_{0}E[(\max[0,S_{T}-X])^{1-\alpha}])^{(1/1-\alpha)}, \tag{1} \]

where \( S_{T} \) is the stock price at the expiration of the option and \( X \) is the exercise price.

---

\(^{4}\) As \( X^{1-\alpha} \) is concave, we encounter Jensen's inequality as we have the expectations operator outside of this function.
To imply the risk aversion parameter, $\alpha$, simulation is required to generate a number of values for the expression within the expectations operator so an average can be found to represent the expected value. There is a problem in Equation (1), however. As $0 < \alpha < \infty$, it can of course be greater than 1. The terminal call option value is occasionally zero when we simulate and this leads to an initial call value of zero with $\alpha > 1$, given the other simulated call values are finite. This is inappropriate given the call must have traded to be in the sample. We assign a value of zero to the expression

$$(\max[0,S_T-X])^{1-\alpha}$$

when $\alpha > 1$ and $S_T < X$. This assigns a value of zero to marginal utility when the call price is zero.$^5$ To estimate $\alpha$ we shall minimize our criterion function, discussed below, which is based on the difference between the observed call values and those generated from our valuation. For our minimization, we shall use a guess of $\alpha$ to find a value for $C_S$. Then we check to see if the simulation call values are close enough to the observed call values with our criterion function. If not, we update our $\alpha$ estimate and repeat the procedure. Hence, the only unknown value we need to know before we start the optimization process is $S_T$. As we are at time zero and valuing the stock at time $T$, we shall use simulation to obtain $S_T$.

---

$^5$ We also try another approach to bound the consumption plan away from zero. We add a small positive constant, $p$, to the terminal call option valuation in Equation (1). However, to keep the equation balanced we need to add an amount to $C_S$ which would be equivalent to $p$ at $C_T$. We add $B_0p$ which when subtracted from both sides of the equation yields

$$C_S = (B_0E[\max[0,S_T-X]+p]^{1-\alpha})^{1/(1-\alpha)} - B_0p.$$ 

We use a value for $p$ of .0001. We find the same $\alpha$ parameter for a range of $p$ values. We find the same $\alpha$ parameter values under the method described here in the footnote and above in the text.
B. SIMULATION

To evaluate the expression within parentheses in Equation (1) we need to obtain $S_T$ and this is done using simulation. For our simulation we follow a method similar to that in Boyle (1977). Given the stock price today, $S_t$, we can generate the stock price at expiration, $S_T$, by assuming a specific return generation process or a specific distribution for stock returns. Boyle assumes risk neutrality. Our approach is the same except that we are interested in estimating the risk aversion parameter of the underlying investors which is embedded in the underlying asset returns. Hence, we need to obtain an estimate of the mean return, $\mu_r$, which we do by taking a simple average of the returns on the underlying asset over the three years of our sample period. Then we simply use this return in place of the riskless rate in the methodology of Boyle (1977).

The equilibrium expected return expression is

$$E(S_T/S_t) = \exp(\mu_r [T-t])$$

where $\mu_r$ is the mean daily return of the underlying asset over the three year period of our study. Then we assume $S_T/S_t$ follows a lognormal distribution with a mean of $\exp(\mu_r (T-t))$. This implies that the mean of the normal distribution of $\log(S_T/S_t)$ with a variance of $\sigma^2(T-t)$ is $(\mu_r - .5\sigma^2)(T-t)$. This results from the fact that the mean of a lognormal variable, $\exp(\mu_r (T-t))$, is equal to the exponential of the sum of the mean and one-

---

6 Another potential method of obtaining an estimate of the mean return of the underlying stock is with the capital asset pricing model ($E[R_j] = R_f + \alpha \text{cov}(R_j, R_m)$). Given estimates of the other parameters one can generate the return on the underlying stock, $E[R_j]$, from the model.
half the variance, \( \mu(T-t) \) and \( \sigma^2(T-t) \) respectively, of the corresponding normal distribution. That is
\[
\exp(\mu_r(T-t)) = \exp((\mu+.5\sigma^2)(T-t)).
\]
Solving for \( \mu \) we find it equals \( \mu_r-.5\sigma^2 \). Thus,
\[
E[\log(S_T/S_t)] = (\mu_r-.5\sigma^2)(T-t).
\]

To generate \( S_T \) we can use the following expression
\[
S_T = S_t \exp[\mu_r(T-t)-.5\sigma^2(T-t)+\sigma x(T-t)^{1/2}]
\]
where \( x \) is a standard normally distributed random variable. We run two series of tests, one in which we estimate \( \sigma^2 \) along with \( \mu_r \) from historical data and imply only \( \alpha \) from our valuation in Equation (1) and a second where we estimate \( \mu_r \) from historical data and imply both \( \alpha \) and \( \sigma^2 \).

C. CRITERION FUNCTION

Using the estimate of \( \mu_r \) from the historical data and an estimate of \( \alpha \) and \( \sigma^2 \), either an initial guess, if on the first pass, or an updated estimate from an optimization algorithm, we can generate \( S_T \) using Equation (2). Repeating the process of generating \( S_T \) a large number of times, 100, we can obtain a simulated value for the expected value in Equation (1) by taking a simple average over these 100 iterations. Similar to Bossaerts and Hillion (1989), we find 100 iterations sufficient for our simulation. We check our simulation errors by comparing theoretical Black-Scholes option values with BS values generated from our valuation in Equation (1) where we set \( \mu_r \) to the riskless rate and \( \alpha \) to zero. For ten and fifty simulations the simulation errors are quite large, but for 100 simulations the errors are less than 3 tenths of one cent. We must also consider the
size of the simulation errors with respect to our model pricing errors, as
the latter need to exceed the former so our conclusions may be based on
the models performance as opposed to the simulation performance. We find
at 100 simulations that the model pricing errors are roughly a magnitude
of ten greater than the simulation errors. Now we can calculate a
simulation\(^7\) value of the call option price at time 0, \(C_s\), from Equation
(1) by substituting the simulated value for the expected value. To find an
estimate of \(\alpha\), or of \(\alpha\) and \(\sigma^2\), we need to specify a criterion function to
optimize.

Our criterion function, the method of simulated moments (MSM), is
discussed in Bossaerts and Hillion (1989), Duffie and Singleton (1989) and
Mcfadden (1989). This process involves the substitution of the simulation
values for the analytical values in the orthogonality conditions of the
generalized method of moments as developed theoretically in Hansen (1982)
and empirically in Hansen and Singleton (1982). The orthogonality
conditions are a set of conditions based on the economic theory which are
expected to be zero when the parameters to be estimated are at their
optimum values. In our model, we wish to imply a value of \(\alpha\) under the
assumption that given the optimum \(\alpha\) value our valuation in Equation (1)
provides the equilibrium call value which is equivalent to the observed
market value. Hence, one set of orthogonality conditions is

\[
E[e_{it}(\alpha)] = 0 \quad \text{for } i = 1 \text{ to } n,
\]

\(^7\) As in Bossaerts and Hillion (1989), we will refer to the estimate
of the call option as the simulation value and the estimate of the stock
price at expiration as the simulated value as the latter is the actual
simulated variable.
where \( n \) is the number of random series chosen and \( e_{it} = C_{0it} - C_{Sit} \). \( C_{0it} \) is the observed call value at time \( t \) and \( C_{Sit} \) is the simulated call value at time \( t \). Additional sets of orthogonality conditions are simply those of the cross-moments, where the cross-moments are formed by the product of the errors, \( e_{it} \), and the values of the corresponding instrumental variable \( z_{it} \).

\[
E[e_{it} z_{it}^j] = 0 \quad \text{for } i = 1 \text{ to } n \text{ and } j = 1 \text{ to } m,
\]

where \( j \) denotes the instrumental variable. The instrumental variables chosen are the strike price, \( X \), and the time-to-maturity, \( T \). To simplify notation we shall consider \( z_{it}^1 = 1 \) so we include the first set of orthogonality conditions within the general notation.

If we let \( h_{it} = e_{it} z_{it} \), then to find the sample counterpart to the population orthogonality conditions we simply take the mean of \( h_{it} \)

\[
h_T = ((1/T) \sum_{t=1}^{T} h_{1t}, \ldots, (1/T) \sum_{t=1}^{T} h_{nt}),
\]

where \( h_T \) is a vector of the \( i \) means of \( h_{it} \). The criterion function, \( f \), which we are to minimize, is the distance between the means of the observed and simulated call values weighted by a symmetric matrix \( W_T \),

\[
f = h_T' W_T h_T.
\]

There are a number of papers that discuss what the optimal \( W_T \) is in that it makes the asymptotic variance-covariance matrix of the parameter estimates as small as possible. As discussed in Hansen (1982), Hansen and Singleton (1982), and Newey and West (1987), this process involves two stages.

In the first stage one uses a suboptimal choice of \( W_T \). In our case, we use the inverse of the sample variance-covariance matrix as our weighting matrix. Hansen and Singleton (1982) point out that a consistent estimator of the parameters is needed to calculate the optimal weighting
matrix $W_T^*$ and Hansen (1982) shows the estimates from the first stage satisfy this criterion. Hence, in the second stage we use the estimate of the parameters resulting from the first stage. Now, we calculate a new minimum based on the criterion function

$$f = h_T'W_T^*h_T.$$ 

$W_T^*$ is now an autocovariance corrected sample variance-covariance matrix.

$$W_T^* = (S_T + \{A_k^1[A_k + A_k']\})^{-1},$$

where $S_T$ is the sample variance-covariance matrix and $A_k$ is the autocovariance matrix for the $l$th lag. The prime on the second $A_k$ represents the transpose. In our case, it simplifies, as we correct for first lag autocorrelation only, to

$$W_T^* = (S_T + A_1^1 + A_1')^{-1}.$$ 

Under our criterion function a number of studies, Hansen (1982), Cumby, Huizinga and Obstfeld (1983), and White and Domowitz (1984), have shown the asymptotic variance-covariance matrix of the parameter estimates to be

$$(H_T'W_TH_T)^{-1}H_TW_TS_TH_T(H_T'W_TH_T)^{-1}.$$ 

$H_T$ is the matrix of the expected value of the partials of $h_T$ with respect to the parameters being estimated or

$$H_T = ((1/T)\Sigma_{t=1}^T(\partial h_{1T}/\partial \alpha), \ldots , (1/T)\Sigma_{t=1}^T(\partial h_{nT}/\partial \alpha)).$$ 

As we use $W_T = S_T^{-1}$ this simplifies to

$$(H_T'W_TH_T)^{-1}.$$ 

The properties of the generalized method of moments (GMM) estimators are addressed in Hansen (1982) where he finds the estimates to be consistent and asymptotically normal. The properties of the MSM estimators are addressed in Duffie and Singleton (1989) for asset prices following
diffusion processes. Duffie and Singleton (1989) show that the estimators are still consistent and asymptotically normal. Bossaerts (1989) addresses the properties of the MSM estimators for contingent claim prices. This presents another problem as pointed out in Bossaerts and Hillion (1989) as there is a kink at the exercise price. They show that consistency carries over from Hansen (1982), but his proof of asymptotic normality does not as the condition of continuous differentiability is violated. Hence, they discuss two methods to satisfy asymptotic normality. The first is to smooth the kink by rewriting the payoff function so it becomes continuously differentiable. This can be accomplished as shown in Bossaerts and Hillion (1989) through an appropriate change of variables. The second method is to appeal to the arguments of Pakes and Pollard (1986) where continuous differentiability is not required, but their arguments depend on the randomness of the sample.

The property of asymptotic normality is important to the hypothesis developed in Hansen (1982) where he shows that the number of observations T times the minimum value of the criterion function approximates a chi-square variable with degrees of freedom equal to the number of orthogonality conditions less the number of parameters being estimated.

\[ T(h_T'W_T'h_T) \sim \chi^2(nm-1). \]

Hansen and Singleton (1982) point out that the number of orthogonality conditions must be no less than the number of parameters to be estimated. In fact, to carry out the \( \chi^2 \) test above on the overidentifying restrictions the number of orthogonality conditions needs to be at least one more than the number of parameters being estimated. To form t-tests for the

---

8 When implying both \( \alpha \) and \( \sigma^2 \) the degrees of freedom number is \( nm-2 \).
parameter estimates one simply multiplies the square root of the sample size, $T$, and the parameter estimate and divides this by the square root of the appropriate diagonal element of the asymptotic variance-covariance matrix for the estimators.\(^9\)

D. DATA

We use the MSM estimator to imply a value of $\alpha$ and then $\alpha$ and $\sigma$ using Standard and Poor's 100 (S&P 100) options contracts on the Chicago Board Options Exchange (CBOE). To obtain our simulated stock value in Equation (2) we require starting values that are simply the series of stock values over our sample period from January 1, 1984 to December 31, 1986. The series of stock values for this three year period is also used to calculate the mean and the variance of the return. We use the daily closing quotes provided by the CBOE for the S&P 100 Index. The observed call values are calculated as the midpoint of the bid-ask spread as quoted on the Berkeley Options Tapes. The corresponding strike prices and maturity dates are also taken from the tapes. The time-to-maturity is calculated as the number of trading days left until the maturity of the option contract. The riskless rates are calculated from the discount rates of the most recently issued 90 day Treasury Bill as quoted in the Federal Reserve Statistical Release H15.\(^{10}\)

\(^9\) For the case in which we only imply the $\alpha$ parameter the asymptotic variance-covariance matrix is a scalar as only one parameter is being estimated.

\(^{10}\) The equation used to calculate the yield from the discount rate is $y = 36500d/[m(100-d)]$, where

(continued...)
To obtain our sample of call options, we find the last nine observations of the day from the combinations of three strike price categories, the nearest-to-the-money, nearest-in-the-money, and nearest-out-of-the-money with three maturity categories consisting of the three furthest months to expiration. As opposed to normal equity option contract cycles, the S&P 100 follows a monthly expiration cycle with four contracts trading at any one time over the period of our sample. Over the sample period there are 733 trading days on the CBOE. From this dataset we draw three random series by choosing randomly for each series one observation from among the nine for each of the 733 trading days. In terms of our previous notation $n = 3$ and $T = 733$.

E. RESULTS

Applying the methodology discussed to our data, we run two series of tests. In the first, we imply $\alpha$ as the only unknown parameter and estimate $\sigma$ along with $\mu_p$ from our S&P 100 closing value series. In the second, we imply both $\alpha$ and $\sigma$ as unknown parameters and estimate only $\mu_p$. Both the

---

$^{10}$ (...continued)
y = annualized yield in percent,  
d = discount in dollars, and 
m = days to maturity.

$^{11}$ A call option is at-the-money when the underlying stock price is equal to the exercise price. A call option is in-the-money when the stock price is greater than the exercise price. A call option is out-of-the-money when the stock price is less than the exercise price.

The nearest-in(out-of)-the-money option is actually the next nearest-in(out-of)-the-money option if the nearest-to-the-money option is in(out-of)-the-money.
strike price, \( X \), and the time-to-maturity, \( T \), are used as instrumental variables which yields three sets of orthogonality conditions:

\[
E[e_{it}] = 0 \quad (3)
\]

\[
E[e_{it}X_{it}] = 0 \quad (4)
\]

\[
E[e_{it}T_{it}] = 0 \quad (5)
\]

with \( i = 1 \) to \( 3 \) representing the random series and \( t = 1 \) to \( 733 \) the observation in a particular series \( i \).

There are a number of possible tests to run with MSM which yield additional information. Looking first at the three series represented by Equation (3) the outcome of the \( \chi^2 \) test tells us whether or not the weighted average errors resulting from our valuation in Equation (1) and the observed call prices are close enough to zero. If not close enough to zero, our valuation is rejected. However, our null hypothesis is joint, such as that in Bossaerts and Hillion (1989), based on having: (1) the correct model valuation, (2) accurate parameter estimation, (3) market efficiency, and (4) no measurement error. If the outcome of the \( \chi^2 \) test is not significantly different from zero, we fail to reject our valuation. If we fail to reject our valuation and if the \( t \)-statistic corresponding to the parameter estimate is significant, then the parameter value is an acceptable estimate. Chi-square tests on the three series represented by Equation (4) or (5) can be interpreted to tell us whether or not the previously reported biases of strike price and time-to-maturity for the Black-Scholes formula may exist in our valuation.\(^2\) If we reject the \( \chi^2 \) test then the value is not close enough to zero and the bias, strike if

\(^2\) For a discussion of this extant literature on biases in the BS formula see Chapter 4.
(4) or time-to-maturity if (5), may exist. If we fail to reject the $\chi^2$ test, we interpret this as a lack of the corresponding bias. These moment conditions can also be tested jointly in groups of two sets or in a group of three sets.

For our first run, implying only $\alpha$, the results of the tests using only the sample variance-covariance matrix for weighting are in Table 3-1. For each of the tests of (3), labeled as N, (4), labeled as X, or (5), labeled as T, individually the degrees of freedom for the $\chi^2$ statistic is two. The number of orthogonality conditions is three, corresponding to the number of random series, from which we subtract the number of parameters estimated which is one, $\alpha$. All three of these tests fail to reject our joint hypothesis and yield risk aversion parameter values of .2042, .2244, and .2070 with significant t-statistics. When we test them in groups of two, the number of degrees of freedom for the $\chi^2$ statistic is five. The number of orthogonality conditions is six, from which we subtract one, which is the number of parameters estimated. Now we find we reject our joint hypothesis when we test either (3) and (5) or (4) and (5) jointly, but again fail to reject when we test (3) and (4) jointly.

In the second stage of our two stage estimation process, we calculate and correct for the autocovariances. These results for implying $\alpha$ only are reported in Table 3-2. The autocovariance corrected sample variance-covariance matrices are reported in Tables 3-3, 3-4, and 3-5. The results in Table 3-2 are very similar to Table 3-1. Again we find we do not reject any of the sets of orthogonality conditions tested separately, however, their $\chi^2$ statistics have increased with the use of the autocovariance corrected sample variance-covariance matrix. The results
for the groups of two sets of orthogonality conditions are the same also except the $\chi^2$ value has dropped significantly for the group including the sets of orthogonality conditions (3) and (4).

We also estimate both $\alpha$ and $\sigma$. The results of tests using only the sample variance-covariance matrix for weighting are provided in Table 3-6. For each of the tests of (3), (4), or (5) individually the degrees of freedom for the $\chi^2$ statistic are one. The number of orthogonality conditions is three from which we subtract, two, the number of parameters estimated. We fail to reject our joint hypothesis in all three of these tests of individual sets of orthogonality conditions. The $\alpha$ parameter estimates with their corresponding t-statistics in parentheses are .2582 (6.744), .2823 (6.775), and .2653 (6.093). The respective $\sigma^2$ parameter estimates are .9753E-04 (5.578), 1.009E-04 (4.676), and 1.005E-04 (4.847). The $\sigma^2$ value we calculated directly from the first run is .7690E-04. Both the $\alpha$ and $\sigma^2$ values are higher than their counterparts when we use an estimate of $\sigma^2$ and imply only $\alpha$. Considering the joint tests of a group of two sets of orthogonality conditions, we find results similar to the first study in terms of the $\chi^2$ statistics. For the tests of (3) and (5) and (4) and (5) we reject our hypothesis, but we fail to reject our valuation with (3) and (4). For the two groups of two sets of orthogonality conditions we reject, the parameter estimates of $\alpha$ and $\sigma^2$ along with their respective t-statistics have large increases.

In the second stage of our two stage estimation process, we calculate the autocovariances and compute the autocovariance corrected sample variance-covariance matrix. The results are in Table 3-7. The autocovariance corrected sample variance-covariance matrices are reported
in Tables 3-8, 3-9, and 3-10. The results in Table 3-7 are only slightly different from those of Table 3-6. The most notable difference is as is found in moving from Table 3-1 to Table 3-2. We again find that the $\chi^2$ statistic for the test of the group of two sets of orthogonality conditions corresponding to (3) and (4) has significantly decreased. The risk aversion parameter estimate has increased to 0.2389 (6.841) and the variance parameter estimate has increased to 0.8679E-04 (6.730) from 0.1978 (4.961) and 0.7328E-04 (6.313), respectively.

F. CONCLUSION

The constant proportional risk aversion parameter values which are implied when our valuation model is not rejected have significant t-statistics and range from about 0.20-.28. The consumption capital asset pricing model shows the excess expected return on an asset is simply a product of the constant proportional risk aversion parameter and the covariance of consumption changes and excess returns. Given the covariance of consumption changes and excess returns, we can see our smaller $\alpha$ value will result in lower excess expected returns than the $\alpha$ values of most of the previous studies which use consumption data.

Our low risk aversion values also imply that a puzzle similar to the "equity premium puzzle" does not exist in the options market, as the estimation of our model produces "reasonable" risk aversion values. A potential explanation for the contrast in equity and option markets may come from the noisy trader literature of Shleifer and Summers (1990) and Cutler, Poterba, and Summers (1991) which relax the rational investor
assumption. It can be argued in their framework that a basic security such as equity is valued by the fundamentals of the economy and subject to investor whims. Whereas, a derivative security, such as an option, which is subject to many bounding relationships like those of Merton (1973) and put-call parity is influenced less by investor whims.
Table 3-1

Implying \( \alpha \) Using the Generalized Method of Moments with Variance-Covariance Matrix Used as the Weighting Matrix

<table>
<thead>
<tr>
<th>Moment Condition</th>
<th>( \alpha )</th>
<th>t-stat</th>
<th>( \chi^2 ) Sample</th>
<th>Degrees of Freedom</th>
<th>P-value of ( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>.2042</td>
<td>28.30</td>
<td>1.607</td>
<td>2</td>
<td>.5522</td>
</tr>
<tr>
<td>T</td>
<td>.2244</td>
<td>30.73</td>
<td>2.064</td>
<td>2</td>
<td>.6437</td>
</tr>
<tr>
<td>X</td>
<td>.2070</td>
<td>29.21</td>
<td>1.620</td>
<td>2</td>
<td>.5551</td>
</tr>
<tr>
<td>N,T</td>
<td>.1888</td>
<td>25.82</td>
<td>178.4</td>
<td>5</td>
<td>.9999</td>
</tr>
<tr>
<td>N,X</td>
<td>.2105</td>
<td>30.09</td>
<td>10.12</td>
<td>5</td>
<td>.9281</td>
</tr>
<tr>
<td>T,X</td>
<td>.2027</td>
<td>28.71</td>
<td>97.55</td>
<td>5</td>
<td>.9999</td>
</tr>
<tr>
<td>N,T,X</td>
<td>.1775</td>
<td>24.73</td>
<td>204.8</td>
<td>8</td>
<td>.9999</td>
</tr>
</tbody>
</table>
Table 3-2

Implying $\alpha$ Using the Generalized Method of Moments with Variance-Covariance Matrix Corrected for Autocovariance Used as the Weighting Matrix

<table>
<thead>
<tr>
<th>Moment Condition</th>
<th>$\alpha$</th>
<th>$t$-stat</th>
<th>$\chi^2$ Sample</th>
<th>Degrees of Freedom</th>
<th>P-value of $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>.2038</td>
<td>16.84</td>
<td>2.114</td>
<td>2</td>
<td>.6525</td>
</tr>
<tr>
<td>$T$</td>
<td>.2240</td>
<td>18.55</td>
<td>2.569</td>
<td>2</td>
<td>.7232</td>
</tr>
<tr>
<td>$X$</td>
<td>.2075</td>
<td>17.53</td>
<td>2.280</td>
<td>2</td>
<td>.6802</td>
</tr>
<tr>
<td>$N,T$</td>
<td>.1648</td>
<td>13.37</td>
<td>120.8</td>
<td>5</td>
<td>.9999</td>
</tr>
<tr>
<td>$N,X$</td>
<td>.2095</td>
<td>18.03</td>
<td>5.877</td>
<td>5</td>
<td>.6816</td>
</tr>
<tr>
<td>$T,X$</td>
<td>.2073</td>
<td>17.87</td>
<td>56.09</td>
<td>5</td>
<td>.9999</td>
</tr>
<tr>
<td>$N,T,X$</td>
<td>.1494</td>
<td>12.27</td>
<td>133.4</td>
<td>8</td>
<td>.9999</td>
</tr>
</tbody>
</table>
Table 3-3

Variance-Covariance Matrices for Each Set of Three Orthogonality Conditions (α):

Three orthogonality conditions of deviations:

<table>
<thead>
<tr>
<th></th>
<th>3.623</th>
<th>3.423</th>
<th>3.421</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.890</td>
<td>3.628</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.925</td>
<td></td>
</tr>
</tbody>
</table>

Three orthogonality conditions of deviations times time-to-maturity:

<table>
<thead>
<tr>
<th></th>
<th>10716.</th>
<th>9338.</th>
<th>9928.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10986.</td>
<td></td>
<td>10163.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13206.</td>
</tr>
</tbody>
</table>

Three orthogonality conditions of deviations times exercise price:

<table>
<thead>
<tr>
<th></th>
<th>129171.</th>
<th>121155.</th>
<th>121898.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>137644.</td>
<td>128900.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>142806.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3-4

Variance-Covariance Matrices for Each Combination of Two Sets of
Three Orthogonality Conditions (α):

Six orthogonality conditions of deviations times time-to-maturity and deviations:

<table>
<thead>
<tr>
<th></th>
<th>11544</th>
<th>9553</th>
<th>10065</th>
<th>202</th>
<th>185</th>
<th>183</th>
</tr>
</thead>
<tbody>
<tr>
<td>11971</td>
<td>11971</td>
<td>179</td>
<td>215</td>
<td>190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14644</td>
<td>190</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Six orthogonality conditions of deviations times exercise price and deviations:

<table>
<thead>
<tr>
<th></th>
<th>129258</th>
<th>120044</th>
<th>120550</th>
<th>679</th>
<th>640</th>
<th>637</th>
</tr>
</thead>
<tbody>
<tr>
<td>137970</td>
<td>137970</td>
<td>637</td>
<td>732</td>
<td>670</td>
<td></td>
<td></td>
</tr>
<tr>
<td>141649</td>
<td>640</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Six orthogonality conditions of deviations times time-to-expiration and deviations times exercise price:

<table>
<thead>
<tr>
<th></th>
<th>11084</th>
<th>9453</th>
<th>10028</th>
<th>36743</th>
<th>33952</th>
<th>34246</th>
</tr>
</thead>
<tbody>
<tr>
<td>11317</td>
<td>11317</td>
<td>10332</td>
<td>32676</td>
<td>38230</td>
<td>34647</td>
<td></td>
</tr>
<tr>
<td>13786</td>
<td>35268</td>
<td>37109</td>
<td>42951</td>
<td>122320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>129313</td>
<td>121067</td>
<td>138800</td>
<td>128406</td>
<td>143859</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3-5

Variance-Covariance Matrix for Three Sets of Three Orthogonality Conditions (α):

Nine orthogonality conditions of deviations times time-to-maturity,

deviations times exercise price, and deviations

\[
\begin{array}{ccccccc}
11493 & 9488 & 9994 & 38005 & 33934 & 34028 & 202 & 183 & 182 \\
12029 & 10704 & 33391 & 40190 & 35590 & 179 & 215 & 189 \\
14667 & 35618 & 38249 & 45029 & 191 & 206 & 238 \\
133264 & 12214 & 122264 & 699 & 649 & 645 \\
143029 & 130417 & 646 & 755 & 686 & 774 \\
148224 & 648 & 4 & 3 & 4 & 4 \\
\end{array}
\]
Table 3-6

Implying $\alpha$ and $\sigma$ Using the Generalized Method of Moments with Variance-Covariance Matrix Used as the Weighting Matrix

<table>
<thead>
<tr>
<th>Moment Condition</th>
<th>$\alpha$</th>
<th>t-stat</th>
<th>$\sigma$</th>
<th>t-stat</th>
<th>$\chi^2$</th>
<th>Sample</th>
<th>P-value of $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>.2582</td>
<td>6.744</td>
<td>.9753E-04</td>
<td>5.578</td>
<td>.1402</td>
<td>.2919</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>.2823</td>
<td>6.775</td>
<td>1.009E-04</td>
<td>4.676</td>
<td>1.702</td>
<td>.8112</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>.2653</td>
<td>6.093</td>
<td>1.005E-04</td>
<td>4.847</td>
<td>.2097</td>
<td>.3530</td>
<td></td>
</tr>
<tr>
<td>N,T</td>
<td>.5025</td>
<td>26.95</td>
<td>5.109E-04</td>
<td>5.323</td>
<td>39.60</td>
<td>.9999</td>
<td></td>
</tr>
<tr>
<td>T,X</td>
<td>.5147</td>
<td>29.41</td>
<td>5.769E-04</td>
<td>5.433</td>
<td>38.60</td>
<td>.9999</td>
<td></td>
</tr>
<tr>
<td>N,T,X</td>
<td>.4994</td>
<td>29.31</td>
<td>4.786E-04</td>
<td>5.976</td>
<td>133.6</td>
<td>.9999</td>
<td></td>
</tr>
</tbody>
</table>
Table 3-7

Implying $\alpha$ and $\sigma$ Using the Generalized Method of Moments with Variance-Covariance Matrix Corrected for Autocovariance Used as the Weighting Matrix

<table>
<thead>
<tr>
<th>Moment Condition</th>
<th>$\alpha$</th>
<th>$t$-stat</th>
<th>$\sigma$</th>
<th>$t$-stat</th>
<th>$\chi^2$</th>
<th>P-value of $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>.2595</td>
<td>7.057</td>
<td>9.828E-04</td>
<td>5.934</td>
<td>.1840</td>
<td>.3320</td>
</tr>
<tr>
<td>T</td>
<td>.2804</td>
<td>6.207</td>
<td>1.009E-04</td>
<td>4.365</td>
<td>1.987</td>
<td>.8413</td>
</tr>
<tr>
<td>X</td>
<td>.2686</td>
<td>6.694</td>
<td>1.013E-04</td>
<td>5.295</td>
<td>.1230</td>
<td>.2742</td>
</tr>
<tr>
<td>N,T</td>
<td>.5643</td>
<td>49.15</td>
<td>1.088E-04</td>
<td>6.194</td>
<td>63.33</td>
<td>.9999</td>
</tr>
<tr>
<td>N,X</td>
<td>.2389</td>
<td>6.841</td>
<td>8.679E-04</td>
<td>6.730</td>
<td>4.877</td>
<td>.6998</td>
</tr>
<tr>
<td>T,X</td>
<td>.3797</td>
<td>12.85</td>
<td>1.967E-04</td>
<td>5.371</td>
<td>34.18</td>
<td>.9999</td>
</tr>
<tr>
<td>N,T,X</td>
<td>.5773</td>
<td>56.82</td>
<td>1.085E-04</td>
<td>6.719</td>
<td>99.83</td>
<td>.9999</td>
</tr>
</tbody>
</table>
Table 3-8

Variance-Covariance Matrices for Each Set of Three Orthogonality Conditions (α and σ):

Three orthogonality conditions of deviations:

<table>
<thead>
<tr>
<th></th>
<th>3.596</th>
<th>3.412</th>
<th>3.410</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.851</td>
<td>3.586</td>
<td>3.883</td>
</tr>
</tbody>
</table>

Three orthogonality conditions of deviations times time-to-maturity:

<table>
<thead>
<tr>
<th></th>
<th>10494.</th>
<th>9154.</th>
<th>9733.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10898.</td>
<td>9936.</td>
<td>12594.</td>
</tr>
</tbody>
</table>

Three orthogonality conditions of deviations times exercise price:

<table>
<thead>
<tr>
<th></th>
<th>126493.</th>
<th>117735.</th>
<th>119220.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>133183.</td>
<td>125132.</td>
<td>138860.</td>
</tr>
</tbody>
</table>

Table 3-9

Variance-Covariance Matrices for Each Combination of Two Sets of
Three Orthogonality Conditions (α and σ):

Six orthogonality conditions of deviations times time-to-maturity
and deviations:

<table>
<thead>
<tr>
<th></th>
<th>13836</th>
<th>8983.</th>
<th>8955.</th>
<th>253.</th>
<th>183.</th>
<th>167.</th>
</tr>
</thead>
<tbody>
<tr>
<td>13263</td>
<td>13263</td>
<td>9761.</td>
<td>171.</td>
<td>239.</td>
<td>182.</td>
<td></td>
</tr>
<tr>
<td>13317</td>
<td>13317</td>
<td>172.</td>
<td>189.</td>
<td>5.</td>
<td>4.</td>
<td>3.</td>
</tr>
</tbody>
</table>

Six orthogonality conditions of deviations times exercise price and
deviations:

<table>
<thead>
<tr>
<th></th>
<th>128622</th>
<th>12105.</th>
<th>121260.</th>
<th>678.</th>
<th>646.</th>
<th>641.</th>
</tr>
</thead>
<tbody>
<tr>
<td>136238</td>
<td>136238</td>
<td>128156.</td>
<td>643.</td>
<td>726.</td>
<td>676.</td>
<td></td>
</tr>
<tr>
<td>143628</td>
<td>143628</td>
<td>645.</td>
<td>683.</td>
<td>4.</td>
<td>3.</td>
<td>3.</td>
</tr>
</tbody>
</table>

Six orthogonality conditions of deviations times time-to-expiration and
deviations times exercise price:

<table>
<thead>
<tr>
<th></th>
<th>11303.</th>
<th>9446.</th>
<th>9958.</th>
<th>37441.</th>
<th>32560.</th>
<th>35005.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10983.</td>
<td>10983.</td>
<td>9963.</td>
<td>34267.</td>
<td>37015.</td>
<td>36026.</td>
<td></td>
</tr>
<tr>
<td>13287.</td>
<td>13287.</td>
<td>33862.</td>
<td>33536.</td>
<td>41867.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>131685.</td>
<td>131685.</td>
<td>121004.</td>
<td>120898.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>134445.</th>
<th>125307.</th>
<th>142113.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3-10

Variance-Covariance Matrix for Three Sets of Three Orthogonality Conditions ($\alpha$ and $\sigma$):

Nine orthogonality conditions of deviations times time-to-maturity, deviations times exercise price, and deviations

<table>
<thead>
<tr>
<th></th>
<th>13623.</th>
<th>8423.</th>
<th>8640.</th>
<th>45780.</th>
<th>31728.</th>
<th>30279.</th>
<th>249.</th>
<th>174.</th>
<th>162.</th>
</tr>
</thead>
<tbody>
<tr>
<td>12799.</td>
<td></td>
<td>9248.</td>
<td>29229.</td>
<td>41892.</td>
<td>31697.</td>
<td>160.</td>
<td>232.</td>
<td>172.</td>
<td></td>
</tr>
<tr>
<td>12511.</td>
<td>12511</td>
<td>30339.</td>
<td>40103.</td>
<td>167.</td>
<td>182.</td>
<td>215.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>165384.</td>
<td>113585.</td>
<td>108726.</td>
<td>884.</td>
<td>613.</td>
<td>577.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>151502.</td>
<td>118160.</td>
<td>607.</td>
<td>815.</td>
<td>626.</td>
<td>752.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>142495.</td>
<td>588.</td>
<td>588.</td>
<td>636.</td>
<td>636.</td>
<td>752.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.</td>
<td>5.</td>
<td>4.</td>
<td>3.</td>
<td>4.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4
LITERATURE REVIEW FOR ESTIMATION OF VOLATILITY

The Black-Scholes (BS) European call option valuation formula is very appealing as even though the call price is a function of five variables, the underlying asset value \( S \), exercise price \( X \), riskless rate \( r \), time-to-maturity \( T \), and the standard deviation of the underlying asset \( \sigma \), two are given in the option contract itself, \( X \) and \( T \), and two others, \( S \) and \( r \), can be observed in the economy. This leaves only the variance of the rate of return on the underlying asset to estimate. There has been much research in this area with the variance of the rate of return estimated in essentially three categories. The first category is an unconditional estimate of the standard deviation based on historical returns data. Comprising the second is a series of methods proposed to imply a conditional estimate of the standard deviation from the BS model. The third category consists of conditional estimates for the volatility based on autoregressive conditional heteroscedasticity and some of its extensions.

A. HISTORICAL VOLATILITY

The first category and method for obtaining an estimate of the volatility on the underlying asset is the calculation of the observed variance from the past stock returns series. This is the method applied in the first test of the Black-Scholes model by Black and Scholes (1972). They use the volatility from the past returns data of the underlying
security for one year previous to the day the option is written as an estimate. Two other early studies by Galai (1977) and Finnerty (1978) also employ an historical estimate of the volatility based on the underlying stock's past returns data. Black and Scholes (1972) compare their historical estimate of the volatility with the volatility of the underlying stock's returns data calculated over the actual life of the contract. They find the observed volatility is a better input for the BS model than the historical. Thus, others began looking for a better estimate of the volatility than a simple unconditional historical estimate.

B. BLACK-SCHOLES IMPLIED VOLATILITY

The second category for estimating the standard deviation of the rate of return by implying it from the BS model has numerous methods. These methods are proposed in an attempt to account for the various pricing biases manifested in the many studies of option valuation with BS. In the first study of the BS model, Black and Scholes (1972) found a bias in the model with respect to the volatility of the rate of return. The model overprices options on high variance stocks and underprices options on low variance stocks. Finnerty (1978) also finds evidence to corroborate that of Black and Scholes. Other studies have found evidence of time-to-maturity and striking price biases as well. MacBeth and Merville (1979) find a striking bias that increases as the options move away from at-the-money in either direction. They also find the BS model overprices relative to the market for out-of-the-money options and underprices relative to the
market for in-the-money-options. Black (1975) finds the reverse of these results in his study. He also finds a time-to-maturity bias as the closer an option gets to expiration the worse the BS prices reflect market realizations. Finally, Rubinstein (1985) finds a time-to-maturity bias as well. He also finds the same results as Black (1975) and MacBeth and Merville (1979) with respect to the exercise price bias. They each consider a different time interval and he considers them both to find that the direction of the exercise price bias switches between the intervals.

The second category for estimation of the standard deviation of the rate of return uses an estimate implied from the BS model. The methods in this category use the given information of the exercise price and the time-to-maturity combined with the observable information of the stock price, riskless rate, and option price to leave a problem consisting of one equation, the BS model, and one unknown, the volatility of the rate of return. Due to the complexity of the BS model an analytic solution is not obtained, nevertheless, a volatility estimate can be implied from the model using numerical techniques. The methods in the second category are various attempts at adjusting for the biases discussed above. To get an estimate of the volatility of the rate of return at a particular point in time with n call option observations some form of an average is taken. This has been addressed, generally, in one of two ways. i) The first method used to calculate an average is to find a weighted average of the n calls which is accomplished again in one of two ways. a) One approach to the weighted average method, referred to hereafter as a volatility weighting scheme, is to first calculate the standard deviations individually by implying them from the BS formula and then applying
weights to these standard deviations to obtain one estimate. b) The second approach to the weighted average method, referred to hereafter as an error weighting scheme, is one in which weights, designed to account for strike price and time-to-maturity biases, are applied directly to the options. Then an estimate of the standard deviation for a set of options is implied by an optimization process that minimizes the weighted deviation of the observed call price from the call price of the BS formula with respect to standard deviation. ii) The second method of calculating an average is to eliminate those options purported to create the biases and then take a simple average of the standard deviations implied from BS with the remaining data.

The first weighted average method with a volatility weighting scheme is proposed by Latane and Rendleman (1976). They propose the use of a weighted average implied standard deviation (WISD) with weights based on the partial derivative of the call price with respect to the standard deviation of the stock return. Allowing j to denote the different options from 1 to n and d the partials, the weights \( w_j \) are

\[
 w_j = \frac{dC_j}{d\sigma_j}.
\]

Having obtained the implied standard deviations (ISD), \( \sigma_j \), for the different options separately, they find

\[
 WISD = \left( \sum_j w_j^2 \sigma_j^2 \right)^{1/2} / \left( \sum_j w_j \right).
\]

The second weighted average method with a volatility weighting scheme is that of Chiras and Manaster (1978). They are unsatisfied with Latane and Rendleman's (1976) method as their weights are based on absolute changes in the prices of options relative to changes in the implied standard deviations. They suggest that this is inappropriate because returns are
not measured by absolute price changes, but, rather are measured by price changes relative to total investment. Hence, they use the price elasticity of options with respect to their ISDs as it considers the percentage change in the price of an option with the percentage change in its ISD. Their weights are

\[ w_j = \left(\frac{dC_j}{d\sigma_j}\right) \left(\frac{\sigma_j}{C_j}\right) \]

and with the implied standard deviations obtained from the options separately they use the following WISD equation:

\[ \text{WISD} = \left(\frac{\Sigma_j w_j \sigma_j}{\Sigma_j w_j}\right). \]

The last weighted average method with a volatility weighting scheme is that of MacBeth and Merville (1979). They get a simple estimate for each stock for each day by assuming an at-the-money option yields the best estimate. Thus, the intercept, \( b_0 \), of a regression is the estimate of the ISD for a specific day. The dependent variable is an individual option \( j \)'s implied standard deviation, \( \sigma_j \), and the independent variable is the ratio of the underlying stock price, \( S_j \), minus the present value of the exercise price, \( X_j \), to the present value of the exercise price. That is

\[ \sigma_j = b_0 + b_1 m_j + e_j, \]

where

\[ m_j = \frac{[S_j - X_j \exp(-rT)]}{X_j \exp(-rT)}, \]

with \( T = \) time to expiration for option \( j \).

This intercept, \( b_0 \), is analogous to Latane and Rendleman's and Chiras and Manaster's WISDs.

The first weighted average method with an error weighting scheme is developed in Whaley (1982). Using the fact that the market price of an option is equal to the model valuation price plus an error term, he
minimizes the sum of these square errors to imply his standard deviation estimate. He uses a first order linearization process to solve this nonlinear problem. He expands the call value in a Taylor series around an initial value \( \sigma_0 \)

\[
C_j = C_j(\sigma_0) + (dC_j/d\sigma)|_{\sigma_0} (\sigma - \sigma_0) + \text{higher order terms} + e_j.
\]

\( C_j \) is the observed call value and \( C_j(\sigma_0) \) is the BS call value evaluated at \( \sigma_0 \). In the volatility weighting schemes where the ISD's, \( \sigma_j \), are implied for each option \( j \) separately from the BS model and then the WISD is simply a weighted average of these ISD's, \( C_j = C_j(\sigma_j) \). In the error weighting schemes where the ISD's are not implied separately for each option \( j \) from the BS model, but rather, a single ISD or WISD, \( \sigma \), is implied from all of the options on an underlying stock over the relevant time period or a single ISD or WISD, \( \sigma \), is implied for all the options of the same maturity on an underlying stock over the relevant time period, \( C_j \) does not equal \( C_j(\sigma) \). Shifting known values to the left and disregarding higher order terms he gets

\[
C_j - C_j(\sigma_0) + \sigma_0 (dC_j/d\sigma)|_{\sigma_0} = \sigma (dC_j/d\sigma)|_{\sigma_0} + e_j.
\]

Using ordinary least squares (OLS) to obtain an estimate of \( \sigma \), he checks for convergence of \( \sigma \) to \( \sigma_0 \). If the convergence test is not satisfied the process is redone with the estimate of \( \sigma \) replacing \( \sigma_0 \) to find a new estimate of \( \sigma \).

A second weighted average method with an error weighting scheme is developed in Beckers (1981). He finds that the ISD's for deep in-the-money close to maturity options can differ from the ISD of an at-the-money option by as much as ten times. Hence, he chooses to use a weighting scheme which emphasizes at-the-money options' ISD's. Allowing \( j \) to denote
the different options on a given stock with the same maturity for a given day, the weights \( w_j \) are

\[
w_j = \frac{dC_j(\sigma)}{d\sigma}.
\]

However, these weights are not directly applied to ISD's to find a WISD. Instead, they are used in the minimization of the following function

\[
\sum_j w_j [C_j - C_j(\sigma)]^2 / \sum_j w_j,
\]

where \( C_j \) is the observed call value and \( C_j(\sigma) \) is the BS call value evaluated at the ISD. Beckers (1981, p. 370) points out "... the actual weights used in this procedure are proportional to the squared values of the Latane-Rendleman weights." Hence, his method weights more heavily options more sensitive to variance specification, namely at-the-money options.

The last weighted average method using an error weighting scheme is in Day and Lewis (1988) and is a variation of the method in Whaley (1982). They minimize a sum of weighted errors squared where the errors are the differences between the model option value given the implied standard deviation estimate and the observed option price. Whereas in Whaley (1982) the updated estimate is calculated with OLS, here it is calculated with generalized least squares (GLS). The weights are assigned to options with different exercise prices on the same underlying stock with the same expiration according to the percent of the day's trading volume they represent within that particular expiration. Hence, they imply out a separate volatility estimate for each different expiration of options on an underlying stock as in Whaley. However, with GLS they weight these estimates according to the percent of the day's trading volume for the different exercise prices. They show this latter part places more weight
on at-the-money or one strike price out-of-the-money options. Since this method assigns little weight to thinly traded and far out-of-the-money options, they expect to lessen the effect of asynchronous trading and the size of the bid-ask spread.

The second way used to imply a standard deviation from the BS model for a sample of options is to eliminate those options creating the biases and then take a simple equally weighted average of the standard deviation of the rates of return implied form the BS model with the remaining data. This method is used by Schmalensee and Trippi (1978) who simply choose a set of criteria to eliminate options for which the BS model does not price well relative to the market. They use four criteria to exclude options with: "1) premia less than $1.00, 2) premia less than 1% of S, 3) if X < S, premia less than 1.5(S-X), and 4) remaining lives less than three weeks." The premium is the amount the option price exceeds the value S-X when the call is in-the-money and zero when the call is out-of-the-money. The first two criteria remove deep out-of-the-money options and options for which transactions costs may be important. The third eliminates deep in-the-money options and the fourth eliminates options too near expiration. Then to form their estimate of the standard deviation of the rate of return from the volatilities implied from the BS model they take a simple average.

Trippi (1977) also uses this second way to find an estimate of the volatility of the rate of return. In fact, to deal with the biases he uses criteria 1), 4) and a modified version of 3), where he excludes options having premia less than 1.3(S-X), of the Schmalensee and Trippi (1978) study. Finally, Patell and Wolfson (1979) use this method as well. They
exclude options with less than thirty days to expiration or whose option price is less than one dollar. They check the robustness of their results by repeating the study excluding options whose price is less than fifty cents instead of a dollar and ceterus paribus find the results to be consistent. They replicate the study again only this time using only one option that is nearest-to-the-money and find consistent results.

Although these studies using implied volatility are the result of a belief that the implied volatility is a better estimate of the future volatility than that based on historical data, only three studies address the veracity of this belief. They are Latane and Rendleman (1976), Chiras and Manaster (1978), and Beckers (1981). Latane and Rendleman (1976) compare the estimate of the variance of the rate of return given by their WISDs with the estimate given by standard deviations calculated on the past returns data of the underlying stock. They use stock and option prices for twenty four companies listed on the Chicago Board Options Exchange (CBOE) over a thirty-eight week sample period, October 5, 1973 to June 4, 1974. Riskless rates are calculated from Treasury Bills whose payment date is closest to the maturity date of the option. Comparisons are made by finding correlations on four series:

1) The WISD averaged over the 38 week sample period (39 observations) for each of the 24 companies.

2) The standard deviation of monthly log price relative returns calculated over the four year period ending September 30, 1973 for each company.
3) The standard deviation of weekly log price relative returns calculated over the 38 week sample period time adjusted to a monthly basis for each of the 24 companies.
4) The standard deviation of monthly log price relatives for each of the 24 companies calculated over the two year period ending March 31, 1975."

The highest correlation found is that between series 1), WISD averaged over the sample period, and series 4), observed standard deviation over the sample period and into the future. They also find series 3), the observed standard deviation over the sample period, and series 4), the observed standard deviation over the sample period and into the future, are more highly correlated with series 1), WISDs averaged over the sample period, than with series 2), historical standard deviation. Based on these results, Latane and Rendleman (1976) conclude that the WISD is a better estimate of future volatility than an historical estimate.

Chiras and Manaster (1978) calculate an implied standard deviation, using numerical techniques, from Merton's (1973) dividend corrected version of the BS model. Their data consist of monthly closing stock and option prices, for firms trading on the CBOE on June 29, 1973, over a twenty-three month period, June 1973 to April 1975. They also collect data on the dividends paid over the period, and riskless rates are calculated from Treasury Bills whose payment date is closest to the expiration date of the option. They also calculate an estimate of the future volatility based on the standard deviation of the returns data for the underlying stock over the sample period. They compare these two estimates of future volatility to a standard deviation of returns data on the underlying stock.
succeeding the sample period with three simple cross-sectional regressions across the firms in the study for each month. The future standard deviation is the dependent variable in all and the historical standard deviation is the only independent variable in one, their WISD is the only independent variable in a second, and finally, the third has both as independent variables. By comparing the R-squared values of the first two regressions and finding the second to be consistently higher, Chiras and Manaster (1978) conclude that their WISD is a better estimate of future volatility than historical volatility. As the R-squared from the second and third sets of regressions remain nearly the same, they conclude that using historical standard deviations adds no information not already contained in their WISDs.

Beckers (1981) partially corrects for dividends with the method suggested by Black (1975). He reduces the stock price by the present value of some fraction of the dividends whose ex-dividend dates are before the option is exercised. Then he compares the values of the option if exercised before the ex-dividend date with the value of the option exercised on the expiration date. The maximum of these two values is taken as the estimate of the observed call price. This method has been criticized as it only allows the probability of early exercise to be zero or one. However, Beckers finds it provides a reasonable approximation to the true value.

With daily closing stock and option prices for firms trading on the CBOE over 75 days from October 13, 1975 to January 23, 1976. Beckers eliminates options that do not trade at least 25 contracts on 50 trading days to try and reduce problems associated with infrequently traded
options. For this dataset Beckers compares the WISD of Latane and Rendleman, his method of weighting, and a method using the ISD of the single most sensitive option or the option with the greatest weight under his weighting scheme. The comparison is based on how well these three methods predict the future volatility over the remaining life of the option determined by a set of three cross-sectional regressions at each time period. All regressions have the volatility over the remaining life of the option as their dependent variable. The independent variables for each of the three regressions are the three methods of computing an ISD estimate, where the ISD for each method serves as the only independent variable in one regression. Comparing R-squared values, he finds his method of weighting superior to that of Latane and Rendleman, however, both are inferior to the method using the ISD implied from the at-the-money option.

Finding ISD's volatile over time, he tests and finds a simple average of the ISD's over time 'significantly increases' their forecasting ability. He points out this result may be explained by the fact that this averaging diminishes the errors-in-measurement problem. The errors-in-measurement problem stems from the lack of knowledge as to whether a transaction took place at the bid or the ask, the closing quotes on the option and stock exchanges are asynchronous, and the stock and option prices are rounded. These problems may be important on a given day, but over a series of days they would be expected to average out.

Expanding his sample to data from all U. S. option exchanges obtained from the Interactive Data Corporation database and a longer time period, slightly over two years, Beckers (1981) chooses to study the
options in ten five day intervals which trade at least 25 contracts on each of the five days within an interval. To compare the predictive abilities of various methods used to estimate ISD's, he sets up regressions and compares R-squared values. This time he uses six different regressions in which the dependent variable for all is the observed volatility over the remaining life of the option. The sets of independent variables for the six regressions are: 1) observed volatility over previous quarter, 2) the simple average of his weighting method over five days in interval, 3) the simple average of the method using a single at-the-money option over five days in interval, 4) a combination of 1) and 3), 5) Fisher Black's volatility estimate, and 6) a combination of 1) and 5).

His results show that in general, not in every case but in the majority (7 out of 10), the implied standard deviation predicts future volatility better than past standard deviations do. He also concludes from his results that 3) contains at least as much information as 2) from which he infers adding options to the at-the-money option only lessens the predictive ability. Finally, he considers the regressions for 4) and 6). As the R-squared are not usually significantly increased from 5) to 6), he concludes Black has included historical volatility estimates. On the other hand, as the R-squared values increase in every period from 3) to 4), he suggests some information from historical volatility may not have been revealed in observed market prices. However, this issue is nebulous as the systematic bias introduced by the dividend correction method chosen may account for this. The evidence from these three studies supports the
belief that the implied volatility is a better estimate of the future volatility than that based on historical data.

C. GARCH VOLATILITY

The second category for estimating the standard deviation of the rates of return by implying them from the BS model grew from a search for a better estimate of future volatility and an attempt to account for biases found in empirical studies. The third category of estimating the variance based on ARCH and variations of GARCH also grows from a search for a better estimate of future volatility and an attempt to account for a changing variance. In early studies by Black and Scholes (1972), Black (1975), and Latane and Rendleman (1976), all find evidence that the standard deviation of the rate of return changes over time in contrast to the assumption posited by Black and Scholes (1973) of a constant variance of the rate of return. Merton (1973) shows that the BS model can be extended to incorporate a changing variance as long as it is a deterministic function of time.

With evidence against the assumption of a non-stochastic variance, alternative models are developed to accommodate a stochastic variance. One such model is the constant elasticity of variance model introduced by Cox (1973) of which BS is a special case. Other cases of this model allow the variance to change stochastically with the stock price. Another set of models is developed with an independent stochastic process to allow implied volatilities to vary randomly. These models are developed by Hull and White (1987), Johnson and Shanno (1987), and Scott (1987). Now, in
addition to the stochastic process describing stock price movements necessary to derive the BS model, a second stochastic process describing the movement of variance is also part of the model. This significantly complicates the problem as the differential of the call price now has two random components as opposed to the one in BS which is eliminated with riskless hedging by a portfolio of stock and an appropriate amount of a call. An analogous result is not as tractable here, as the portfolio would require a stock and two options, but the value of one option would be required to find the value of the other which yields a problem in determining which comes first.

Johnson and Shanno (1987) assume an asset exists with the same stochastic process as the variance to form a hedge portfolio to derive a differential equation they are unable to solve. They use Monte Carlo simulation to obtain numerical results for the price of a call as the variance of the rate of return changes stochastically. Hull and White (1987) start with the differential equation an option must satisfy developed by Cox, Ingersoll, and Ross (1985). Unwilling to assign much likelihood to finding an asset with the same stochastic process as the variance, they rule out the possibility of forming a hedge portfolio to eliminate risk. Instead, they use the multi-factor Capital Asset Pricing Model to describe the return generating process and assume volatility is uncorrelated with aggregate consumption and the stock price to obtain a risk neutral valuation relationship from which they derive an analytic solution to the call option pricing problem with stochastic variance. Scott (1987) tried to solve his set of two stochastic differential equations originally by forming a riskless hedge with a stock and two
options on that stock to find the problem of needing to know the value of one option in order to value the other. He then followed a process similar to Hull and White to find a solution.

Hull and White (1987) obtain option values from their analytic solution. However, finding the assumption of no correlation between volatility and aggregate consumption too restrictive, they relax it and obtain option values through Monte Carlo simulation as they no longer can obtain an analytic solution. They find that relative to their option values based on a stochastic volatility uncorrelated with the stock price the BS model overvalues at-the-money options and undervalues deep in and out-of-the-money options. They use numerical techniques, when the volatility is correlated with the stock price, to find that with positive correlation out-of-the-money options are underpriced by BS and in-the-money options are overpriced. With negative correlation out-of-the-money options are overpriced and in-the-money options are underpriced. They also find these effects to be intensified by increases in volatility, volatility of volatility, or time-to-maturity.

Evidence from the options field of a stochastic variance is corroborated by a series of papers dealing with the description of the returns generation process. Epps and Epps (1976) and Tauchen and Pitts (1983) find evidence of a changing variance. A whole series of papers model this changing variance with a mixture of distributions. Epps and Epps (1976) and Tauchen and Pitts (1983) try to model the distribution of the variance as a function of the trading volume. Blattberg and Gonedes' (1974) model assumes the distribution of the variance is an inverted gamma
distribution. Kon (1984) uses a discrete mixture of normal distributions. These are just a few examples from this literature.

Akgiray (1989, p. 56) finds two common assumptions throughout the literature on the distribution of returns data: "1) returns are independent, and 2) the return generating process is a linear process with parameters that are independent of the past realizations of the process." Financial models typically include the mean and variance and Akgiray cites Poterba and Summers (1986) as providing evidence against the assumption of constant conditional means and variances. Akgiray (1989) himself provides evidence against both the independence and linearity assumptions. Using daily stock returns data on the Center for Research in Security Prices, CRSP, value-weighted and equal-weighted indices over a twenty four year period, January 1963 to December 1986, he finds evidence against independence of daily returns and in support of a nonlinear dependence structure in the daily returns series. Through further testing, he finds the residual series intertemporally dependent, but uncorrelated. Hence, specification of a simple autoregressive, AR(1), process to describe the autocorrelation in daily return series is inadequate as it will fail to take into account the information available from the intertemporal dependence of the squared residuals.

A model which can account for most of this dependence, called autoregressive conditional heteroscedasticity, ARCH, has been developed in Engle (1982). ARCH improves on previous time series techniques because in addition to allowing the mean of a process to depend on past realizations of the underlying series, it also allows the variance of a process to
depend on past realizations of the squared residuals. The ARCH(p) model is:

\[R_t | I_{t-1} \sim F(\mu_t, h_t),\]

\[R_t = \mu_t + e_t,\]

\[E_{t-1}[e_t] = 0, \text{ and}\]

\[h_t = \sum_{i=1}^{p} \alpha_i e_{t-i}^2,\]

where \(R_t\) is a random variable at time \(t\), \(I_{t-1}\) is the information set at time \(t-1\), \(F\) is the conditional distribution of the random variable \(R_t\), \(\mu_t\) is the conditional mean, \(h_t\) is the conditional variance, and \(e_t\) is the residual. The order of the ARCH process \(p\) must be an integer greater than zero and specified in advance of the model estimation. The \(\alpha\) parameters must satisfy \(\alpha_0 > 0\) and \(\alpha_i \geq 0\).

Bollerslev (1986), citing early empirical work finding long lags in the conditional variance equation and using fixed lag structures to circumvent the negative parameter estimates for the variance, generalized the ARCH model to allow for longer memory and a more flexible lag structure. He accomplishes this by allowing the variance to depend not only on past realizations of the squared residuals, but also on the past realizations of the conditional variance. His generalized ARCH model is GARCH(p,q), which is simply the ARCH model above with the following \(h_t\):

\[h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i e_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}.\]

The parameters \(p\) and \(q\) are the orders of the process that need to be specified in advance with \(p\) greater than zero and \(q\) greater than or equal to zero. The ARCH(p) process is simply a GARCH(p,q) process with \(q\) equal to zero.
To estimate the GARCH\((p,q)\) model the conditional distribution, \(F\), of the random variable needs to be specified. Much of the early literature uses a normal distribution. However, Bollerslev (1987) uses a Student's t-distribution specification on daily foreign exchange price data for GARCH\((1,1)\) and finds it fits better than a normal GARCH\((1,1)\). Engle (1982) and Bollerslev (1986) find maximum likelihood estimates. The log likelihood equation for a sample of size \(n\) with \(t\) going from 1 to \(n\) is

\[
L(\theta|p,q) = \frac{1}{n} \sum_{t=1}^{n} \ln[f(e_t,h_t)]
\]

where \(\theta=(\mu, \sigma_0, \ldots, \sigma_p, \beta_1, \ldots, \beta_q)\), \(f\) is the density of the specified conditional distribution for the random variable, and \(e_t\) and \(h_t\) are recursively calculated from the model. In Engle (1982), the method of scoring algorithm is used and in the case of ARCH it is shown to be simply a least squares regression on transformed variables.\(^{13}\) However, in Bollerslev (1986), the partials of the likelihood function have recursive terms that complicate the problem. He uses the Berndt, Hall, Hall, and Hausman (1974), BHHH, algorithm to update his parameter estimates.\(^{14}\) Engle (1982) and Bollerslev (1986) point out that the iterations on the mean and variance parameters can be done separately as a result of their assumption of the block diagonality of the information matrix.

\(^{13}\) The method of scoring algorithm updates parameters by adding the inverse of the information matrix evaluated at the current estimate times the derivative of the log likelihood function evaluated at the current estimate times a variable step-length to the current estimate to get a new estimate. The information matrix is just the negative of the expected value of the Hessian.

\(^{14}\) The BHHH algorithm updates parameter estimates by adding the product of a variable step-length, the sum over the number of the observations of the gradients of the logarithm of the function, and the inverse of the sum over the number of observations of the gradients of the logarithm of the function times their transposes to the current estimate.
Bollerslev (1986) suggests the use of the Box-Jenkins (1976) framework to identify the orders $p$ and $q$. A likelihood ratio test can also be used to select between alternative model specifications as well as Lagrange multiplier tests developed in Engle (1982) and Bollerslev (1986). Akgiray (1989) implicitly uses likelihood ratio tests for his daily returns series data and finds that GARCH processes fit the data much better than ARCH processes. He also shows with likelihood ratio and Lagrange multiplier tests that ARCH models are significant against a null hypothesis of a normal process with non-stochastic variance. Hence, GARCH processes fit the best. Testing among GARCH processes with likelihood ratio tests he finds no significant improvement in fit over the GARCH(1,1).

Engle and Bollerslev (1986) describe forecasting with ARCH and GARCH. They argue that the use of these models with conditional variances for forecasting should bring gains similar to those gains from using conditional means in forecasting. One step ahead forecasts for the conditional variance, $h_t$, for ARCH(1) are given explicitly by the ARCH(1) model once the parameters have been estimated as

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2.$$  

To calculate multistep forecasts of the conditional variance for ARCH it must be recalled that $E_t[e_t] = 0$ and $h_t = E_t[e_t^2] - (E_t[|e_t|])^2$ to provide the result that $h_t = E_t[e_t^2]$. Then taking expectations of

$$h_{t+1} = \alpha_0 + \alpha_1 e_{t+1-1}^2,$$

yields

$$E_t[h_{t+1}] = \alpha_0 + \alpha_1 E_t[e_{t+1-1}^2].$$

Using the law of iterative expectations
\[ E_t[h_{t+1}] = a_0 + a_1 E_t[e_{t+1-2}^2] \]

and substituting we get
\[ E_t[h_{t+1}] = a_0 + a_1 E_t[h_{t+1-1}] \]

which is valid for \( i \geq 2 \). For \( i = 1 \), the \( e \) is observed. To forecast for GARCH(1,1) take expectations of
\[ h_{t+1} = a_0 + a_1 e_{t+1-1}^2 + \beta_1 h_{t+1-1} \]

to yield
\[ E_t[h_{t+1}] = a_0 + a_1 E_t[e_{t+1-1}^2] + \beta_1 E_t[h_{t+1-1}] \]

Now applying the law of iterative expectations they obtain
\[ E_t[h_{t+1}] = a_0 + a_1 E_t[E_{t+1-2}[e_{t+1-1}^2]] + \beta_1 E_t[h_{t+1-1}] \]

With a conditional mean of zero using the definition of variance we have
\[ E_t[h_{t+1}] = a_0 + (a_1 + \beta_1) E_t[h_{t+1-1}] \]

which holds for \( i \geq 2 \). For \( i = 1 \), \( e \) and \( h \) are observed.

Akgiray (1989) compares forecasts of volatility for each of his four periods by splitting the series into two parts each consisting of 480 days. He uses the first 480 days to estimate the models and then forecasts the next 20 days with these model estimates. Then he drops the first 20 observations from the 480 and adds 20 new observations to get a new set of estimates for the models and forecasts the next 20 days. He continues this process until the days in the period are exhausted which yields 24 sets of forecasts. He calculates the actual variance to be compared with the forecasts taking into account first-lag autocorrelation of daily returns. He compares this with a simple historical average, an exponentially weighted moving average, an ARCH(2) forecast, and a GARCH(1,1) forecast. Comparisons are done by mean error, root mean square error, mean absolute error, and mean absolute percent error. He finds GARCH(1,1) does much
better than the other three and offers 'substantial improvement' relative to those normally used, such as a simple unconditional historical average.

A recent study by Day and Lewis (1990), also in the vein of GARCH estimates of future volatility, compares the information content of estimates of future volatility from volatilities implied from a dividend corrected version of the BS model with the information content of estimates of future volatility from two extended versions of GARCH. Assuming premature exercise is not optimal, they use the dividend corrected version of the BS formula, for a finite number of known dividends over the option's life, that simply reduces the underlying asset value by the present value of these dividends. They imply both the volatility and the underlying asset value from this model using a technique similar to Day and Lewis (1988), but the updated estimates from GLS now must account for the underlying stock also being estimated.

To deal with asynchronous trading, Day and Lewis (1990) imply the underlying stock price from their option valuation model along with the volatility arguing that these implied asset values should be free of asynchronous trading effects. They find little difference when comparing sample statistics of the actual and implied returns series. They also find small and insignificant values for the autocorrelations for all lags in both series. This supports the use of the GARCH methodology. They also use the likelihood ratio test for various specifications of GARCH and find results in agreement with Akgiray (1989) that GARCH(1,1) is not significantly improved upon by any higher order specification.

One extended version of GARCH considered in Bollerslev, Engle, and Wooldridge (1988), referred to as GARCH in mean or GARCHM, that they use
has the conditional volatility in the mean equation. They also use an exponential GARCH or EGARCH model proposed by Nelson (1989) to account for empirical evidence found in earlier papers by Black (1976), Christie (1982), and French, Schwert, and Stambaugh (1987). The latter group identified a negative relationship between returns data and unanticipated volatility increases which Nelson interprets as an asymmetric relation between volatility and historical data. Day and Lewis (1990) also cite the work of Pagan and Schwert (1989) in support of the choice of EGARCH as the latter find EGARCH performs better than GARCH in forecasting out of sample using returns data on Standard and Poor's 500 Index.

The returns are described by the two following equations for all the models they test:

\[ R_t = \mu + \epsilon_t + \lambda h_t^{1/2}, \]
\[ \epsilon_t \sim N(0, h_t). \]

The rest of the model descriptions are as follows:

**GARCHM**

\[ h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}, \]
\[ \alpha_0 > 0, \alpha_1 > 0, \text{ and } \beta_1 > 0. \]

**EGARCH**

\[ \ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 (\Phi \tau_{t-1} + \gamma(|\tau_{t-1}| - (2/\pi)^{1/2})), \]
\[ \tau_{t-1} = e_{t-1}/h_{t-1}^{1/2}. \]

**GARCHM with implied volatility \( \sigma_{t-1}^2 \)**

\[ h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} + \delta \sigma_{t-1}^2. \]

**EGARCH with implied volatility \( \sigma_{t-1}^2 \)**

\[ \ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 (\Phi \tau_{t-1} + \gamma(|\tau_{t-1}| - (2/\pi)^{1/2})) + \delta \ln(\sigma_{t-1}^2). \]
Their study is conducted on closing prices of stocks and options on the Standard and Poor's 100 Index over a period from March 11, 1983 thru December 31, 1988. Riskless rates are calculated as the average of bid and ask discounts on the Treasury Bill with maturity date closest to that of option's maturity date and actual dividends are used to proxy expected dividends. They estimate the implied volatility using the options with the shortest maturity trading at the beginning of the week as they suggest these represent the closest ex-ante forecast of future volatility available to proxy a weekly time interval.

Conclusions can be drawn from the models on the additional information provided by implied volatilities relative to the model of conditional variance it is in by the statistical significance of $\delta$. The appropriate likelihood ratio tests can provide the information on which conclusions can be drawn for tests of usefulness of implied volatilities to forecast future volatility and of whether or not the separate extensions of GARCH provide additional information relative to implied volatilities. The latter requires two additional conditional variance specifications, for GARCHM, $h_t = \alpha_0 + \delta \sigma_{t-1}^2$, and for EGARCH, $\ln(h_t) = \alpha_0 + \delta \ln(\sigma_{t-1}^2)$.

Their results suggest that implied volatilities contain information not in the conditional volatility of GARCHM or EGARCH, but the latter also appear to contain information not available in the former. Thus, none of the models appear to fully explain the conditional volatility of stock returns. Their evidence suggests a model combining ex-ante forecasts can explain in sample volatility better than any of the three models.
separately, implied volatility, or conditional volatility from either GARCHM or EGARCH.

This study points out that the two conditional forecasts of volatility used in the past, BS implied volatilities and modified ARCH models, contain some different information. The previous studies all compare unconditional forecasts of volatility with one of these two conditional volatility forecasts. In the next chapter, we compare the two conditional volatility forecasts.
This study focuses on forecasting the volatility of the underlying stock on which an option is written. The existing literature that has compared volatility forecasts has compared an unconditional volatility estimate, historical volatility, with a conditional volatility estimate, either an implied volatility from the Black-Scholes European call option pricing model or some variation of the autoregressive conditional heteroscedasticity (ARCH) model. This extant literature finds the conditional estimate outperforms the unconditional estimate. In this study, we are going to compare the ability of the two conditional volatility estimates to forecast volatility.

There are essentially three categories of estimates that have been used in the past to estimate the future volatility. The first category is an estimate of the unconditional standard deviation of the rate of return based on historical returns data. Comprising the second is a series of methods proposed to imply an estimate of the standard deviation from the Black-Scholes (1973), (BS), European call option pricing model. As these standard deviation estimates are based on data from the options market, they are also based on the information contained in these prices and are hence conditional standard deviation estimates. The third category consists of estimates for the conditional volatility based on the generalized autoregressive conditional heteroscedasticity (GARCH) methodology proposed in Bollerslev (1986).
The historical estimate is the one originally employed in Black and Scholes (1972). Trying to improve the volatility forecast, Latane and Rendleman (1976) are the first to imply an estimate from the BS formula itself. With the call price a function of five variables, the exercise price and the time-to-maturity set out in the option contract, the underlying stock price and riskless rate observable in the economy, only the standard deviation is unknown. Using the BS equation and observed call prices one can use numerical techniques to imply an estimate of the standard deviation. In this category, numerous methods have been developed in an attempt to account for the various biases, such as time-to-maturity and striking price, reported in past empirical studies of options. To get an estimate of the standard deviation at a particular point of time with multiple option observations some form of an average is taken. This has been addressed, generally, in one of two ways: (i) find a weighted average or (ii) eliminate those options purported to create the empirical biases. Once the options have been eliminated in (ii) a simple average is taken of the standard deviations individually implied from BS with the remaining data. The weighted average method, (i), has also been approached in two ways. (a) First calculate the standard deviations individually by implying them from the BS formula and then apply weights to these standard deviations to obtain one estimate. (b) The second approach is one in which weights, designed to account for strike price and time-to-maturity biases, are applied directly to the options. Then an estimate of the standard deviation for a set of options is implied by a numerical optimization process. This process minimizes the weighted deviations of the observed
call prices from the call prices of the BS formula by searching over the standard deviation parameter of the BS formula.

Three earlier studies have compared the forecasting performance of volatility of the historical volatility based on the returns series of the underlying stock and the implied volatility from BS based on options data. Latane and Rendleman (1976), Chiras and Manaster (1978), and Beckers (1981) all find that the estimate based on the volatility implied from BS performs better than that based on the historical returns series. Another study performed by Akgiray (1989) compares the estimates of the standard deviation using the GARCH methodology with an estimate based on the historical returns series. Among various ARCH and GARCH specifications he finds GARCH(1,1) performs the best. In a comparison of this GARCH specification with the historical estimate, he finds GARCH(1,1) provides a better estimate of future volatility than the historical estimate. Considering the information content of option prices, Day and Lewis (1990) find that while volatilities implied from the BS model contain information not in their GARCHM(1,1) or EGARCH(1,1) volatility estimates, the latter also contain information not in the former. They include the implied volatilities in their modified GARCH models and use likelihood ratio tests to compare the information content of the volatility estimates.

These last two studies provide the motivation for this study. The main purpose of this chapter is to compare the ability to forecast the volatility of the conditional volatility estimates, GARCH(1,1) and volatilities implied from the BS formula. In the first section, we discuss the data used. The second section discusses the methodology used for the
forecasts and their evaluation. Our results for daily data are presented in the third section.

A. DATA

The sample has 25 of the thirty firms used in Rubinstein (1985). These firms are listed in Table 5-1 along with the averages of their GARCH parameters over our 506 model estimations as discussed in the methodology section. The period for which forecasts are studied is from January 3, 1983 to December 31, 1984. For the sample period, options prices, taken to be the midpoint of the bid-ask spreads, are obtained from the Berkeley option tapes. Also obtained from the Berkeley option tapes are the corresponding exercise prices and maturity dates. The time-to-maturity is calculated as the number of trading days on the Chicago Board Options Exchange (CBOE) between the observation date and the expiration date. Riskless rates are proxyed by the yields as calculated from the discount rates of the most recently issued 90 day Treasury Bill as quoted in the Federal Reserve Statistical Release H15.

The data on stock returns are obtained from the Center for Research in Securities Prices (CRSP) tapes. The period for which the daily

---

15 The five firms from Rubinstein's thirty firm sample not included in our study are Houston Oil and Mineral (HOI), Kennecott (KN), Polaroid (PRD), Syntex (SYN), and Xerox (XRX).

16 The returns for the CRSP tape are calculated as described in the CRSP Stock File Guide as follows:
"For time t (a holding period), let
\[ r(t) = \text{return on purchase at } t', \text{ sale at } t \]
\[ p(t) = \text{last sale price or closing bid/ask average at time } t \]

(continued...)
returns data are extracted is from January 3, 1978 to June 21, 1985. The estimation period for the GARCH and historical forecasts starts at January 3, 1978, which is five years prior to the start of the forecast comparison period. Dividend information is also obtained from the CRSP tapes.

B. METHODOLOGY

1. Ex-post Volatility Estimation

To determine which series is a better forecast of volatility, the actual\textsuperscript{17} future volatility, FUT, is calculated from the ex-post returns. We calculate four separate future volatility estimates with which we shall compare our volatility forecasts. The intervals chosen for our forecast horizons are 20, 40, 80 and 120 days. These are chosen so that they are less than the maximum time-to-maturity of stock options. We choose equal intervals of 40, 80 and 120 days with an additional interval of 20 days as our primary focus in this study is on the conditional estimates that are expected to be better estimates in the short-term. As for the unconditional historical estimates, the distinction between these unconditional estimates and the conditional estimates is expected to be more pronounced in the short-term. The actual future volatility over the

\begin{itemize}
\item \textsuperscript{16} (...continued)
\end{itemize}
\begin{align*}
d(t) &= \text{cash adjustment for } t \\
f(t) &= \text{price adjustment factor for } t \\
\text{then } r(t) &= \left[ \frac{(p(t)f(t)+d(t))/p(t')}{p(t')} \right] - 1.\textsuperscript{17}
\end{align*}
Dividends are taken care of with \texttt{d(t)} and stock splits with \texttt{f(t)}.

\textsuperscript{17} Actual is italicized as the situation is similar to that pointed out in Whaley (1982) where he emphasizes that the volatility based on observed returns is only an estimate of the real volatility.
given number of trading days is calculated by taking a simple standard deviation of the returns over this same number of days following the day of interest. For example, for day 1 in Figure 5-1, the future standard deviation for the next 20 days is calculated from the observations for days 2 through 21 or equivalently $i = 1266$ to 1285. Similarly, for the 40 day estimation period the future standard deviation is calculated from the observations for days 2 through 41.

2. Implied Volatility from the Black-Scholes Formula

The first forecast of volatility we consider is that obtained by using numerical techniques to imply an estimate of volatility from the BS formula. These estimates are to be based on call option observations from the day of interest. With respect to Figure 5-1 the estimates for day 1 are based on observations from day 1 or $i = 1265$. In implying an estimate from BS, we take into account the empirical evidence of striking price bias and time-to-maturity bias manifested in the extant literature. We will account for these biases using two methods. In the first, an implied volatility is found for the nearest-to-the-money mid-maturity option.\textsuperscript{18} The inclusion of this estimate is a result of Beckers (1981) study in which he found the at-the-money option implied volatility estimate outperformed the weighted average implied volatility estimates.

In the second method, an implied volatility is found for a set of no more than nine options on a given day for an underlying stock. These nine options are the combinations of three maturity categories (near, mid and

\textsuperscript{18} If there is no observation for the mid-maturity option, we use the far maturity option.
far) and three strike price categories (nearest-out-of-the-money, nearest-to-the-money and nearest-in-the-money).\textsuperscript{19} We use this method of implying a single volatility estimate from multiple option observations on a single underlying stock for two reasons. First, we want a volatility estimate that is based on the larger pool of information available from multiple options. Second, the extant literature also uses an estimate based on multiple options. As a first step to try and minimize the biases we have chosen the three strike price categories closest to at-the-money. The at-the-money options according to the extant literature are least biased and the time-to-maturity bias is exacerbated as we move away from at-the-money. Hence, we exclude deep in or out-of-the-money options from the sample.

Day and Lewis (1988) find implied volatilities for each expiration series of options. They find implied volatilities of expiring options are significantly different from other options in the four day interval prior to expiration. As one moves further before the expiration day, the statistical significance declines monotonically and becomes insignificant outside of their event window. They also find the difference of non-expiring options implied volatilities are insignificant. As our estimates do not contain observations of options within the month of their expiration, their results would indicate that we should not have a problem aggregating options of different expiration.

\textsuperscript{19} The nearest-in(out-of)-the-money option is actually the next nearest-in(out-of)-the-money option if the nearest-the-money option is in(out-of)-the-money.

There may be less than nine options if options did not trade in all nine of the categories we were looking for on a given day.
Stein (1989) also looks at the differences in implied volatility for options of differing expirations, specifically one and two month options on the Standard and Poor's 100 Index. Using an average volatility until expiration interpretation developed theoretically in Merton (1973) and interpreted empirically in Patell and Wolfson (1979), he finds that the implied volatilities are not as different as expected under a regime of a mean reverting stochastic volatility process and rational expectations. Although he does not discuss the statistical significance of the difference between the one and two month implied volatilities, the fact that they are not as different as expected provides weak support for aggregating the implied volatilities of options of different expirations on the same underlying stock.

The implied volatility for the second method using multiple option observations is found by minimizing the weighted errors in the loss function

$$\min \ e'\Omega^{-1}e.$$  

Here $e$ is a column vector of the errors, $e_i$, which are the deviations of the observed call option prices, $C_{Oi}$, from those generated by BS, $C_{BSi}$, for a given standard deviation, that is,

$$e_i = C_{Oi} - C_{BSi}, \quad i = 1 \ to \ 9.$$  

$\Omega$ is the variance-covariance matrix of the errors. Assuming BS is correct we would expect a better volatility estimate the smaller the error. As the errors manifest how close the BS value is to the observed value, we weight by the inverse of the variance-covariance matrix.
This involves a two stage estimation method. On the first pass we find an implied volatility by minimizing the sum of squared errors on each day for each firm, that is,

$$\min_{\sigma} e' e.$$

Then a variance-covariance matrix is calculated based on the estimated errors over these days. Finally, a new implied volatility is found by minimizing the loss function $e' \Omega^{-1} e$.

In estimating the above implied volatilities we drop some observations. We eliminate puts because we are using the BS formula for European call options and puts introduce a more complex early exercise problem. We also eliminate those options whose price is less than fifty cents. This last filter eliminates options for which transactions costs are important.\(^{20}\) The last observation of an option satisfying the above criteria is found each day for each maturity on an underlying stock. Then, following Rubinstein (1985), applying the test for early exercise potential of call options proposed by Black (1975), two call option values are calculated, with actual dividends used to proxy expected dividends. One is based on exercising the option just before the last ex-dividend date at time $t$ previous to the expiration of the option

$$C_1 = \max(0, S - X e^{-rT}).$$

The second is based on holding the option until expiration at time $T$

\(^{20}\) Evidence of the significance of transactions costs on options priced below fifty cents is provided in Phillips and Smith (1980). They find an average percentage bid-ask spread of thirty percent for all call options, but only a four and one-half percent average bid-ask spread for call options trading above fifty cents.
\[ C_2 = \max(0, S - Xe^{-rT} - De^{-rt}). \]

If the first value, \( C_1 \), exceeds the second value, \( C_2 \), the observation is eliminated from the sample as early exercise is probable and we are faced with the problem of valuing an American call.

It is from this filtered data set that implied volatilities are found using the BS formula. In the BS formula the stock value needs to be reduced by the present value of the future dividends, again proxied by actual dividends, to be paid over the remaining life of the option. For each underlying stock an implied volatility from the BS formula is obtained for each day it is traded for each of the two methods used to account for the empirical biases. BS1 represents the estimate based on the one at-the-money option while BS3X3 represents the estimate based on the nine options.

Corresponding to the 25 firms there are 25 such sets of two estimates for each day, one estimate for each of the two methods. The interpretation of these estimates varies. Black and Scholes (1973) originally interpret the volatility as instantaneous given their assumption of constant volatility. Merton (1973) interprets the volatility as the average future volatility of the underlying stock for the period between the observation date and the expiration date given his extension of BS to incorporate a variance that is a deterministic function of time. Stein (1989) has shown volatilities of different expirations are not as different as expected. Day and Lewis (1988) show that the difference of

---

21 If there is more than one dividend payment before the expiration of the option we can just subtract the present value of the additional dividend payments from \( S \) in \( C_1 \) and \( C_2 \) above.
implied volatilities between options with different expirations is significant if one of the options is expiring within 4 days, but declines significantly as one moves back from day -2 to day -4. They also provide evidence that the differences of the implied volatilities of options with different expirations are insignificant if neither of the options is expiring. As we do not have options within their month of expiration, we aggregate the different expiration options on the same underlying stock to yield a single volatility estimate. We then interpret this as the average volatility estimate over any subinterval within the longest maturity option's expiration date. Hence, our BS1 and BS3X3 estimates are the same for the volatility forecast horizons of 20, 40, 80 and 120 days. The implied volatilities from the BS formula for the first method are computed using the IMSL subroutine DNEQNF which employs the Levenberg-Marquardt algorithm to solve a system of non-linear equations for the same number of unknowns as equations in the system. The implied volatilities from the BS formula for the second method are computed using the IMSL subroutine BCONF which employs a quasi-Newton optimization algorithm.

---

22 The Levenberg-Marquadt algorithm updates parameter estimates by adding the product of the sum over the number of observations of the products of the gradients and the variable vectors with the inverse of the sum over the number of observations of the gradients times their transposes plus a variable step-length times the identity matrix. This is a variant of the Gauss-Newton method.

23 IMSL's BCONF subroutine uses a quasi-Newton algorithm with a finite difference gradient or numerical approximation to the derivative. A quasi-Newton method updates estimates by subtracting the product of a variable step-length, a positive definite matrix to approximate the Hessian, and the gradient from the current estimate. The initial matrix supplied to approximate the Hessian can be any positive definite matrix as it will converge to the inverse of the Hessian as the algorithm progresses.
3. GARCH Model Estimation

Future volatilities are also forecast with a GARCH(1,1) model. The GARCH model as introduced in Bollerslev (1986) is a generalization of the ARCH model introduced by Engle (1982). The particular specification, GARCH(1,1), is used because we think it is reasonably general while at the same time parsimonious. It is also chosen as it is shown to perform best among ARCH and other GARCH specifications in Akgiray (1989) and Day and Lewis (1990). Parsimony is important in our study because as we discuss below we estimate over 10000 GARCH models. The specific GARCH model to be estimated is

\[ R_t | I_{t-1} \sim N(\mu_t, h_t), \]
\[ R_t = \mu_t + e_t, \]
\[ E_{t-1}(e_t) = 0, \text{ and} \]
\[ h_t = \alpha_0 + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1}, \]
\[ \alpha_0 > 0, \alpha_1 \geq 0, \text{ and } \beta_1 \geq 0, \]

where

- \( R_t \) is the return at time \( t \) on the underlying stock,
- \( I_{t-1} \) is the information set at time \( t-1 \),
- \( N \) is the conditional cumulative normal distribution,
- \( \mu_t \) is the conditional mean,
- \( h_t \) is the conditional variance,
- \( e_t \) is the residual at time \( t \), and
- \( E \) is the expectations operator.

With a program that uses the Berndt, Hall, Hall and Hausman algorithm, estimates of the GARCH model for day 1, are computed as discussed in

\[ ^{24} \text{There are other potential distribution assumptions such as the Student-t distribution.} \]
Bollerslev (1986) using maximum likelihood estimation. The GARCH model parameters are estimated from the series of 1264 observations from January 3, 1978 to December 31, 1982 for each of the 25 firms. This corresponds to $i = 1$ to 1264 in Figure 5-1. For subsequent days we use a rolling GARCH model. To estimate a new GARCH model for day 2 the first observation in the series, January 3, 1978, is dropped and that for day 1, January 3, 1983, is added. This corresponds to $i = 2$ to 1265 in Figure 5-1. This process is performed 506 times for each firm, so that a GARCH model is estimated for each trading day on the NYSE over the study period from January 1, 1983 to December 31, 1984.

These estimated models are then used to forecast the volatility of the returns series. The GARCH model for a particular day is used to forecast the volatility at 20 (GARCH20), 40 (GARCH40), 80 (GARCH80) and 120 (GARCH120) days to be compared with the corresponding actual future volatility calculations. For example, to get an estimate of the future volatility over the next 20 trading days on day 1 for a specific firm, the estimated GARCH model is used to forecast the conditional variance as described in Engle and Bollerslev (1986), for the next 20 trading days which correspond to days 2 through 21 in Figure 5-1. That is, taking conditional expectations of both sides of the conditional variance equation, we have the expected future conditional variance for day $t+s$ is

$$E_t[h_{t+s}] = \alpha_0 + \alpha_1 E_t[e_{t+s-1}^2] + \beta_1 E_t[h_{t+s-1}].$$

Applying the law of iterative expectations, we have

$$E_t[h_{t+s}] = \alpha_0 + \alpha_1 E_t[E_{t+s-2}[e_{t+s-1}^2]] + \beta_1 E_t[h_{t+s-1}].$$

Noting that $E_{t+s-2}[e_{t+s-1}^2] = h_{t+s-1}$ as $E_{t+s-2}[e_{t+s-1}] = 0$, we then find

$$E_t[h_{t+s}] = \alpha_0 + (\alpha_1 + \beta_1) E_t[h_{t+s-1}].$$
We can get $h_{t+1}$ from the GARCH model itself and calculate $h_{t+s}$ for $s$ between 2 and 21. Then a simple average of these 20 forecasted conditional variances is taken to yield an average forecast of the volatility for the next 20 days. A similar procedure is followed for the 40, 80 and 120 day forecast intervals as well. This process is then repeated for all 506 days for each of the 25 firms.

4. Estimation of Historical Volatility

Finally, historical volatilities, calculated as the standard deviations of historical return series, are estimated for comparison. We calculate historical standard deviations based on the historical return series 20, 40, 80 and 120 days previous to the day of interest to be compared with the actual future volatility over the corresponding number of days. For example, HIST20 on day 1 is calculated as a simple standard deviation from $i = 1245$ to 1264 in Figure 5-1. On day 1, HIST40 is a simple standard deviation from $i = 1225$ to 1264. HIST80 and HIST120 are calculated similarly. For day 2, HIST20 is a simple standard deviation from $i = 1246$ to 1265.

Another historical volatility estimate is to be calculated. This is because some preliminary results indicate the superior performance found in previous studies of the implied volatilities from BS relative to historical volatility estimates may be a result of the period of the returns data chosen to calculate the historical volatility. Thus, an historical estimate, HIST, is calculated using the preceding five years data. It is compared to each of the actual future volatility forecast
horizons. With respect to Figure 5-1, for day 1 HIST is a simple standard deviation of returns from $i = 1$ to 1264.

5. Forecast Comparisons

In our study, for comparison of forecasts we use methodology similar to that in previous studies, such as Latane and Rendleman (1976), Beckers (1981), and particularly Chiras and Manaster (1978). There are four different forecast horizons used, 20, 40, 80 and 120 days. For each forecast horizon, we employ the following methods of comparison:

1) In the regression method we use the regression

$$FUT = a + b \times X.$$  

Here $X$ is one of the five forecasts described in the preceding sections corresponding to the appropriate forecast horizon. For example if the forecast horizon is twenty days, with $FUT_{20}$, $X$ is one of $BS1$, $BS3X3$, GARCH20, HIST, and HIST20. Each regression is cross-sectional over the 25 firms in our sample. Over our sample period there are 496 such regressions for our forecasts calculated from options data and 506 such regressions for our forecasts calculated from stock returns data. We compare the average slope, average intercept, and average R-squared from the regressions over our sample period for each of the five forecasts.

---

25 A thorough discussion of the appropriate techniques for forecast comparison is contained in Granger and Newbold (1986).

26 For options data the number of firms with observations on a particular day is never less than 18. In fact, there are only 7 days of the 496 for which the number of firms is less than 24.
2) We also calculate an average mean absolute deviation (MAD). The average MAD for the 20 day volatility forecast horizon for example is computed as

\[ \text{AVGMAD} = \frac{1}{d} \sum_j \sum_i \frac{|FUT20_{ij} - X_{ij}|}{n} \]

j goes from 1 to d and i goes from 1 to n, and

X = future standard deviation forecast BS1, BS3X3, GARCH20, HIST, and HIST20,

d = number of days, 496 for BS1 and BS3X3 and 506 for GARCH20, HIST, and HIST20, and

n = number of firms trading on a given day d.

3) The average MSE for the 20 day volatility forecast horizon is computed as

\[ \text{AVGMSE} = \frac{1}{d} \sum_j \sum_i \left( \frac{(FUT20_{ij} - X_{ij})^2}{n} \right) \]

For the 40 day forecast horizon we repeat the above three methods using FUT40, BS1, BS3X3, GARCH40, HIST, and HIST40. The calculations are similar for the series based on 80 and 120 day volatility forecast horizons.

C. RESULTS

Applying our methodology to the data yields the following results for analysis. The R-squared statistic measures the percentage of the total variation of the dependent variable explained by the variation in the independent variable. Thus, the higher the R-squared the better the
estimate. As for the intercept, $a$, and the slope, $b$, the optimal outcome consists of a forty-five degree line through the origin, which implies $a = 0$ and $b = 1$, as this represents a perfect forecast. Hence, the closer $a$ is to zero and $b$ is to one the better the forecast. The regression results for the 20, 40, 80 and 120 day forecast horizons are in Tables 5-2, 5-3, 5-4 and 5-5, respectively.

The average MAD results are presented in Tables 5-6, 5-8, 5-10 and 5-12 for the 20, 40, 80 and 120 day forecast intervals, respectively. As these average MADs are calculated from about 500 observations, we also report the standard deviation and the 5, 10, 50, 90 and 95 percent quantiles to get some idea of the distribution. The average MSE results for the 20, 40, 80 and 120 day forecast horizons are in Tables 5-7, 5-9, 5-11 and 5-13, respectively. As with the average MADs, the average MSEs are calculated from about 500 observations so we again report some information on the distribution in the standard deviation and the 5, 10, 50, 90 and 95 percent quantiles.

Let us first compare our conditional forecasts of volatility, specifically GARCH versus the two BS forecasts, BS1 and BS3X3. In our 20 day forecast horizon, GARCH20 outperforms BS1 and BS3X3 using the R-squared criterion as the former has a higher R-squared than the latter. GARCH20 also outperforms BS1 and BS3X3 using the AVGMAD and AVGMSE criteria. In both instances the GARCH AVGMAD and AVGMSE are smaller than the BS implied volatility AVGMAD and AVGMSE. We might also note that the GARCH MAD and MSE for each of the quantiles reported is less than either of its BS counterparts in all comparisons except for the GARCH20 MSE versus BS3X3 and the GARCH40 MAD versus BS1 and BS3X3. In terms of the
slope and the intercept coefficients we cannot distinguish between GARCH20 and BS1 or BS3X3 as none of them are statistically significantly different from perfect forecasts. With these results we conclude GARCH20 forecasts FUT20 better than BS1 or BS3X3. As similar results hold for 40, 80 and 120 day forecast horizons, we conclude that GARCH forecasts of volatility are better than BS implied volatility forecasts.

In contrast to Beckers (1981), we find that BS3X3 outperform BS1 implied volatility forecasts. Our results show that BS3X3 performs as well as, if not better than, BS1 in all four forecast horizons, 20, 40, 80 and 120, in each of the four comparisons, R-squared, regression coefficients, AVGMAD, and AVGMSE.

As we mentioned earlier, we expect the differences between the conditional and unconditional estimates to be more pronounced in the short-term relative to the long-term. With this in mind, let us compare the GARCH estimates with the short-term historical standard deviation estimates. At the 20 day forecast horizon GARCH20 outperforms HIST20 in the average R-squared and the average slope and intercept comparison. That is to say, GARCH20 has a higher R-squared than HIST20 does and its slope and intercept are not statistically significantly different from one and zero, respectively, whereas the slope and intercept of HIST20 are significantly different from one and zero. We also find GARCH20 outperforms HIST20 with a smaller AVGMAD and a smaller AVGMSE. Hence, GARCH20 forecasts volatility in the 20 day forecast horizon better than HIST20.

In both the 40 and 80 day forecast horizons we find the same results as above. GARCH40 and GARCH80 outperform HIST40 and HIST80, respectively,
in forecasting the volatility over the 40 and 80 day horizons, with the exception that the intercept coefficient is not statistically different from zero for HIST80. We conclude from this information that GARCH40 outperforms HIST40 in forecasting volatility over the 40 day forecast horizon and draw the same conclusion for GARCH80 versus HIST80. Only when we reach the longest short-term historical standard deviation forecast are the results indeterminant. In the forecast horizon of 120 days, GARCH120 outperforms HIST120 with the AVGMAD and AVGMSE criterion, however, the R-squared of HIST120 is much larger than the R-squared of GARCH120. As for the forecasting ability in terms of the regression coefficients no significant difference is detected between the two forecasts. Thus, in this case, the evidence does support the expectation that the conditional estimates outperform unconditional estimates with a general trend of this difference declining as the forecast horizon increases in length as well as the period over which the historical standard deviation is calculated.

The results indicate GARCH outperforms the short-term historical standard deviation forecast in the 20, 40 and 80 day forecast horizons. The 120 day forecast horizon does not yield any conclusive results.

Comparing GARCH forecasts with long-term historical, in the 20 day forecast horizon GARCH20 has a higher R-squared than does HIST. There is no significant difference in the regression parameters in the sense that for both GARCH20 and HIST the forecasts are not significantly different from perfect forecasts. GARCH20 also has a smaller AVGMAD and AVGMSE than HIST. As the forecast horizon increases to 40, 80 and 120 days, the R-squared of the GARCH estimates is greater than that of HIST, but the difference declines to .1. The AVGMAD and AVGMSE results shift to favor
the HIST estimates and the regression coefficients in each case yield the same conclusions with respect to the difference from perfect forecasts for both GARCH and HIST. Hence, we conclude GARCH outperforms the long-term historical standard deviation in the very short-term, 20 day, forecast horizon. Even this small distinction declines as the forecast horizon is increased.

We now compare the BS implied volatility estimates to the short-term historical standard deviations. At the 20 day forecast horizon, BS forecasts outperform HIST20 in the R-squared, regression coefficients, and AVGMSE comparisons. In AVGMAD, BS3X3 outperforms HIST20, but HIST20 outperforms BS1. From these results we conclude BS implied volatilities forecast volatility in the 20 day forecast horizon better than HIST20. As we move to the 40 day forecast horizon comparisons, they are inconclusive as the regression coefficient comparison goes to the BS implied volatilities, the R-squared favors BS3X3 over HIST40 but HIST40 to BS1, the AVGMAD comparison favors HIST40 to BS, and finally the AVGMSE results favor BS3X3 to HIST40 and HIST40 to BS1. When we consider the 80 and 120 day forecast intervals, HIST80 and HIST120 outperform the BS implied volatilities in forecasting their respective future horizons in the R-squared, AVGMAD and AVGMSE comparisons. BS is only better than the short-term historical standard deviations at forecasting volatility in the very near term, specifically 20 days. However, the short-term historical standard deviations are better forecasts of volatility than BS implied volatilities in the longer short-term horizons of 80 and 120 days.

The results of comparisons of the BS implied volatilities versus the long-term historical standard deviation forecasts are similar in all four
forecast horizons to the conclusions drawn for the 80 and 120 day forecast horizons above. The BS implied volatilities are outperformed in every interval in the R-squared, AVGMAD, and AVGMSE comparisons. The regression coefficient conclusions are the same for BS and HIST except in the FUT120 comparison where BS regression coefficients are indistinguishable from a perfect forecast and the slope parameter of HIST is significantly different from one. So here we conclude HIST forecasts volatility better than BS implied volatilities.

The extant literature finds BS outperforms historical standard deviation in forecasting volatility. This literature however, uses monthly data. We perform tests similar to ours above on monthly data with a forecast horizon of 20 months and a 20 month historical standard deviation forecast and BS volatility forecasts based on observations from the last day of the month. We find results consistent with the extant literature.

D. CONCLUSION

In this chapter we have compared forecasts of the volatility of equity returns based on three different methods. These three forecasts are obtained: (i) by estimating a GARCH model, (ii) by implying the volatility from the Black-Scholes formula, and (iii) by calculating the simple standard deviation of a historical return series. Our primary focus has been on the first two which are conditional estimates of future volatility while the third is an unconditional estimate.

We have found that GARCH forecasts of volatility outperform BS implied volatility forecasts. Perhaps, the GARCH model forecasts perform
better than the Black-Scholes implied volatilities because (a) the latter requires that the option and equity markets are constantly in equilibrium, and (b) the latter is based on stronger assumptions about the equity returns process. We also find that GARCH forecasts do better than short-term historical standard deviation forecasts. This is particularly true for the very short-term forecast horizon and diminishes as the forecast horizon increases. Finally, we find that historical standard deviation estimates based on daily data forecast volatility better than BS implied volatilities.
Figure 5-1
Sample Calculations of Standard Deviations

<table>
<thead>
<tr>
<th>1/3/78</th>
<th>1/3/83</th>
<th>12/31/84</th>
<th>6/21/85</th>
</tr>
</thead>
<tbody>
<tr>
<td>i=1</td>
<td>i=1265</td>
<td>i=1771</td>
<td>i=1791</td>
</tr>
</tbody>
</table>

The days below our time line represent the period over which we compare forecasts.

The i’s above our time line represent the period over which we collected data from the daily CRSP tapes.

Standard deviation estimate for 20 day forecast horizon for day 1 is based on:

FUT20: days 2 through 21 or equivalently i = 1266 to 1285.

BS1 and BS3X3: day 1 option observations from Berkeley tapes.

HIST: i = 1 to 1264.

HIST20: i = 1245 to 1264.

GARCH20: model estimated on i = 1 to 1264 with conditional forecast made over days 2 through 21 or i = 1266 to 1285.
Table 5-1

GARCH Parameters Averaged over the 506 Estimated Models

<table>
<thead>
<tr>
<th></th>
<th>AVG $\alpha_0$</th>
<th>AVG $\alpha_1$</th>
<th>AVG $\beta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. AT&amp;T</td>
<td>.100</td>
<td>.152</td>
<td>.765</td>
</tr>
<tr>
<td>2. Avon</td>
<td>.404</td>
<td>.077</td>
<td>.771</td>
</tr>
<tr>
<td>3. Boeing</td>
<td>.753</td>
<td>.094</td>
<td>.756</td>
</tr>
<tr>
<td>4. Bally Mfg</td>
<td>1.23</td>
<td>.094</td>
<td>.757</td>
</tr>
<tr>
<td>5. Bethlehem Steel</td>
<td>.814</td>
<td>.139</td>
<td>.646</td>
</tr>
<tr>
<td>6. Citicorp</td>
<td>.634</td>
<td>.088</td>
<td>.746</td>
</tr>
<tr>
<td>7. Control Data</td>
<td>.674</td>
<td>.076</td>
<td>.797</td>
</tr>
<tr>
<td>8. Digital Equipment</td>
<td>.704</td>
<td>.088</td>
<td>.743</td>
</tr>
<tr>
<td>9. Dow Chemical</td>
<td>.989</td>
<td>.084</td>
<td>.657</td>
</tr>
<tr>
<td>10. Eastman Kodak</td>
<td>.326</td>
<td>.083</td>
<td>.786</td>
</tr>
<tr>
<td>11. Exxon</td>
<td>.252</td>
<td>.074</td>
<td>.775</td>
</tr>
<tr>
<td>12. Ford</td>
<td>.328</td>
<td>.099</td>
<td>.839</td>
</tr>
<tr>
<td>13. General Electric</td>
<td>.097</td>
<td>.096</td>
<td>.854</td>
</tr>
<tr>
<td>14. General Motors</td>
<td>.240</td>
<td>.095</td>
<td>.822</td>
</tr>
<tr>
<td>15. Gulf &amp; Western</td>
<td>1.02</td>
<td>.134</td>
<td>.600</td>
</tr>
<tr>
<td>16. Holiday Inns</td>
<td>1.07</td>
<td>.078</td>
<td>.721</td>
</tr>
<tr>
<td>17. Homestake Mining</td>
<td>1.14</td>
<td>.102</td>
<td>.769</td>
</tr>
<tr>
<td>18. IBM</td>
<td>.346</td>
<td>.073</td>
<td>.765</td>
</tr>
<tr>
<td>19. ITT</td>
<td>.663</td>
<td>.066</td>
<td>.652</td>
</tr>
<tr>
<td>20. McDonald's</td>
<td>.400</td>
<td>.079</td>
<td>.756</td>
</tr>
<tr>
<td>21. National Semiconductor</td>
<td>2.31</td>
<td>.071</td>
<td>.713</td>
</tr>
<tr>
<td>22. Occidental Petroleum</td>
<td>.869</td>
<td>.085</td>
<td>.716</td>
</tr>
<tr>
<td>23. RCA</td>
<td>.474</td>
<td>.105</td>
<td>.780</td>
</tr>
<tr>
<td>24. Sears</td>
<td>.349</td>
<td>.103</td>
<td>.779</td>
</tr>
<tr>
<td>25. United Airlines</td>
<td>1.28</td>
<td>.081</td>
<td>.754</td>
</tr>
</tbody>
</table>
Table 5-2
Regression Results for Daily Data
20 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Average r-square</th>
<th>Average Adjusted r-square</th>
<th>Intercept a (t)</th>
<th>Slope b (t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>42.8</td>
<td>40.7</td>
<td>.0015 (.34)</td>
<td>.813 (-.93)</td>
</tr>
<tr>
<td>BS3X3</td>
<td>44.8</td>
<td>42.3</td>
<td>.0017 (.48)</td>
<td>.797 (-1.10)</td>
</tr>
<tr>
<td>GARCH20</td>
<td>45.8</td>
<td>43.4</td>
<td>.0030 (.85)</td>
<td>.765 (-1.40)</td>
</tr>
<tr>
<td>HIST</td>
<td>44.6</td>
<td>42.2</td>
<td>.0033 (.94)</td>
<td>.750 (-1.48)</td>
</tr>
<tr>
<td>HIST20</td>
<td>35.2</td>
<td>32.5</td>
<td>.0076 (2.50)</td>
<td>.586 (-2.57)</td>
</tr>
</tbody>
</table>

* This test is for a significant difference of b from one. At the 95% level with 25 observations those t-values greater than 2.06 or less than -2.06 are statistically significantly different from one.
Table 5-3
Regression Results for Daily Data
40 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Average r-square</th>
<th>Average Adjusted r-square</th>
<th>Intercept a</th>
<th>(t)</th>
<th>Slope b</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>48.4</td>
<td>46.7</td>
<td>.0021</td>
<td>(.49)</td>
<td>.796</td>
<td>(-1.21)</td>
</tr>
<tr>
<td>BS3X3</td>
<td>51.0</td>
<td>48.7</td>
<td>.0021</td>
<td>(.67)</td>
<td>.786</td>
<td>(-1.37)</td>
</tr>
<tr>
<td>GARCH40</td>
<td>53.4</td>
<td>51.3</td>
<td>.0031</td>
<td>(1.05)</td>
<td>.764</td>
<td>(-1.68)</td>
</tr>
<tr>
<td>HIST</td>
<td>52.3</td>
<td>50.2</td>
<td>.0034</td>
<td>(1.15)</td>
<td>.752</td>
<td>(-1.75)</td>
</tr>
<tr>
<td>HIST40</td>
<td>48.6</td>
<td>46.4</td>
<td>.0056</td>
<td>(2.11)</td>
<td>.687</td>
<td>(-2.25)</td>
</tr>
</tbody>
</table>

* This test is for a significant difference of b from one. At the 95% level with 25 observations those t-values greater than 2.06 or less than -2.06 are statistically significantly different from one.
Table 5-4

Regression Results for Daily Data
80 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Average r-square</th>
<th>Average Adjusted r-square</th>
<th>Intercept a (t)</th>
<th>Slope b (t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>48.6</td>
<td>46.7</td>
<td>.0026</td>
<td>(.65) .768</td>
</tr>
<tr>
<td>BS3X3</td>
<td>52.6</td>
<td>50.4</td>
<td>.0025</td>
<td>(.79) .763</td>
</tr>
<tr>
<td>GARCH80</td>
<td>57.0</td>
<td>55.1</td>
<td>.0033</td>
<td>(1.25) .751</td>
</tr>
<tr>
<td>HIST</td>
<td>56.4</td>
<td>54.5</td>
<td>.0035</td>
<td>(1.30) .747</td>
</tr>
<tr>
<td>HIST80</td>
<td>55.4</td>
<td>53.5</td>
<td>.0043</td>
<td>(1.71) .730</td>
</tr>
</tbody>
</table>

* This test is for a significant difference of b from one. At the 95% level with 25 observations those t-values greater than 2.06 or less than -2.06 are statistically significantly different from one.
Table 5-5
Regression Results for Daily Data
120 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Average r-square</th>
<th>Average Adjusted r-square</th>
<th>Intercept $a$</th>
<th>$b$ * (t)</th>
<th>Slope $b$ (t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>49.1</td>
<td>47.1</td>
<td>.0026</td>
<td>(.76)</td>
<td>.764 (-1.49)</td>
</tr>
<tr>
<td>BS3X3</td>
<td>53.1</td>
<td>50.9</td>
<td>.0027</td>
<td>(.94)</td>
<td>.752 (-1.81)</td>
</tr>
<tr>
<td>GARCH120</td>
<td>58.6</td>
<td>56.8</td>
<td>.0036</td>
<td>(1.40)</td>
<td>.732 (-2.20)</td>
</tr>
<tr>
<td>HIST</td>
<td>58.5</td>
<td>56.7</td>
<td>.0036</td>
<td>(1.43)</td>
<td>.735 (-2.15)</td>
</tr>
<tr>
<td>HIST120</td>
<td>64.5</td>
<td>62.9</td>
<td>.0031</td>
<td>(1.39)</td>
<td>.769 (-2.12)</td>
</tr>
</tbody>
</table>

* This test is for a significant difference of $b$ from one. At the 95% level with 25 observations those $t$-values greater than 2.06 or less than -2.06 are statistically significantly different from one.
Table 5-6*  
Mean Absolute Deviation (MAD) Results for Daily Data  
20 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean Deviation</th>
<th>Quantiles</th>
<th>5%</th>
<th>10%</th>
<th>50%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>4.75</td>
<td>1.08</td>
<td>3.26</td>
<td>3.48</td>
<td>4.52</td>
<td>6.14</td>
<td>6.67</td>
</tr>
<tr>
<td>BS3X3</td>
<td>4.64</td>
<td>.96</td>
<td>3.23</td>
<td>3.49</td>
<td>4.57</td>
<td>5.90</td>
<td>6.31</td>
</tr>
<tr>
<td>GARCH20</td>
<td>4.36</td>
<td>.88</td>
<td>3.17</td>
<td>3.31</td>
<td>4.21</td>
<td>5.65</td>
<td>6.05</td>
</tr>
<tr>
<td>HIST</td>
<td>4.43</td>
<td>.87</td>
<td>3.21</td>
<td>3.35</td>
<td>4.35</td>
<td>5.63</td>
<td>6.05</td>
</tr>
<tr>
<td>HIST20</td>
<td>4.71</td>
<td>1.30</td>
<td>3.00</td>
<td>3.41</td>
<td>4.50</td>
<td>6.41</td>
<td>7.68</td>
</tr>
</tbody>
</table>

*All values in table are X10^{-3}.
Table 5-7*

Mean Square Error (MSE) Results for Daily Data
20 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>5%</th>
<th>10%</th>
<th>50%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>4.17</td>
<td>3.97</td>
<td>1.68</td>
<td>2.01</td>
<td>3.19</td>
<td>6.04</td>
<td>7.95</td>
</tr>
<tr>
<td>BS3X3</td>
<td>4.12</td>
<td>3.59</td>
<td>1.66</td>
<td>1.93</td>
<td>3.26</td>
<td>5.57</td>
<td>7.26</td>
</tr>
<tr>
<td>GARCH20</td>
<td>3.58</td>
<td>3.12</td>
<td>1.55</td>
<td>1.75</td>
<td>2.78</td>
<td>5.18</td>
<td>7.92</td>
</tr>
<tr>
<td>HIST</td>
<td>3.67</td>
<td>3.11</td>
<td>1.60</td>
<td>1.81</td>
<td>2.90</td>
<td>5.06</td>
<td>7.88</td>
</tr>
<tr>
<td>HIST20</td>
<td>4.96</td>
<td>4.32</td>
<td>1.66</td>
<td>2.04</td>
<td>3.64</td>
<td>8.25</td>
<td>18.13</td>
</tr>
</tbody>
</table>

*All values in table are X10^{-5}.
Table 5-8*

Mean Absolute Deviation (MAD) Results for Daily Data
40 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Standard Deviation</th>
<th>Quantiles</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>50%</td>
<td>90%</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS1</td>
<td>4.20</td>
<td>2.59</td>
<td>2.83</td>
<td>4.27</td>
<td>5.42</td>
<td>5.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS3X3</td>
<td>4.09</td>
<td>2.59</td>
<td>2.84</td>
<td>4.11</td>
<td>5.24</td>
<td>5.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH40</td>
<td>3.72</td>
<td>2.66</td>
<td>2.81</td>
<td>3.72</td>
<td>4.63</td>
<td>4.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIST</td>
<td>3.73</td>
<td>2.64</td>
<td>2.78</td>
<td>3.74</td>
<td>4.58</td>
<td>4.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIST40</td>
<td>3.79</td>
<td>2.63</td>
<td>2.74</td>
<td>3.54</td>
<td>5.43</td>
<td>5.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*All values in table are X10^{-3}.
Table 5-9*

Mean Square Error (MSE) Results for Daily Data
40 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean Deviation</th>
<th>Quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>BS1</td>
<td>3.34</td>
<td>2.48</td>
</tr>
<tr>
<td>BS3X3</td>
<td>3.24</td>
<td>2.17</td>
</tr>
<tr>
<td>GARCH40</td>
<td>2.61</td>
<td>1.76</td>
</tr>
<tr>
<td>HIST</td>
<td>2.64</td>
<td>1.76</td>
</tr>
<tr>
<td>HIST40</td>
<td>3.31</td>
<td>2.64</td>
</tr>
</tbody>
</table>

*All values in table are X10^-5.
Table 5-10*

Mean Absolute Deviation (MAD) Results for Daily Data
80 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean Deviation</th>
<th>5%</th>
<th>10%</th>
<th>50%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>4.01</td>
<td>2.52</td>
<td>2.82</td>
<td>4.14</td>
<td>4.99</td>
<td>5.27</td>
</tr>
<tr>
<td>BS3X3</td>
<td>3.87</td>
<td>2.43</td>
<td>2.75</td>
<td>3.96</td>
<td>4.78</td>
<td>5.12</td>
</tr>
<tr>
<td>GARCH80</td>
<td>3.33</td>
<td>2.35</td>
<td>2.47</td>
<td>3.25</td>
<td>4.25</td>
<td>4.37</td>
</tr>
<tr>
<td>HIST</td>
<td>3.29</td>
<td>2.43</td>
<td>2.50</td>
<td>3.28</td>
<td>4.19</td>
<td>4.32</td>
</tr>
<tr>
<td>HIST80</td>
<td>3.38</td>
<td>2.21</td>
<td>2.49</td>
<td>3.10</td>
<td>4.32</td>
<td>5.55</td>
</tr>
</tbody>
</table>

*All values in table are X10^-3.
Table 5-11*

Mean Square Error (MSE) Results for Daily Data
80 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Standard Mean Deviation</th>
<th>Quantiles</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5%</td>
<td>10%</td>
<td>50%</td>
<td>90%</td>
<td>95%</td>
</tr>
<tr>
<td>BS1</td>
<td>2.97</td>
<td>1.46</td>
<td>1.14</td>
<td>1.33</td>
<td>2.73</td>
<td>5.19</td>
</tr>
<tr>
<td>BS3X3</td>
<td>2.83</td>
<td>1.35</td>
<td>1.03</td>
<td>1.25</td>
<td>2.53</td>
<td>4.98</td>
</tr>
<tr>
<td>GARCH80</td>
<td>2.07</td>
<td>.90</td>
<td>.89</td>
<td>.99</td>
<td>1.88</td>
<td>3.47</td>
</tr>
<tr>
<td>HIST</td>
<td>2.05</td>
<td>.89</td>
<td>.93</td>
<td>1.03</td>
<td>1.90</td>
<td>3.45</td>
</tr>
<tr>
<td>HIST80</td>
<td>2.36</td>
<td>1.19</td>
<td>.92</td>
<td>1.06</td>
<td>2.01</td>
<td>4.07</td>
</tr>
</tbody>
</table>

*All values in table are X10^{-5}.
Table 5-12*
Mean Absolute Deviation (MAD) Results for Daily Data
120 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean Deviation</th>
<th>5%</th>
<th>10%</th>
<th>50%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS1</td>
<td>4.01</td>
<td>2.57</td>
<td>2.85</td>
<td>3.93</td>
<td>5.28</td>
<td>5.64</td>
</tr>
<tr>
<td>BS3X3</td>
<td>3.88</td>
<td>2.53</td>
<td>2.83</td>
<td>3.83</td>
<td>5.03</td>
<td>5.36</td>
</tr>
<tr>
<td>GARCH120</td>
<td>3.25</td>
<td>2.34</td>
<td>2.50</td>
<td>3.32</td>
<td>4.10</td>
<td>4.20</td>
</tr>
<tr>
<td>HIST</td>
<td>3.17</td>
<td>2.28</td>
<td>2.41</td>
<td>3.17</td>
<td>3.97</td>
<td>4.18</td>
</tr>
<tr>
<td>HIST120</td>
<td>3.32</td>
<td>2.00</td>
<td>2.14</td>
<td>3.11</td>
<td>5.61</td>
<td>6.10</td>
</tr>
</tbody>
</table>

*All values in table are X10^{-3}.
Table 5-13*

Mean Square Error (MSE) Results for Daily Data
120 Day Forecast Horizon

<table>
<thead>
<tr>
<th>FUT vs.</th>
<th>Mean Deviation</th>
<th>Standard Quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>BS1</td>
<td>2.85</td>
<td>1.16</td>
</tr>
<tr>
<td>BS3X3</td>
<td>2.73</td>
<td>1.06</td>
</tr>
<tr>
<td>GARCH120</td>
<td>1.92</td>
<td>.61</td>
</tr>
<tr>
<td>HIST</td>
<td>1.86</td>
<td>.58</td>
</tr>
<tr>
<td>HIST120</td>
<td>2.19</td>
<td>1.21</td>
</tr>
</tbody>
</table>

*All values in table are X10^-5.
CHAPTER 6

SUMMARY

This dissertation considers the information on risk available from options data. In one study, we use observations from the Chicago Board Options Exchange to obtain an estimate of a constant proportional risk aversion parameter. The second study compares the conditional forecast of volatility on an underlying security implied from the Black-Scholes European call option pricing formula using observed call prices with the conditional forecast of volatility using generalized autoregressive conditional heteroscedasticity.

In Chapter 2, we briefly review the development of option pricing. We discuss the Black-Scholes European call option pricing model, its contribution to the literature, and the literature previous to Black-Scholes. Then we discuss the development of the equivalent martingale measure. The equivalent martingale measure framework is used in Chapter 3 to derive a valuation from which we imply an estimate of the constant proportional risk aversion parameter using call option data. Hence, we can avoid the problems in measuring consumption data that exist in all of the previous studies on estimating the risk aversion parameter by using options data. We find an estimate of constant proportional risk aversion in the range of .20-.28. This implies that the investors are only slightly risk averse.

Chapter 4 is a review of the literature on forecasting the volatility of returns on the asset underlying an option. We discuss the three previous forecasts used. One estimate is unconditional and is
calculated as a simple standard deviation of the historical returns. There are two conditional estimates. One is based on options data, the volatility implied from Black-Scholes (BS), and the other is based on historical returns data, generalized autoregressive conditional heteroscedasticity (GARCH). The extant literature has compared only a conditional estimate with an unconditional estimate. In Chapter 5, we compare the two conditional estimates. We describe the methodology used to obtain our forecasts and that used to compare them. We find that our GARCH forecasts of volatility are better than BS.
REFERENCES


VITA
KENNETH S. BARTUNEK

Department of Finance 336 W. Parker Blvd. #3
2163 CEBA  Baton Rouge, LA 70808
College of Business Administration (504) 769-2352
Louisiana State University
Baton Rouge, LA 70803
(504) 388-6291

Education

Louisiana State University  Finance 1987- PhD (1991)
University of Michigan-Flint Mathematics 1983-87 BS
United States Naval Academy  1983

Dissertation

Topic: "Implied Volatility and Risk Preference from Option Prices"

Presentations at Professional Meetings

Presented dissertation topic at the Financial Management
Association’s Doctoral Student Seminar in Orlando, October, 1990.

Presented "Estimating the Variance of the Underlying Asset of an
Option: Conditional versus Unconditional" at the Midwest Finance
Association meeting in Saint Louis, April, 1991.

Teaching Experience

Intro. to Corp. Fin.  Louisiana State University  Spring 1991

Experience

Michigan Bell Telephone  Economic Analysis Department Intern 1987
   Summer
University of Michigan-Flint  Chemistry Research Assistant 1984-85
   Part-time
Honors

Who's Who of Emerging Leaders in America 1991
Beta Gamma Sigma 1991
Alumni Federation Fellowship Louisiana State University 1987-
Graduate summa cum laude University of Michigan-Flint 1987
James B. Angell Scholar University of Michigan-Flint 1985-87
Outstanding Mathematics Student University of Michigan-Flint 1985-87
Student
Branstrom Scholar University of Michigan-Flint 1984
Outstanding Freshman Chemist University of Michigan-Flint 1984
Award
GM-Best of Class-1983 1983
American Legion Award 1983
Bausch and Lomb Science Award 1983

Organizations

Mathematical Association of America 1985-
New Orleans Museum of Art 1988-
Financial Management Association 1989-
Eastern Finance Association 1990-
Midwest Finance Association 1990-
International Platform Association 1991-

Personal

Enjoy physical fitness activities.
Travels:
Have been to forty-nine of the fifty states including Alaska and Hawaii.
Have journeyed to twenty-one countries since 1986:
Austria Holland Mexico
Canada Hong Kong Nepal
Denmark India Scotland
Egypt Ireland Spain
England Italy Sweden
France Kenya Switzerland
Germany Macao Thailand

References

Furnished upon request.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Kenneth S. Bartunek

Major Field: Business Administration (Finance)

Title of Dissertation: Implied Volatility and Risk Preference from Option Prices

Approved:

[Signatures]

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

May 3, 1991