8-2010

COMPARISON OF IMPLIED VOLATILITY APPROXIMATIONS USING 'NEAREST-TO-THE-MONEY' OPTION PREMIUMS

Joseph Ewing
Clemson University, jewing@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses
Part of the Economics Commons

Recommended Citation
https://tigerprints.clemson.edu/all_theses/868

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
COMPARISON OF IMPLIED VOLATILITY APPROXIMATIONS USING “NEAREST-TO-THE-MONEY” OPTION PREMIUMS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Applied Economics and Statistics

by
Joseph Alexander Ewing
August 2010

Accepted by:
Dr. Olga Isengildina-Massa, Committee Chair
Dr. William Bridges, Jr.
Dr. Charles Curtis, Jr.
ABSTRACT

Implied volatility provides information which is useful for not only investors, but farmers, producers, manufacturers and corporations. These market participants use implied volatility as a measure of price risk for hedging and speculation decisions. Because volatility is a constantly changing variable, there needs to be a simple and quick way to extract its value from the Black-Scholes model. Unfortunately, there is no closed form solution for the extraction of the implied volatility variable; therefore its value must be approximated. This study investigated the relative accuracy of six methods for approximating Black-Scholes implied volatility developed by Curtis and Carriker, Brenner and Subrahmaniam, Chargoy-Corona and Ibarra-Valdez, Bharadia et al., Li (2005) and Corrado and Miller. Each of these methods were tested and analyzed for accuracy using nearest to the money options over two data sets, corn and live cattle, spanning contract years 1989 to 2008 and 1986 to 2008, respectively. This study focuses on accuracy for nearest-to-the-money options because the majority of traded options are concentrated at or near-the-money and several of the approximations were developed for at-the-money options.

Rather than following only the traditional measures of testing approximations for accuracy, this study considered several alternative ways for testing accuracy. In addition to analyzing mean errors and mean percent errors, other moments of the error distributions such as variance and skewness were analyzed. Beyond this, measures of goodness of fit, determined through an adjusted $R^2$, and accuracy over observed changes
in market variables, such as moneyness, time to maturity and interest rates, were analyzed.

The results were divided into three distinct groups, with the first group comprised of only the Corrado and Miller approximation. This method was clearly the most accurate, followed by Bharadia et al. and Li (2005) in the second group and finally the Curtis and Carriker, Brenner and Subrahmanyam, Chargoy-Corona and Ibarra-Valdez methods in the third group.
DEDICATION

I would like to dedicate my Thesis to all the friends I have made over the past two years at Clemson University. Each one of them has helped me through the good times and the bad. I look forward to continuing these relationships into the future.
ACKNOWLEDGMENTS

I would like to thank Dr. Patrick Gerard who has given me great guidance through many aspects over the last year. His support and willingness to answer my never ending questions is appreciated more than he knows.

I would also like to extend my gratitude to my committee for their guidance through the entire process of accomplishing this Thesis. I have learned lessons from them that I will take with me as I continue in my academic journey.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
</tbody>
</table>

## CHAPTER

I. INTRODUCTION ......................................................... 1

II. LITERATURE REVIEW ............................................... 6
   Approximations ................................................. 6
   Accuracy Analysis .......................................... 11
   Other contributions ................................. 14

III. DATA ................................................................. 20

IV. METHODS ............................................................ 31
   Error Histograms ............................................. 31
   Adjusted $R^2$ ............................................... 32
   Changes in Error over Observed Market Variables 33

V. RESULTS ............................................................. 37
   Error Histograms ............................................. 37
   Adjusted $R^2$ ............................................... 43
   Changes in Error over Observed Market Variables 44

VI. SUMMARY AND CONCLUSIONS .............................. 68
Table of Contents (Continued)

APPENDICES ............................................................................................................... 72

| A: | SAS Code Used to Merge Futures with Calls/Puts | 73 |
| B: | SAS Code Used to Merge Calls and Puts | 75 |
| C: | SAS Code Used to find a Benchmark Black-Scholes Implied Volatility for Call options | 76 |
| D: | SAS Code Used to find a Benchmark Black-Scholes Implied Volatility for Put Options | 78 |

REFERENCES .............................................................................................................. 80
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive Statistics for Corn</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive Statistics for Live Cattle</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Analysis of Variance, Corn Calls Means, Moneyness</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>Corn Calls Means, Moneyness, LS Means Differences</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>Analysis of Variance, Live Cattle Calls Means, Moneyness</td>
<td>51</td>
</tr>
<tr>
<td>6</td>
<td>Live Cattle Calls Means, Moneyness, LS Means Differences</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>Analysis of Variance, Corn Calls Variance, Moneyness</td>
<td>53</td>
</tr>
<tr>
<td>8</td>
<td>Analysis of Variance, Live Cattle Calls Variance, Moneyness</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>Analysis of Variance, Corn Calls Means, Time to Maturity</td>
<td>57</td>
</tr>
<tr>
<td>10</td>
<td>Corn Calls Means, Time to Maturity, LS Means Differences</td>
<td>57</td>
</tr>
<tr>
<td>11</td>
<td>Analysis of Variance, Live Cattle Calls Means, Time to Maturity</td>
<td>59</td>
</tr>
<tr>
<td>12</td>
<td>Analysis of Variance, Corn Calls Variance, Time to Maturity</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>Effects Test, Corn Calls Variance, Time to Maturity</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>Corn Calls Variance, Time to Maturity, LS Means Differences</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>Analysis of Variance, Live Cattle Calls Variance, Time to Maturity</td>
<td>61</td>
</tr>
<tr>
<td>16</td>
<td>Analysis of Variance, Corn Calls Means, Interest Rate</td>
<td>64</td>
</tr>
<tr>
<td>17</td>
<td>Analysis of Variance, Live Cattle Calls Means, Interest Rate</td>
<td>64</td>
</tr>
<tr>
<td>18</td>
<td>Analysis of Variance, Corn Calls Variances, Interest Rate</td>
<td>65</td>
</tr>
<tr>
<td>19</td>
<td>Analysis of Variance, Live Cattle Calls Variances, Interest Rate</td>
<td>65</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Black-Scholes Implied Volatility</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>Corn Calls Error Histograms</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Live Cattle Calls Error Histograms</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>Corn Calls Approximations Percent Error</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>Live Cattle Calls Approximations Percent Error</td>
<td>42</td>
</tr>
<tr>
<td>6</td>
<td>Corn Calls Percent Error and Moneyness</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>Live Cattle Calls Percent Error and Moneyness</td>
<td>47</td>
</tr>
<tr>
<td>8</td>
<td>Corn Calls Percent Error and Time to Maturity</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>Live Cattle Calls Percent Error and Time to Maturity</td>
<td>55</td>
</tr>
<tr>
<td>10</td>
<td>Corn Calls Percent Errors and Interest Rate</td>
<td>62</td>
</tr>
<tr>
<td>11</td>
<td>Live Cattle Calls Percent Errors and Interest Rate</td>
<td>63</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The ability to correctly determine price risks and appropriately make investment decisions is fundamental for successful market trading. From Wall Street investors to average American farmers there is a need to understand risk, whether for pure speculation or to assist hedging decisions. In order to do this, a reliable measure of price risk, or a measure of the uncertainty in future price movements, must be identified (Hull). While numerous measures of risk are available, implied volatility stands out as one of the best measures to determine price risk. For example, in their analysis of 93 studies of volatility forecasting models, Poon and Granger (2003) found that implied standard deviations, or implied volatility methods, provide the best forecast of risk (volatility). This is shown by the result that of 34 studies, 26 or 76% indicate that implied volatility models were better at forecasting volatility than historical volatility models when compared directly (Poon and Granger). Implied volatility is the market’s expectation of volatility over the life of an option, which is used for investment decisions (Poon and Granger). This measure of risk is used in a variety of investment decisions and is found through volatility implied from option pricing models.

The most widely used option pricing model was developed by Fisher Black and Myron Scholes (1973). The Black-Scholes model was one of the first models to price European equity option contracts, defined as the right to buy (sell) an asset at a certain price on a certain date, and it continues to be the industry standard today. The Black-
Scholes model describes the relationship between the stock option’s call premium and several market variables:

\[
C = MN(d_1) - (Xe^{-\tau r}N(d_2)),
\]

\[
d_1 = \frac{\ln \left( \frac{M}{X} \right) + (r + \frac{\sigma^2}{2})\tau}{\sigma \sqrt{\tau}}, \quad d_2 = \frac{\ln \left( \frac{M}{X} \right) + (r - \frac{\sigma^2}{2})\tau}{\sigma \sqrt{\tau}}
\]

Where, \(C\) is the call premium,

\(N\) is the cumulative normal distribution function

\(M\) is the settle price of the underlying asset,

\(X\) is the option strike price,

\(r\) is the daily interest rate,

\(\tau\) is time to maturity, \(\tau = [(T-t)/365]\),

\(\sigma\) is implied volatility.

While the model is developed for pricing options, it is most often used for calculating implied volatility because volatility is the only unobservable component of this model. Each of the above variables, with the exception of implied volatility can be put into the Black-Scholes model to derive the volatility implied by the market using a backward induction technique (Poon and Granger). Black and Scholes first constructed this formula to calculate equity option premiums for common stocks and bonds, widely used by corporations and speculators.

Stemming from the original formula presented in 1973, Fisher Black extended it to compute option prices for underlying futures contracts in 1976. This development extended the use of this formula to a much larger pool of commodity options contracts.
widely used for the purpose of hedging. Black’s formula, comprised of the same inputs, follows the spot-futures parity condition, which replaces the original discounted spot price with a futures price, $S$, or $S=Me^{rt}$ (CMIV).

\[ C = e^{-rt} [SN(d_1) - XN(d_2)], \]  
\[ d_1 = \frac{\ln \left( \frac{S}{X} \right) + \frac{\sigma^2 T}{2} }{\sigma \sqrt{T}}, \quad d_2 = \frac{\ln \left( \frac{S}{X} \right) - \frac{\sigma^2 T}{2} }{\sigma \sqrt{T}} \]

With the majority of hedging decisions made using futures contracts, Black’s formula provides hedging guidance for producers, distributors and users of commodities, in addition to corporations (Black).

Unfortunately, Black’s formula (2) is a nonlinear function which has no closed form solution for implied volatility. Therefore, an iterative process must be performed to calculate implied volatility. This is done by taking each observable variable and solving to find the volatility value associated with the zero difference between a predicted call premium and the actual call premium. Doing this is often tedious, requiring the use of sophisticated statistical software, and cannot be done quickly through the use of simple calculations in a spreadsheet. The utility of implied volatility as a measure of price risk and the difficulty of solving the original formula for implied volatility has motivated extensive research and attempts to find an accurate approximation. Rather than the tedious iterative process, these approximations of implied volatility can be easily and quickly calculated in a spreadsheet form.

There are two main groups of approximations; the first group is comprised of approximations which make the starting assumption that the options are exactly at-the-
money, $S = Xe^{-rT}$. Although this assumption greatly simplifies the Black-Scholes model it is rarely the case that options will be exactly at-the-money. Several formulas analyzed in this study like the Direct Implied Volatility Estimate, the Brenner and Subrahmanyam method, and the Chargoy-Carona Ibarra-Valdez method, starts with this assumption. Other methods considered in this study, which allow for strike prices to vary, are the Corrado-Miller method, the Bharadia et al. method, and the method provided by Li (2005)

Although each approximation method is tested for accuracy individually, they have yet to be fully tested for accuracy against vast market data in comparison to an iterated, or benchmark, Black-Scholes implied volatility value. When testing approximation accuracy individually, each method has unique assumptions and limitations. The limitations among the methods include: testing accuracy using different benchmarks; as well as accuracy test using both real and hypothesized option values. Some tests only use at-the-money options (Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Carona Ibarra-Valdez), while others consider options that vary across strike prices (Corrado-Miller, Bharadia et al., and Li (2005)). Also, when testing accuracy, only select methods are analyzed together, rather than a comprehensive study of several approximation methods. Finally, all of these methods for testing accuracy are limited by primary analysis using mean percent and raw errors. These limitations show why these studies are not directly able to be compared. Hence, the goal of this study is to analyze six approximation methods and test their relative accuracy over two extensive real market data sets; using a single benchmark or Black-Scholes implied
volatility. The data used in this study is comprised of daily, nearest-to-the-money, December call and put options for corn data from November 24th 1989 through November 19th 2008 and live cattle data from March 27th 1986 through November 28th 2008.

Traditional measures of accuracy are primarily limited to analysis of mean percent and raw errors. Stephen Figlewski (2001) notes “The statistical properties of a sample mean make it a very inaccurate estimate of the true mean;” therefore, this study considers additional moments and measures for testing approximation accuracy. These include: analysis of mean percent and raw errors, variance and skewness in errors, an adjusted $R^2$ value for goodness of fit, and accuracy measures over changes in the observed variables time to maturity, $\tau$, interest rates, $r$, and moneyness, $(S/X)$. These methods go beyond traditional measures of accuracy to ensure robust results.

For the first time, this study takes six of the best methods for approximating implied volatility and tests the accuracy of these methods against real market data to determine which method is most accurate and how it performs given changes in observed variables. This study will provide farmers, producers, manufacturers and even speculators with the most accurate method for approximating volatility when determining hedging strategies. Next, a thorough review of each method and tests for accuracy are presented, along with a review of other contributing literature. From there, a discussion of the data and methods used to conduct this study is provided, followed by the results.
CHAPTER TWO
LITERATURE REVIEW

The six approximations tested and presented here include methods by Curtis and Carriker; Brenner and Subrahmanyam; Corrado and Miller; Bharadia, Chrstitofides, and Salkin; Li (2005); and Chargoy-Corona and Ibarra-Valdez. This chapter also describes other approximation methods and relevant studies.

Approximations

The first approximation method included in this study is the Direct Implied Volatility Estimate, or DIVE (Curtis and Carriker). In 1988 Curtis and Carriker proposed a non-iterative method which easily approximates implied volatility for at-the-money options \( S = Xe^{-rT} \). Black’s formula, given the at-the-money assumption, is simplified to:

\[
C = S\left[ N\left( \sigma \sqrt{r} / 2 \right) - N\left( \sigma \sqrt{r} / 2 \right) \right] = S\left( 2N\left( \sigma \sqrt{r} / 2 \right) \right) - 1
\] (3)

This is then solved for,

\[
\sigma = \frac{(2/\sqrt{r}) \varphi ((C + S)/2S)}{}
\] (4)

Where \( \varphi = N^{-1} \).

The result is an approximated implied volatility for a call option on an underlying futures contract. Curtis and Carriker take this approximation along with the approximated implied volatility from a put option and average the two to arrive at the Direct Implied Volatility Estimate. The main limitation of Direct Implied Volatility Estimate is that the approximation assumes the options are exactly at-the-money. As
options get further away from being exactly at-the-money this approximation method becomes increasingly less accurate.

Later in 1988, Brenner and Subrahmanyam provide another simplified approximation of the implied volatility calculation. Similarly this approximation method assumed options to be at-the-money, \( S = X e^{-rt} \), for European call options. Brenner and Subrahmanyam use a quadratic expansion of the standard normal distribution of \( d_1 \) to yield:

\[
\sigma \approx \sqrt{\frac{2 \pi}{\tau}} \frac{C}{S}
\]

The authors suggest that there might be “nontrivial estimation errors when the option is not exactly at-the-money” and that taking the straddle, or an average of a put and a call premium; will improve the accuracy of the approximation (Brenner and Subrahmanyam). Again, this model is limited by the fact that it relies on the assumption that futures prices are equal to discounted strike price (at-the-money). This is important to note because this assumption motivated several other approximation methods which use the Brenner and Subrahmanyam method as a starting point, then go further to calculate a method for options where futures price does not equal the discounted strike price.

In 1995, Bharadia et al. developed their approximation under the assumption that options are not always strictly at-the-money. This was the first approximation method which was not limited by the at-the-money assumption. The authors base their derivation on a linear approximation of the cumulative normal distribution, and then use this
approximation to find the parameters $d_1$ and $d_2$. These parameters inserted into equation (2) are then solved for implied volatility. This approach is summarized as:

$$\sigma \approx \sqrt{\frac{2\pi}{\tau} \frac{C-S-K}{2} \frac{S}{S-K}}$$

(6)

Where $K$ is the discounted strike price, $K = Xe^{-rt}$

An advantage of this formula is the improved accuracy of the approximation when options are not exactly at-the-money.

In 1996 Corrado and Miller extended the Brenner and Subrahmanyam method to approximate near-the-money, rather than exactly at-the-money options. The authors follow the same quadratic approximation of the standard normal probabilities, which reduces to the original formula, (5), as calculated by Brenner and Subrahmanyam. It is here that the authors simplify this quadratic formula to accommodate options that are “in the neighborhood of where the stock price is equal to the discounted strike price” (Corrado and Miller). The improvement to the quadratic formula simplifies to:

$$\sigma \approx \sqrt{\frac{2\pi}{\tau} \frac{1}{S+K} \left[ C - \frac{S-K}{2} + \sqrt{\left( C - \frac{S-K}{2} \right)^2 - \frac{(S-K)^2}{\pi}} \right]}$$

(7)

This improved quadratic formula to compute implied standard deviation uses not only discounted strike prices, but also discounted futures prices; represented as $K = Xe^{-rt}, S = Se^{-rt}$.

The next approximation method provided by Li in 2005 follows the progression of formulas starting with Brenner and Subrahmanyam then to Bharadia et al. and finally Corrado and Miller. When options are near-the-money, Li (2005) provides an
improvement on the Brenner and Subrahmanyam formula by using a Taylor series expansion to the third order and substituting the expansions into the cumulative distribution functions; resulting in:

\[
\sigma \approx \frac{2\sqrt{2}}{\sqrt{\tau}} z - \frac{1}{\sqrt{\tau}} \sqrt{8z^2 - \frac{6\alpha}{\sqrt{2}}}
\]

(8)

Where \( z = \cos \left[ \frac{1}{3} \cos^{-1} \left( \frac{3\alpha}{\sqrt{3\eta}} \right) \right] \) and \( \alpha = \frac{\sqrt{2\pi\epsilon}}{S} \) (Li).

For options that are deeper in or out-of-the-money Li (2005) provides an alternative formula, which includes a variable to weigh the moneyness of an option (Li (2005)); \( \eta = \frac{K}{S} \), where \( \eta = 1 \) represents an at-the-money option, \( \eta > 1 \) represents an out-of-the-money option and \( \eta < 1 \) represents an in-the-money option. If \( \sigma \ll \sqrt{\frac{\left| \eta - 1 \right|}{\tau}} \), where “\( \ll \)” means “far less than” and \( \alpha = \frac{\sqrt{2\pi\epsilon}}{1+\eta} \left[ \frac{2\epsilon}{S} + \eta - 1 \right] \), then implied volatility can be approximated as:

\[
\alpha + \sqrt{\alpha^2 - \frac{4(\eta-1)^2}{1+\eta}} \quad \sigma \approx \frac{1+\eta}{2\sqrt{\tau}}
\]

(9)

Note that this formula reduces to the Brenner and Subrahmanyam formula (5) when \( \eta = 1 \). Li (2005) then presents another variable to combine the two formulas. He defines \( \rho = \frac{|\eta - 1|}{\epsilon^2} = \frac{|K-S|S}{\epsilon^2} \) then provides a framework for selecting an appropriate formula. If \( \rho > 1.4 \) formula (9) should be used, and if \( \rho \leq 1.4 \) formula (8) should be used. The primary advantage of Li (2005)’s method is his consideration of the impact moneyness has on implied volatility. Although Li (2005) analyses his model in comparison to
The accuracy of the results is limited by the use of hypothesized option premiums.

The authors of the next and most recent approximation method have a different perspective of the Black-Scholes formula, and approach the extraction of implied volatility from a new angle. The article “A Note on Black-Sholes Implied Volatility” was published in *Physica A*, where the authors Chargoy-Corona and Ibarra-Valdez chose to approach the approximation of implied volatility from a mathematical framework. They employ the Galois Theory to obtain a closed form solution for approximating implied volatility. (Chargoy-Corona and Ibarra-Valdez)

Although the authors begin their approximation from an alternative mindset, they also start with an assumption that options are at-the-money, or as they define it “zero-log-moneyness,” where $S = X e^{-rT}$. Here it is noted that the standard Black-Scholes formula simplifies to:

$$C = S \left[ N \left( \frac{\sigma \sqrt{T}}{2} \right) - N \left( - \frac{\sigma \sqrt{T}}{2} \right) \right]$$

(10)

From this simplified Black-Scholes formula, the authors use the Galois Theory to reduce the number of variables. By doing so, they derive an asymptotic formula for Black-Scholes which is used to define their approximated option value:

$$\sigma = \left( \frac{2}{\sqrt{T}} \right) \Phi \left( \frac{C e^{-rT} + X}{2X} \right)$$

(11)

Note that this formula makes the assumption of “zero-log-moneyness” options, or where the option is exactly at-the-money. This assumption presents the same limitation as previous methods, where the authors only consider options which are at-the-money.
Accuracy Analysis

Most studies reviewed in the first part of this chapter that derive a method for approximating implied volatility also provide a measure of the accuracy of their model. This section discusses the tests of accuracy applied in the previous studies as well as their limitations, followed by suggested improvements.

Curtis and Carriker used two strategies to analyze the Direct Implied Volatility Estimate. First is analysis of raw and mean errors between the averages of put and call approximated volatilities and average iterated, or Black-Scholes, implied volatility. The second compared raw and mean errors for the five day moving average prediction of premiums for both the approximated implied volatility and Black-Scholes iterated volatility. For both strategies, the raw and mean errors were analyzed to measure approximation accuracy for the two datasets. The data includes 331 daily November Soybean option premiums from 1986 to 1988 and 366 daily December Corn option premiums for the same contract years.

The first comparison used by Curtis and Carriker resulted in mean errors of 0.5973 for December corn and 0.4283 for November soybeans. The second comparison resulted in mean errors of -0.000818 and -0.00146 for December corn put and call options, respectively; and mean errors of -0.000876 and -0.004205 for November soybean put and call options, respectively. The authors note that their approximation is accurate except in the days prior to expiration where the approximations and benchmark values differ. This will be the case not only for the Direct Implied Volatility Estimate approximation, but for all approximations due to the nature of options contracts near to
expiration. Although this method tests accuracy against real market data, the data sets are relatively small containing only a few years of data.

Brenner and Subrahmanyam provide little analysis of the accuracy of their model. However, they do suggest that there might be “nontrivial estimation errors when the option is not exactly at-the-money” and that taking the straddle, or a put and a call together; will improve the accuracy of the approximation. The authors use this straddle approach to improve the accuracy of their approximation.

The accuracy of the Bharadia et al. model was evaluated by comparing their model to the Brenner and Subrahmanyam approximation, the Manaster-Koehler approximation, as well as an iterated Black-Scholes benchmark. Manaster and Koehler provide an algorithm which converges monotonically and quadratically to an implied variance, which is essentially an additional benchmark rather than a pure approximation method (Manaster and Koehler). The authors found that their model was closer to the Black-Scholes volatility than both the Brenner and Subrahmanyam method and the Manaster-Koehler method. They tested their model for accuracy against a set of hypothesized call options with times to maturity of 0.25, 0.5, 0.75, and one year; fixed interest rates; a fixed annualized volatility of 35%; and a fixed stock/strike price ratio (Bharadia et al.). The errors (actual-estimated volatility) were found and plotted against moneyness (S/X) for each of the three models. Using these plots to analyze accuracy, the authors show that their technique obtains very accurate results for options that are at-the-money as well as when options are deeper in or out-of-the-money. Whereas, the Brenner and Subrahmanyam and Manaster-Koehler methods only provide accurate estimates
when the options are very close-to-the-money, with accuracy deteriorating as option values move away from the money.

Corrado and Miller analyzed the accuracy of their approximation by comparing their method with the Brenner and Subrahmanyam method and a benchmark of the Black-Scholes model. These three methods were used to calculate implied volatilities for a small set of American style options, or options which can be exercised anytime prior to expiration, on real stocks using the two closest strike prices on either side of the actual stock price (Corrado and Miller). Calculation of implied volatility was done using time to maturity of 29 days and an interest rate of 3%. It was found that the Corrado and Miller method was very close to the benchmark, where the Brenner and Subrahmanyam method was only accurate when approximating volatility for options very close-to-the-money.

In analyzing the accuracy of his model, Li (2005) notes that Corrado and Miller’s method provides the most accurate approximation and that it will be used as a benchmark for testing his model. This is done with two sets of hypothesized options, one for in-the-money call options, $\eta = 0.95$, and one set for out-of-the-money calls, $\eta = 1.05$. The two data sets contain Black-Scholes benchmark volatilities ranging from 15% to 135%, and times to maturity from 0.1 to 1.5 years, with all other variables held constant. Li (2005) calculated estimation errors (estimated volatility-Black-Scholes volatility) for both his method and the Corrado and Miller method over the two data sets. Each data set reveals that the error using Li (2005)’s method is, on average, about 0.021 less than when using Corrado and Miller’s method.
Chargoy-Corona and Ibarra-Valdez analyze accuracy using mathematical proofs with no application to actual market data. The authors claim “Our contribution… is mainly theoretical; hence we did not test our results against market data” (Chargoy-Corona and Ibarra-Valdez).

Each of the methods presented here make various assumptions which limit the accuracy of approximating implied volatility. This study will overcome these limitations by analyzing each method over two extensive real market data sets. In addition, the accuracy of each method will be analyzed considering three different observed variables, moneyness, time to maturity and changing interest rates. By testing all of these methods over the same data set a true determination of which method provides the most accurate approximation will be found.

Other Contributions

Although the following papers did not result in an approximation method tested in this study, their contribution to the literature is deemed significant and is therefore included. The first contributing paper is provided by Don Chance (1996), where he presents an improvement to the Brenner and Subrahmanyam method. He notes the importance of implied volatility calculations for at-the-money options but then asserts that the implied volatility calculation for an at-the-money option will not be the same as one for another strike price due to strike price bias (Chance). Strike price bias is represented by the under prediction of out-of-the-money option premiums using the Black-Scholes model, where under prediction increases as the ratio of strike price to spot price increases (Borensztein and Dooley). Chance presents an improved approximation
stemming from the Brenner and Subrahmanyam approximation for the calculation of implied volatility at varying strike prices. In doing so, Chance takes the Brenner and Subrahmanyam method as a starting value and adds a variable which represents the change in volatility due to changes in strike price.

Chambers and Nawalkha start their discussion of implied volatility approximations by pointing out a shortfall of Chance’s approximation method. Specifically Chance’s model requires a starting option price, then derives an approximation for the at-the-money option including two variables. Chance’s second order Taylor series expansion:

\[ \Delta c^* = \frac{\partial c^*}{\partial X^*} (\Delta X^*) + \frac{1}{2} \frac{\partial^2 c^*}{\partial X^*^2} (\Delta X^*)^2 + \frac{1}{2} \frac{\partial^2 c^*}{\partial \sigma^*^2} (\Delta \sigma^*)^2 + \frac{1}{2} \frac{\partial^2 c^*}{\partial \sigma^* \partial X^*} (\Delta \sigma^* \Delta X^*) \] (12)

Where \( \Delta X = X - X^* \), \( \Delta \sigma^* = \sigma - \sigma^* \)

The first variable used in Chance’s Taylor series approximation is one that allows for the exercise price to stray from exactly at-the-money, the other is an approximation of volatility as the option’s strike price strays from exactly at-the-money. Chambers and Nawalkha simplify Chance’s approach by removing the strike price variable from the Taylor series relying only on the volatility variable shown as:

\[ \Delta c^* = \frac{\partial c^*}{\partial \sigma^*} (\Delta \sigma^*) + \frac{1}{2} \frac{\partial^2 c^*}{\partial \sigma^*^2} (\Delta \sigma^*)^2 \] (13)

This improvement of Chance’s formula provides a more accurate approximation represented by the reduction of mean absolute values of estimation error for hypothesized options.
Chambers and Nawalkha also describe a limitation in the Corrado and Miller model which requires no initial starting point; however, the authors mention one possible shortcoming of the Corrado and Miller model. By including a square root term in the approximation method, the model is opened to cases where there might not be a real solution, or where there might be division by zero resulting in no solution in some cases (Chambers and Nawalkha). This shortcoming is observed to happen in less than 1% of the data for the present study. Chambers and Nawalkha then modify the Corrado and Miller method by replacing the square root term with a term that provides real solutions. This modified Corrado and Miller method is then tested against the same data set and the results show that this modified method is far less accurate than the modified Chance model.

Chambers and Nawalkha also review the Bharadia et al. approximation method in comparison to the Corrado and Miller method and modified Chance model. The Bharadia et al. method is then tested over the same data set resulting in mean absolute errors which are far less accurate than the modified Chance model and the modified Corrado and Miller model. By using a hypothesized set of options, Chambers and Nawalkha can clearly demonstrate the accuracies and impacts of changing variables on the methods, but hypothesized options do not show the frequency of accuracy and impacts from changing variables in real data. This paper is also limited to the requirement that an estimate of volatility be used as a starting value. For these reasons, the Chance model and the modification of Chance’s model provided by Chambers and Nawalkha are not included in this study.
Latane and Rendleman’s study was the first to provide valuable information on how changes in the observable variables affect not only the calculation of a call premium, but also the accuracy of the implied volatility approximation.

Latane and Rendleman first noted in 1976 that each observable variable has a changing impact on the resulting call premium (Latane and Rendleman). This is an important fact because it points out how the accuracy of the implied volatility approximation will be impacted by these changing variables. For example, as an option gets closer to its expiration there is great difficulty in accurately approximating implied volatility. Another example is the effect of volatility where options are close to, or at-the-money, versus when they stray further away from the money. As options stray away from the money the accuracy of volatility begins to diminish relative to near-the-money options. These facts of implied volatility from this early approximation method by Latane and Rendleman are facts which hold for all further approximation methods. Their model approximates volatility by taking the implied volatilities for all options traded on a given underlying asset and weighting them by the partial derivative of the Black-Scholes equation with respect to each implied volatility. Due to the complexities of their study which no longer make it a simple approximation method, the Latane and Rendleman method was not included in the analysis.

Another method provided in the paper “Approximate inversion of the Black-Scholes formula using rational functions” by Minqiang Li (2006). Here, Li presents an approximation method which is claimed to be a simple method which can be executed using spreadsheets. However, this rational approximation method is far from simple;
requiring the use of 31 numerical parameters. Although Li presents an approximation method it becomes cumbersome and tedious when attempting to apply it to a spreadsheet form. For this reason it was not included in the analysis of accuracy conducted in this study.

The next topic which deserves mention is an accuracy analysis by Isengildina-Massa, Curtis, Bridges and Nian (Isengildina-Massa et al.). The authors provide a study which serves as the foundation for the present study by their similar accuracy analysis over some of the same approximation methods. The options used by the authors were closest to the money, but not in-the-money options. This resulted in strong biases towards overestimated implied volatility in the data. These biases in data are overcome by the use of similar datasets that have additional observations through the 2008 contract year which use nearest-to-the-money options, both in and out-of-the-money.

The discussion in this section demonstrated that each of the approximation methods presented here use different benchmarks as well as different hypothesized option values as a means of testing accuracy. This study overcomes these limitations by testing the Curtis and Carriker, Brenner and Subrahmanyam, Chargoy-Corona and Ibarra-Valdez, Corrado and Miller, Bharadia et al. and Li (2005) methods for approximating implied volatility using two large real market data sets which contain all of the natural market conditions which might affect a model’s accuracy. The present study analyzes the accuracy of these approximation methods together through the use of a single Black-
Scholes benchmark volatility using improved measures of accuracy. The extensive nature of the data used for this study is discussed in the following chapter.
CHAPTER III

DATA

The aim of this study is to test accuracy of six implied volatility approximation methods developed in the previous studies. These methods will be analyzed together using real market data which contains all of the necessary input variables over which the methods will be tested for accuracy.

The data sets comprised of 20 years of data are necessary in order to ensure robust results which capture a wide range of market conditions. The first decision made was to have both storable and non-storable commodity types, and therefore two data sets; a crop commodity, corn, and a live stock commodity, live cattle. The second important decision made was to use December contracts for each of these commodities. By confining the data to one contract month it is easy to compare data and approximation performance, as well as assess accuracy in various market conditions.

The futures data was gathered from INFOTECH and resulted in a data set comprised of a single futures closing price for each day from April of 1985 through November of 2008. Options data from 1985 through 2005 was gathered from INFOTECH, and options data from 2006 through 2008 was obtained from Barchart.

The SAS code presented in Appendix A.1 shows the procedures used to combine the calls with the futures as well as the puts with futures. An important decision made here was how to appropriately combine the extensive call and put data with the daily futures prices. The decision commanded SAS to merge the call option premiums with the futures prices by finding the minimum difference between the various strike prices.
and the single futures price for each day. Here, the minimum difference is represented by
the closest strike price to futures price; a value no greater than +/- $5, for both corn and
live cattle. There were a few observations in the early years of the data where fewer strike
prices were traded and therefore the closest to the money options were further away from
the money. These select observations were removed due to the reduced accuracy of
approximating implied volatility. This resulted in a data set where the strike price
available for each day was combined with the single futures price. Doing this ensured a
dataset where only closest-to-the-money options were used. This was done for several
reasons, the most important of which being, as mentioned previously, that the majority of
the approximations are defined for at-the-money options, or where futures equal a
discounted strike. The low likelihood of futures equaling exactly a discounted strike price
allowed for the use of closest-to-the-money options to be used as a guideline for selecting
the data.

Now that both the call options and the put options were merged with futures, an
important decision on how to properly combine the two data sets was made to ensure
uniformity of the data. Again, this called for the use of SAS (Appendix A.2), where the
two datasets were merged by date, resulting in each observation containing the following
variables: date, contract, futures settle price, closest-to-the-money strike price for calls
and puts, a call premium and a put premium. Unfortunately, as is the nature of the
options markets, there are several days where the closest-to-the-money strike prices for
calls and puts did not match because one or the other might not have been traded on the
same day. It was found that this frequently occurred in the early years of the data as well
as in the beginning of the contract life. This was the first of several methods for cleansing the data; every observation day where the call strike did not match the put strike was removed from the data set. The resulting data sets were then reduced to 4732 observation days for corn and 3949 observations for live cattle.

Next, a time to maturity variable was introduced into the corn data set. This was done in Microsoft Excel by finding the distance between the current date $t$, and the expiration date $T$, then dividing by 365 for a resulting proportion of a year, $\tau = \frac{(T - t)}{365}$. Here, the second method of cleansing the data was used. In order to have all of the data as uniform as possible, time to maturity was restricted to one year or less, ($\tau \leq 1$). The remaining piece of information necessary for a calculation of each approximation is an interest rate variable. The daily interest rates over the entire data set were found through the Federal Reserve website and merged into the existing data using SAS. Next, the data was cleansed a third time. Again, to ensure uniformity in all of the data, the decision to restrict the data set to complete contract years was made. At this point the corn data set is complete and consists of 4507 observations over 19 contract years.

The exact same procedures were employed for the live cattle data set; however there were a few more obstacles to get over with this data set. Due to the nature of the options there were far more observation days where the call strike price did not match the put strike price, and where the closest-to-the-money options were far away from the futures price. There are a few reasons for this. First, live cattle being a living commodity there were hardly any contracts traded as the time to maturity stretched further away from expiration. In the earlier years in which these options were traded, there were far fewer
strike prices available for calls and puts. It was not till the later years where entire contract years of acceptable data were available. Also, due to inconsistencies in the raw data, the 1997 contract year was removed due to lack of data which met each of the above requirements. Given the methods presented for corn and the data inconsistencies presented here, the live cattle data set consists of 3852 observations over 22 contract years.

The datasets cover the time periods of November 24th 1989 through November 19th 2008 for corn options, and March 27th 1986 through November 28th 2008 for live cattle options. The 19 and 22 years of data for corn and live cattle, respectively, provide many fluctuations in the data which have an impact on volatility. First, these datasets begin at a time when derivatives were not extensively traded and continue into a time when calls and puts on these commodities were heavily traded. This interesting point is shown through the previously mentioned inconsistencies in the early years of the data where the nearest-to-the-money call options have different strike prices than the nearest-to-the-money put options. However, in the later years of the data this inconsistency is much less frequent due to the increase in number of options traded. Next, the length of this dataset covers various bear and bull markets. These bull and bear markets are most noticeable towards the end of each data set with the bull markets of 2006 and 2007 before the bear market of 2008. It is easily seen (Figure 1) that during the bull market volatility decreased and during the bear market of 2008 that volatility sharply increased. These two datasets have some interaction which could affect volatility simultaneously, represented by the fact that corn is used as feed for live cattle.
Figure 1- Black-Scholes Implied Volatility
These two data sets serve as a platform for the accuracy analysis of each of the six approximation methods. As with the formation of the data sets, each approximation method was calculated in Microsoft Excel. Calculating each method resulted in an approximated implied volatility for a call option, a put option, and an average of the two. The six approximation methods were calculated in spreadsheet form with relative ease, which held with the authors' claims.

Now that each approximation method is in place, a benchmark implied volatility value is necessary to study the accuracy. The Black-Scholes implied volatility was calculated using an iterative process in SAS (code in Appendix A.3). A data set containing each of the observable variables was input into SAS along with Black’s formula (2) and a predicted call value was calculated. Due to the size of these data sets and the wide range of approximated implied volatility values, the predicted call premium was calculated by plugging in values of implied volatility over the range 0.001 to .9 for corn call premiums, and 0.001 to .5 for live cattle call premiums by 0.000001. SAS calculated each of these implied volatility values until the difference between the predicted call and actual call (diff = cc - c) price was less than 0.001. This was deemed to be an acceptable difference because the known call values are in dollars and cents; therefore an implied volatility value which predicted a call premium within 0.001 of the actual call premium was taken as the actual Black-Scholes implied volatility value for that observation. A similar procedure was used in SAS (code in Appendix A.4) to find the iterated Black-Scholes implied volatility for put options. The same ranges of implied volatility were used to find predicted put premiums.
The only remaining calculation needed prior to analyzing accuracy is a measure of moneyness. As previously mentioned, the options used in these data sets are closest-to-the-money options; however, a moneyness variable is still necessary for further accuracy analysis. It is important to not only test the data for accuracy against a benchmark Black-Scholes implied volatility but to also test the data over observed changes in market variables. There are measures of moneyness presented in the papers, Li (2005) and Bharadia et al., but the basic definition of moneyness is the distance between the futures price and the option strike price, \((S-X)\) (Hull).

For this study two measures of moneyness were used. The first measure for comparison within each approximation method is defined

\[
M = \frac{d_1 + d_2}{2},
\]

where \(d_1 + d_2\) are the two Black-Scholes parameters. Here, moneyness reduces to \(M = \frac{\ln(S/X)}{\sigma\sqrt{T}}\), or the natural log ratio of futures settle price and option strike price, standardized by \(\sigma\sqrt{T}\) for each approximation method. The resulting values are centered at zero, or when options are exactly at-the-money, with negative values representing out-of-the-money options and positive values representing in-the-money options prices. This measure of moneyness is still a measure of the difference in settle price and strike price but it also takes into account the other variables for each observation. The primary purpose of this definition of moneyness is to obtain a graphical representation of changes in percent errors due to changes in moneyness. Although an alternative definition of moneyness is used in the Bharadia et al. paper, the limited number of observations they were analyzing allowed for a simplified graphical depiction of moneyness. However, with extensive datasets
covering roughly 2 decades, the graphs become unclear and difficult to distinguish changing patterns in error. For this reason, this study employs the use of a modified definition of moneyness for individual analysis and a generalized definition for comparison of all approximations together. Rather than the modified definition, which uses the natural log ratio of futures prices and strike price, and is standardized for each approximation method; the generalized definition is the same across all approximations.

The moneyness variable calculated by Li (2005) was determined to be the best comparison for all the approximations, $\eta = \frac{S}{K}$ where S and K are the discounted futures price and discounted option strike price. Here, moneyness ranges from 0.97561 to 1.0231 for corn, and 0.9466 to 1.0183 for live cattle, with $\eta = 1$ representing at-the-money. This measure serves best because it is uniform throughout the datasets and shows which options are relatively in, out and at-the-money. First, the distribution of moneyness over the entire data set was determined, and because the data is already closest-to-the money, each of these values were very close together. Next, the data sets were broken into separate groups determined by using the first quartile, the middle two quartiles, and the upper quartile. For corn, the middle two quartiles are between moneyness values of 0.99108 and 1.0081, within 1% of being exactly at the money. Within this range all of the approximations are very accurate. However, as moneyness is further in or out of the money, $0.97561 < \eta < 0.99108$, $1.0081 < \eta < 1.0231$ the accuracy of the approximations deteriorates. The same observations are noted for live cattle, with the middle two quartiles between 0.99596 and 1.00478, less than 0.5% of being at-the-
money. These three groups of moneyness will serve to compare accuracy not only between models, but also within each approximation.

Simple descriptive statistics of the approximations and the Black-Scholes benchmark for calls and the average of puts and calls were found and assembled into Table 1 and Table 2, for corn and live cattle. It is easy to see that the difference between the approximation mean and actual Black-Scholes mean is roughly +/- 0.001% for both datasets. On average corn has higher volatility than live cattle. In addition to differences in the means, these statistics show that the variances are lowest for Corrado and Miller, Bharadia et al. and Li (2005). This could be represented by the limiting at-the-money assumptions made by the other three models, which makes these methods less accurate. The difference in the number of observations for Corrado and Miller and the other methods is represented by the case the inclusion of a square root term in this method where there might not be real solutions, as indicated by Chambers and Nawalkha, and discussed previously. This occurs in less than 1% of observations for this study.
### Table 1- Descriptive Statistics for Corn

#### Approximated IV for Calls

<table>
<thead>
<tr>
<th></th>
<th>DIVE</th>
<th>ISD</th>
<th>CCIV</th>
<th>CMIV</th>
<th>BIV</th>
<th>LIIV</th>
<th>BSIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.2402</td>
<td>0.2398</td>
<td>0.2405</td>
<td>0.2396</td>
<td>0.2407</td>
<td>0.241</td>
<td>0.2399</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
</tr>
<tr>
<td>Median</td>
<td>0.233</td>
<td>0.2327</td>
<td>0.2332</td>
<td>0.2304</td>
<td>0.2314</td>
<td>0.2317</td>
<td>0.2307</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.0642</td>
<td>0.064</td>
<td>0.065</td>
<td>0.0597</td>
<td>0.0582</td>
<td>0.0584</td>
<td>0.059</td>
</tr>
<tr>
<td>Sample Var.</td>
<td>0.0041</td>
<td>0.0041</td>
<td>0.0042</td>
<td>0.0036</td>
<td>0.0034</td>
<td>0.0034</td>
<td>0.0035</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.9449</td>
<td>2.9706</td>
<td>2.7607</td>
<td>4.9035</td>
<td>4.048</td>
<td>4.0124</td>
<td>3.9307</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.9017</td>
<td>0.9019</td>
<td>0.8653</td>
<td>1.4855</td>
<td>1.4536</td>
<td>1.4512</td>
<td>1.4054</td>
</tr>
<tr>
<td>Range</td>
<td>0.6277</td>
<td>0.6275</td>
<td>0.6312</td>
<td>0.694</td>
<td>0.548</td>
<td>0.5482</td>
<td>0.5701</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.0069</td>
<td>0.0069</td>
<td>0.0069</td>
<td>0.0588</td>
<td>0.0619</td>
<td>0.0619</td>
<td>0.0399</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.6346</td>
<td>0.6344</td>
<td>0.638</td>
<td>0.7529</td>
<td>0.6099</td>
<td>0.6101</td>
<td>0.61</td>
</tr>
<tr>
<td>Sum</td>
<td>1082.5</td>
<td>1081</td>
<td>1083.8</td>
<td>1076.6</td>
<td>1084.7</td>
<td>1086.1</td>
<td>1081.3</td>
</tr>
<tr>
<td>Count</td>
<td>4507</td>
<td>4507</td>
<td>4507</td>
<td>4493</td>
<td>4507</td>
<td>4507</td>
<td>4507</td>
</tr>
</tbody>
</table>

#### Approximated IV for Average of Put and Call

<table>
<thead>
<tr>
<th></th>
<th>DIVE</th>
<th>ISD</th>
<th>CCIV</th>
<th>CMIV</th>
<th>BIV</th>
<th>LIIV</th>
<th>BSIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.2411</td>
<td>0.2408</td>
<td>0.2411</td>
<td>0.2404</td>
<td>0.2415</td>
<td>0.2418</td>
<td>0.2407</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
</tr>
<tr>
<td>Median</td>
<td>0.2315</td>
<td>0.2312</td>
<td>0.2317</td>
<td>0.2328</td>
<td>0.233</td>
<td>0.2333</td>
<td>0.2325</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.0585</td>
<td>0.0583</td>
<td>0.0585</td>
<td>0.0624</td>
<td>0.0632</td>
<td>0.0634</td>
<td>0.0637</td>
</tr>
<tr>
<td>Sample Var.</td>
<td>0.0034</td>
<td>0.0034</td>
<td>0.0034</td>
<td>0.0039</td>
<td>0.004</td>
<td>0.004</td>
<td>0.0041</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.4678</td>
<td>1.4705</td>
<td>1.4671</td>
<td>1.0976</td>
<td>1.1776</td>
<td>1.1754</td>
<td>1.1014</td>
</tr>
<tr>
<td>Range</td>
<td>0.5349</td>
<td>0.5347</td>
<td>0.5342</td>
<td>0.6156</td>
<td>0.5748</td>
<td>0.5749</td>
<td>0.5558</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.0756</td>
<td>0.0756</td>
<td>0.0755</td>
<td>0.0766</td>
<td>0.0686</td>
<td>0.0686</td>
<td>0.0653</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.6105</td>
<td>0.6103</td>
<td>0.6097</td>
<td>0.6921</td>
<td>0.6435</td>
<td>0.6435</td>
<td>0.6211</td>
</tr>
<tr>
<td>Sum</td>
<td>1086.6</td>
<td>1085.1</td>
<td>1086.9</td>
<td>1080.2</td>
<td>1088.2</td>
<td>1089.8</td>
<td>1084.8</td>
</tr>
<tr>
<td>Count</td>
<td>4507</td>
<td>4507</td>
<td>4507</td>
<td>4493</td>
<td>4507</td>
<td>4507</td>
<td>4507</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)
BSIV represents the iterated Black-Scholes implied volatility
Table 2- Descriptive Statistics for Live Cattle

<table>
<thead>
<tr>
<th></th>
<th>DIVE</th>
<th>ISD</th>
<th>CCIV</th>
<th>CMIV</th>
<th>BIV</th>
<th>LIV</th>
<th>BSIV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.1344</td>
<td>0.1343</td>
<td>0.1345</td>
<td>0.1347</td>
<td>0.1354</td>
<td>0.1354</td>
<td>0.1346</td>
</tr>
<tr>
<td><strong>Std. Error</strong></td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>0.1281</td>
<td>0.128</td>
<td>0.1282</td>
<td>0.1299</td>
<td>0.1306</td>
<td>0.1306</td>
<td>0.1298</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>0.0435</td>
<td>0.0435</td>
<td>0.0439</td>
<td>0.0389</td>
<td>0.0388</td>
<td>0.0388</td>
<td>0.039</td>
</tr>
<tr>
<td><strong>Sample Var.</strong></td>
<td>0.0019</td>
<td>0.0019</td>
<td>0.0019</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>1.6607</td>
<td>1.6601</td>
<td>1.6098</td>
<td>2.4917</td>
<td>2.4887</td>
<td>2.4899</td>
<td>2.4913</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>0.8666</td>
<td>0.8659</td>
<td>0.8597</td>
<td>1.1601</td>
<td>1.1601</td>
<td>1.1607</td>
<td>1.1434</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>0.431</td>
<td>0.4305</td>
<td>0.4299</td>
<td>0.3982</td>
<td>0.4073</td>
<td>0.4078</td>
<td>0.4213</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>0.0016</td>
<td>0.0016</td>
<td>0.0016</td>
<td>0.0421</td>
<td>0.0331</td>
<td>0.0331</td>
<td>0.0195</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>0.4326</td>
<td>0.4321</td>
<td>0.4315</td>
<td>0.4403</td>
<td>0.4403</td>
<td>0.4409</td>
<td>0.4408</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>517.68</td>
<td>517.51</td>
<td>518.07</td>
<td>517.45</td>
<td>521.4</td>
<td>521.57</td>
<td>518.54</td>
</tr>
<tr>
<td><strong>Count</strong></td>
<td>3852</td>
<td>3852</td>
<td>3852</td>
<td>3842</td>
<td>3852</td>
<td>3852</td>
<td>3852</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>DIVE</th>
<th>ISD</th>
<th>CCIV</th>
<th>CMIV</th>
<th>BIV</th>
<th>LIV</th>
<th>BSIV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.1354</td>
<td>0.1353</td>
<td>0.1354</td>
<td>0.1356</td>
<td>0.1363</td>
<td>0.1363</td>
<td>0.1355</td>
</tr>
<tr>
<td><strong>Std. Error</strong></td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0007</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>0.1305</td>
<td>0.1305</td>
<td>0.1304</td>
<td>0.13</td>
<td>0.1309</td>
<td>0.1309</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>0.0385</td>
<td>0.0384</td>
<td>0.0386</td>
<td>0.041</td>
<td>0.0412</td>
<td>0.0412</td>
<td>0.0411</td>
</tr>
<tr>
<td><strong>Sample Var.</strong></td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>1.5111</td>
<td>1.5123</td>
<td>1.4967</td>
<td>1.4981</td>
<td>1.7373</td>
<td>1.7339</td>
<td>1.5221</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>1.0495</td>
<td>1.0491</td>
<td>1.0485</td>
<td>0.9561</td>
<td>1.0109</td>
<td>1.0109</td>
<td>0.9593</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>0.3393</td>
<td>0.3393</td>
<td>0.3379</td>
<td>0.3574</td>
<td>0.3582</td>
<td>0.3582</td>
<td>0.3575</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>0.032</td>
<td>0.032</td>
<td>0.0318</td>
<td>0.0421</td>
<td>0.0421</td>
<td>0.0421</td>
<td>0.0421</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>0.3713</td>
<td>0.3712</td>
<td>0.3698</td>
<td>0.3995</td>
<td>0.4002</td>
<td>0.4003</td>
<td>0.3996</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>521.42</td>
<td>521.24</td>
<td>521.39</td>
<td>520.87</td>
<td>524.93</td>
<td>525.11</td>
<td>522.06</td>
</tr>
<tr>
<td><strong>Count</strong></td>
<td>3852</td>
<td>3852</td>
<td>3852</td>
<td>3842</td>
<td>3852</td>
<td>3852</td>
<td>3852</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)
BSIV represents the iterated Black-Scholes implied volatility
CHAPTER IV

METHODS

Traditional measures of analyzing accuracy include: mean error, root mean squared error, mean absolute error and mean absolute percent error (Poon and Granger). Although these traditional measures provide a determination of an approximation’s accuracy, few studies consider measures other than mean errors and variants of mean errors. To provide a more detailed determination of accuracy it is important to analyze moments in addition to the mean, as well as how errors change given variation of the input variables. This study analyzes the errors, percent errors and mean of percent errors, but also considers variations of these errors, provided by analysis of error histograms, as well as analysis of errors given changes in observed variables. In addition to these, this study also provides a goodness of fit measure, or an adjusted $R^2$ value, to compare method accuracy. By analyzing these additional measures, the present study goes beyond traditional measures to give a redundant and practical determination of accuracy.

Error Histograms

The first step in determining the accuracy of these models was to calculate the raw error (12) and percent error (13) for every observation:

$$e_t = (A_t - B_t)$$  \hspace{1cm} (12)

$$p_t = \left( \frac{(A_t - B_t)}{B_t} \right) \times 100$$  \hspace{1cm} (13)
Where, $A_t$ is the approximated volatility, and $B_t$ is the Black-Scholes implied volatility. The raw errors from each approximation were used to find individual error histograms, each scaled to have the same axes for appropriate comparison. This was done by finding the minimum and maximum error among all 6 approximations then setting the bin size equal to $(\text{max-min})/8$. These histograms give visual measures of traditional accuracy such as mean error, but they also give measures of variance, skewness, minimum, and maximum of the errors.

**Adjusted $R^2$**

A measure of accuracy traditionally used to evaluate accuracy is Root Mean Squared Error, which is defined as the square root of the expected value of the errors.

$$\text{RMSE} = \sqrt{\frac{\sum e^2}{n}} \quad (14)$$

The radicand, or the mean squared error, is the sum of the squared errors between each approximation and the Black-Scholes benchmark volatility. The square root of the resulting mean squared error value is taken to arrive at the root mean squared error. While this provides a measure of the spread of errors about the Black-Scholes benchmark, it serves as a comparison among each approximation method rather than a standardized measure of how closely each method approximates the Black-Scholes implied volatility. Therefore, this study uses a similar accuracy measure, adjusted $R^2$.

The adjusted $R^2$ was found by plotting the approximated implied volatility values on the y-axis and the Black-Scholes implied volatility values on the x-axis. Next, a line
of perfect agreement, or (1:1) line, was drawn. The perfect agreement line was used rather than the predicted least squares line in order to find errors associated with the Black-Scholes implied volatility, rather than a predicted least squares line. The sum of squared errors associated with this line represents the mean squared error previously discussed. The adjusted $R^2$ was defined as:

$$R^2 = 1 - \frac{\text{SSE}(1:1)}{\text{SSE}(\text{mean})} = 1 - \frac{\sum (A_t - B_t)^2}{\sum (A_t - \bar{A}_t)^2}$$  \hspace{1cm} (15)$$

Where SSE (1:1) is the sum of the squared deviations of the perfect agreement line and SSE (mean) is the sum of the squared deviations from a horizontal line representing the mean of the approximation, or $\bar{A}_t$. This calculation provides a standardized measure of the discrepancy between each approximation method and the Black-Scholes implied volatility. The adjusted $R^2$ values, between 0 and 1, provide a measure of how accurate each approximation is individually and how well it compares to the other approximation methods.

**Changes in Error over Observed Market Variables**

The next measure of accuracy is the relationship of each approximation’s percent error and three input variables; time to maturity, $\tau$, interest rates, $r$, and moneyness, $(S/X)$. These relationships can be analyzed graphically by plotting approximation percent error on the y-axis and each input variable on the x-axis. Each table gives a simple visual representation of the relationship of accuracy and the three variables. Additionally,
statistical tests may be used to compare the mean percent errors for different levels of the three variables.

To accommodate statistical analysis, groups of the three variables should be made for moneyness, using Li’s (2005) definition $\eta = \frac{K}{S}$. Three groups were defined based on the first quartile, the middle two quartiles, and the fourth quartile of this variable. By dividing the data this way, it is easy to analyze the accuracy of each approximation not only very close-to-the-money, but how the approximation’s accuracy is affected as the options get further away from the money.

As previously discussed, approximation accuracy decreases as time to maturity approaches expiration. Based on time to maturity, the data is divided into two groups: below .2, or 20% of year, and above .2. This was done because the largest fluctuations of percent errors, above 25%, are all within 20% of a year till expiration. Beyond this the percent errors are consistently low, below 25% error. Next, the interest rate variable was separated roughly in half, or at 5%. The interest rates over the data set ranged from less than 1% to nearly 10% so a break at 5% was used.

The percent errors were separated into groups, as specified above, then three samples of 100 were randomly selected from each approximation over each group using JMP. Because there are no specific well-known tests to analyze other parameters such as skewness, minimums and maximums, random samples were chosen to ensure the Central Limit Theorem held, or that means of each sample are approximately normal. The sample
means allowed for analysis of variance and Fishers Least Significant Difference test to be conducted.

With the random samples of each group, analysis of variance was used to test for overall differences in the methods, overall differences among the groups as well as differences in the interaction of methods and groups.

Statistical differences across groups can be analyzed by first using the F ratio:

\[
F = \frac{MS_1^2}{MS_2^2}
\]  \hspace{1cm} (14)

Where, \( MS_1^2 \) = the mean squared error between the methods and \( MS_2^2 \) = the mean squared error for the interaction of the methods among the groups of the observed market variable (Mendenhall and Sincich). The F ratio along with its associated p-value, allow for a decision to either reject the null hypothesis or fail to reject the null hypothesis; where the null hypothesis is that there are no differences in means among the groups. If the decision is made to reject the null, represented by a p-value less than the level of significance, then Fishers Least Significant Difference (LSD) test is used to determine where there are significant differences among the means. This test provides a pairwise comparison of means for every pair of methods between each group. Fishers Least Significant difference test shown as:

\[
LSD_{ij} = t_{\alpha/2} \sqrt{s_w^2 \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}
\]  \hspace{1cm} (15)
Where $i$ and $j$ represent two different means, $s^2_{pooled}$ is the pooled estimator of population variance, $n_i$, and $n_j$ are the sample sizes from population $i$ and $j$, and $t_{\alpha/2}$ is the critical value (Ott).

It is important to consider approximation accuracy over multiple changing variables represented in the market in addition to traditional measures of mean errors. Therefore, this study considers several tradition measures as well as histograms of errors, adjusted $R^2$ measures, and statistical tests to analyze approximation accuracy over three observed variables. Doing this provides farmers, producers, manufacturers and even speculators a comprehensive and robust determination of which method should be used to approximate implied volatility.
CHAPTER V
RESULTS

This chapter discusses results of analysis of the Black-Scholes methods developed in the previously studies. From these results, a method, or possibly group of methods will emerge as most accurate given analysis of errors, adjusted $R^2$, and accuracy over changing market variables.

Error Histograms

The descriptive statistics of Black-Scholes and the six approximation methods, shown in Tables 1 and 2, demonstrate that all of the approximations appear to be satisfactory methods of approximating Black-Scholes implied volatility. However, these statistics show very little of how well they approximate volatility over the entire data set. Traditional methods of determining accuracy such as analysis of mean absolute and percent errors fail to grasp changes over time in a large data set, or how the errors vary throughout the data. This study considers mean errors, but goes beyond this by plotting histograms of the errors which display much more information, such as variance, skewness, minimum, and maximum of the errors. With each histogram plotted together on the same axes it is easy to see how well each method compares to the others.

The histograms, located in Figures 2 and 3, present three obvious groups within the 6 approximations. The first group, comprised of the Corrado and Miller approximation, has a mean located in the bin which includes zero, and with no other bars present, there is essentially no variation outside of this first bin. With the minimum error
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 2- Corn Calls Error Histograms
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 3- Live Cattle Calls Error Histograms
and maximum errors also in the first positive bin, Corrado and Miller clearly stands out as a very accurate approximation method. The next group comprised of Li (2005) and Bharadia et al., where both methods have mean errors located in the bin closest to zero. Unlike Corrado and Miller, these methods show slight variation in the errors, with a few observations falling in the bin with a midpoint of 0.078 for corn and 0.0605 for live cattle. Although these are still considered very accurate approximations, they are clearly not as accurate as Corrado and Miller. Next is the group comprised of Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez. These approximations have much more variation, with errors ranging from -0.039 to 0.196 for corn and -0.1010 to 0.3296 for live cattle. The majority of the observations have errors located in the same bin as the other two groups, indicating means similar to the two more accurate groups. Rather than analyzing differences in means, these histograms provide more information such as variance, skewness, minimum and maximums of the errors. All of the mean errors for these approximation methods appear to be similar; however, it is easy to see how they differ through the variation. This allows for the first determination of accuracy to be based on more than just a comparison of mean errors.

Each of the approximations was plotted with percent error over the duration of the data set to distinguish patterns in the errors. These graphs display the first patterns of how percent errors vary more as the option approaches expiration with the greatest percent error occurring just prior to expiration (Figures 4 and 5). The errors which occur just before expiration are represented by the large spikes. By analyzing each of these graphs it is easy to distinguish the three groups of approximations as well as the relative
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

**Figure 4- Corn Calls Approximations Percent Error**
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 5- Live Cattle Calls Approximations Percent Error
accuracy of each. Again it is shown that the Corrado and Miller method is the most accurate approximation for both data sets, with the majority of the errors less than -2%. The next group consisting of Bharadia et al. and Li (2005) show the majority of the errors are well less than 10% with only few full of spikes greater than this. The third group represented is comprised of the Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez methods. Each of these graphs has a majority of errors less than 25%, with various spikes greater than this. The pattern of these groups show the most accurate approximation of Black Scholes, represented by lowest percent errors, are the Corrado and Miller method; followed by Bharadia et al. and Li (2005). The remaining three approximations; Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez all have very similar approximations; however, the relative accuracy of these approximations is weak in comparison to the other approximation models. Corn and live cattle show the same patterns in approximation accuracy when analyzing error histograms and therefore live cattle results are the same as the discussed corn results. A noticeable difference between the two datasets is the fact that the Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez approximations have a wider range of percent errors for the live cattle data versus corn.

Adjusted $R^2$

An adjusted $R^2$ value, as mentioned in the previous chapter, was calculated for each approximation. This value demonstrates how closely each approximation measure
is to the actual Black-Scholes implied volatility. The results indicate that the Corrado and Miller model has the strongest correlation of 0.99989 for the corn data set and 0.999971 for the live cattle data. This shows that the Corrado and Miller approximated implied volatility matches the Black-Scholes implied volatility almost one to one. Next, are the Bharadia et al. and Li (2005) approximations with adjusted $R^2$ values of approximately 0.993 for corn and 0.992 for live cattle. These two also have very strong correlations with the Black-Scholes implied volatility, but are slightly less accurate than Corrado and Miller. The remaining three approximations, Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez have much lower adjusted $R^2$ values of roughly 0.8 for both corn and live cattle. These results again show that Corrado and Miller is the most accurate followed by Bharadia et al. and Li (2005), with Curtis and Carriker, Brenner and Subrahmanyam, and Chargoy-Corona and Ibarra-Valdez being relatively less accurate.

Model Accuracy over Observed Market Variables

It has already been shown that the Corrado and Miller approximation is the most accurate overall, as demonstrated by the error histograms, and the very high adjusted $R^2$ values. Now, model accuracy will be analyzed over the observed market variables: moneyness, time to maturity, interest rates by analyzing means and variances of model errors given different market variables. This analysis will be done by first performing a graphical analysis of how each approximation varies given the individual market
variables. Then they will be tested further using statistical analysis to confirm patterns observed in the graphs.

As mentioned previously, there are two variables for moneyness. The first is used to perform a graphical analysis of moneyness for each approximation individually followed by a variable for moneyness which is used to compare each of the approximation methods to each other.

The three groups of approximation accuracy are easily identified with the graphical analysis of percent error versus moneyness, where moneyness is defined as the average of the two Black-Scholes parameters \( M = \frac{d_1 + d_2}{2} \). Figures 6 and 7 clearly show that the Corrado and Miller approximation is only slightly affected by moneyness with percent error dropping almost negligible amounts below zero as when options are not exactly at-the-money.

The next group, of Bharadia et al. and Li (2005), present very similar results. When the options are very close-to-the-money the accuracy is hardly effected. However, as moneyness gets further from being at-the-money the percent error goes above 50%, being slightly higher as options are further out-of-the-money. These models are considered to be accurate, but it is interesting to note the observed declines in accuracy as moneyness gets further from being at-the-money.

The third group presents the strongest changes in accuracy relative to moneyness.
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 6- Corn Percent Calls Errors and Moneyness
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 7- Live Cattle Calls Percent Errors and Moneyness
These changes can be attributed to the fact that these methods were developed for at-the-money options. Curtis and Carriker, Brenner and Subrahmanyam and Chargoy-Corona and Ibarra-Valdez each have very similar graphs of percent error versus moneyness, with errors of roughly -50% and below when options are out-of-the-money, and percent errors above 150% when options are in-the-money. These graphs give a great picture of the limitation of these three models as the errors are drastically affected with marginal changes in moneyness. Looking at the Bharadia et al., Li (2005), and Corrado and Miller methods it is easy to observe the changes made from their starting point of the Brenner and Subrahmanyam method. These methods are developed for options that are not limited to being at-the-money; and therefore the low percent errors which extend further away from being exactly-at-the-money, clearly show the improved accuracy. Given that the majority of all traded options are near-the-money, rather than at-the-money, these three approximation methods all appear to be accurate and useful approximations of implied volatility.

The graphical analysis for the live cattle options reveals the same patterns as the corn options (Figure 7). There are three distinct groups of accuracy: Corrado and Miller as the relatively most accuracy, followed by Bharadia et al. and Li (2005), then Curtis and Carriker, Brenner and Subrahmanyam and Chargoy-Corona and Ibarra-Valdez being relatively less accurate.

To further test these approximations, statistical tests were used to analyze approximation accuracy as moneyness changes. This is easily done using Li’s (2005)
definition of moneyness, \( \eta = \frac{S}{K} \) where \( S \) and \( K \) are the discounted values of the futures settle price and option strike price. It is important to note that this definition of moneyness is a simplified version of the previous definition, by being standardized across all approximation methods, therefore having no impact on results. The use of statistical tests and groups of moneyness were used to find where there are statistically significant differences between each group of moneyness. Doing this gives statistical evidence to support the observations made from the graphs.

With the samples of each approximation for each group of moneyness read into JMP, an analysis of variance, or ANOVA, was run to test the effect that each group has on approximation accuracy. The null hypothesis is that there are no differences between the means of the percent errors for each of the methods, groups, and the interaction between the two. If the null hypothesis is rejected, then there are significant differences between the methods and different groups of moneyness.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>17</td>
<td>3908.56</td>
<td>229.915</td>
<td>381.312</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>21.7065</td>
<td>0.603</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>53</td>
<td>3930.27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ANOVA was conducted for the means of each sample as well as the variances. The first results analyzed were for the means of corn calls percent error and moneyness. The results in Table 3 prove a rejection of the null hypothesis, p-value <.0001, for the
interaction of method and group, which indicates that there is a difference in means of
percent error among the groups of moneyness. From this rejection, it is shown using
Fishers Least Significant Difference Test that three methods, Curtis and Carriker,
Chargoy-Corona and Ibarra-Valdez and Brenner and Subrahmanyam had mean percent
errors which were significantly different among each of the three groups of moneyness;
where L represents the lower quartile of moneyness, B represents the middle two
quartiles and G represents the upper quartile (Table 4). This result indicates that the mean
errors, for those methods, are significantly different for options that are more than 1%
away from being exactly at-the-money. All of the other groups were not significantly
different among any group of moneyness for corn. This confirms the initial results
observed from the graphs. Results for differences in means between the three groups of
moneyness for live cattle also show a rejection of the null, with a p-value<0.0001 (Table
5).

As seen in Table 6, the live cattle data resulted in the same significant differences in the
means of percent errors between the different groups of moneyness. The methods of
Curtis and Carriker, Chargoy-Corona and Ibarra-Valdez and Brenner and Subrahmanyam
all had significant differences between the mean errors of being in, at and out-of-the-
money. This result confirms the graphical analysis, that the percent errors are much
higher for these groups when the options are not in the middle two quartiles of
moneyness, or within 0.5%. The other three methods showed no differences between
groups of moneyness.
Table 4- Corn Calls Means, Moneyness
LS Means Differences

\( \alpha = 0.050 \quad t = 2.02809 \)

<table>
<thead>
<tr>
<th>Level</th>
<th>Least Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCIV,L</td>
<td>18.9124</td>
</tr>
<tr>
<td>ISD,L</td>
<td>15.5604</td>
</tr>
<tr>
<td>DIVE,L</td>
<td>13.7324</td>
</tr>
<tr>
<td>LIIV,L</td>
<td>1.1462</td>
</tr>
<tr>
<td>BIV,G</td>
<td>1.11556</td>
</tr>
<tr>
<td>LIIV,G</td>
<td>0.99979</td>
</tr>
<tr>
<td>BIV,L</td>
<td>0.99778</td>
</tr>
<tr>
<td>CCIV,B</td>
<td>0.5526</td>
</tr>
<tr>
<td>DIVE,B</td>
<td>0.27463</td>
</tr>
<tr>
<td>LIIV,B</td>
<td>0.12653</td>
</tr>
<tr>
<td>ISD,B</td>
<td>0.05419</td>
</tr>
<tr>
<td>BIV,B</td>
<td>-0.0119</td>
</tr>
<tr>
<td>CMIV,B</td>
<td>-0.1305</td>
</tr>
<tr>
<td>CMIV,L</td>
<td>-0.1338</td>
</tr>
<tr>
<td>CMIV,G</td>
<td>-0.1424</td>
</tr>
<tr>
<td>ISD,G</td>
<td>-12.598</td>
</tr>
<tr>
<td>DIVE,G</td>
<td>-12.936</td>
</tr>
<tr>
<td>CCIV,G</td>
<td>-13.917</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chagraye-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)
*Levels not connected by same letter (A-G) are significantly different.

Table 5- Analysis of Variance, Live Cattle Calls Means, Moneyness

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>17</td>
<td>5390.97</td>
<td>317.116</td>
<td>200.386</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>56.9709</td>
<td>1.583</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>53</td>
<td>5447.94</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The next ANOVA was conducted to test changes in the average variances of the six
methods for corn and live cattle. Initial results (Table 7) show a failure to reject the

Table 6- Live Cattle Calls Means, Moneyness
LS Means Differences

\[ \alpha = 0.050 \quad t = 2.02809 \]

<table>
<thead>
<tr>
<th>Level</th>
<th>Least Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCIV,L</td>
<td>21.0264</td>
</tr>
<tr>
<td>DIVE,L</td>
<td>18.2416</td>
</tr>
<tr>
<td>ISD,L</td>
<td>16.9128</td>
</tr>
<tr>
<td>BIV,G</td>
<td>2.18377</td>
</tr>
<tr>
<td>LIIV,G</td>
<td>1.73375</td>
</tr>
<tr>
<td>BIV,L</td>
<td>1.20887</td>
</tr>
<tr>
<td>LIIV,L</td>
<td>0.93759</td>
</tr>
<tr>
<td>DIVE,B</td>
<td>0.41222</td>
</tr>
<tr>
<td>BIV,B</td>
<td>0.17384</td>
</tr>
<tr>
<td>LIIV,B</td>
<td>0.16458</td>
</tr>
<tr>
<td>CMIV,B</td>
<td>-0.0234</td>
</tr>
<tr>
<td>CMIV,G</td>
<td>-0.0554</td>
</tr>
<tr>
<td>CMIV,L</td>
<td>-0.0799</td>
</tr>
<tr>
<td>CCIV,B</td>
<td>-0.1418</td>
</tr>
<tr>
<td>ISD,B</td>
<td>-0.2347</td>
</tr>
<tr>
<td>ISD,G</td>
<td>-15.147</td>
</tr>
<tr>
<td>DIVE,G</td>
<td>-15.614</td>
</tr>
<tr>
<td>CCIV,G</td>
<td>-16.298</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)
*Levels not connected by same letter (A-E) are significantly different.

The next ANOVA was conducted to test changes in the average variances of the
six methods for corn and live cattle. Initial results (Table 7) show a failure to reject the
null hypothesis, demonstrated by a p-value of 0.1785 for corn, which indicates that there
were no significant differences in mean variation of errors among the three groups of
moneyness for corn. The p-value of 0.0273 (Table 8), for live cattle indicates that there
are differences in the parameters tested, however a p-value of 0.1506 for the interaction
of methods and groups leads to a failure to reject that there are significant differences
between the mean variances of groups for live cattle. The resulting effects test for the
variances of moneyness for live cattle show significant differences between the mean
variances between the groups and methods which is acceptable. However the importance
of this test is the analysis of the interaction of methods and groups, therefore these results
are ignored.

<p>| Table 7- Analysis of Variance, Corn Calls Variance, Moneyness |</p>
<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>17</td>
<td>3617410</td>
<td>212789</td>
<td>1.4323</td>
<td>0.1785</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>5348326</td>
<td>148565</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>53</td>
<td>8965736</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| Table 8- Analysis of Variance, Live Cattle Calls Variance, Moneyness |</p>
<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>17</td>
<td>2728306</td>
<td>160489</td>
<td>2.138</td>
<td>0.0273</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>2702353</td>
<td>75065</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>53</td>
<td>5430659</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The next market variable used to analyze approximation accuracy is time to
maturity, or the time till the expiration of the option. Figures 6 and 7 show each
approximation method’s percent error plotted with time to maturity. These plots show the
same patterns of how implied volatility changes as options approach expiration. As time
to maturity is further away, the errors are very small; however as time to maturity
approaches expiration, the errors become much more substantial. This is due to the fact
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 8- Corn Calls Percent Error and Time to Maturity
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 9- Live Cattle Calls Percent Error and Time to Maturity
that as an option nears expiration the time value diminishes; therefore, the value of the option depends more on intrinsic value, or the difference between the strike price and settle price. Stated in terms of the Black-Scholes model, this means that as the value of $\tau$ decreases, changes in option premiums will have a greater effect on the accuracy of approximating implied volatility. Again the graphs are divided into three distinctive groups of accuracy.

The Corrado and Miller method proves again to be a very accurate approximation, which is accurate even near to expiration. The next group consists of Bharadia et al. and Li (2005). Both of these methods have very smooth lines past about .2, or 20% of a year till expiration. Inside of 20% the errors begin to increase, up to about 50% error as the option nears expiration. The final group consists of Curtis and Carriker, Brenner and Subrahmanyam and Chargoy-Corona and Ibarra-Valdez. For this group, it also appears that the error smooths out as the time to maturity approaches a year. This is true by looking at the third group independently; however, if you compare it to the other groups there is more error, both positive and negative as time to maturity approaches a year. The third group appears to be the least accurate inside of 20% of a year with the errors ranging from -50% to over 150% error as the option approaches maturity. These graphs alone demonstrate that Bharadia et al., Li (2005) and Corrado and Miller would all provide accurate approximations if time to maturity is more than 20% of a year away from expiration. However if it is necessary to provide an approximation of implied volatility closer to expiration, the Corrado and Miller method should be used.
While these plots give a great illustration of accuracy over the life of an option, it is necessary to test accuracy using statistical tests. This is done in a similar manner to the statistical tests employed for testing moneyness. ANOVA was conducted to test the effect each group, less than and greater than .2, of time to maturity had on model accuracy.

Using the same null and alternative hypotheses as the test for moneyness, it was determined that the null hypothesis is rejected, p-value=0.0001, which indicates that there are differences in the mean percent errors (Table 9). It is therefore necessary to test which methods are significantly different. Results from Fishers Least Significant Difference test indicate that there are three methods which have significantly different means between the two groups of time to maturity; with L representing time to maturity less than 20% of a year and G representing time to maturity greater than 20% of a year (Table 10).

| Table 9- Analysis of Variance, Corn Calls Means, Time to Maturity |
|---------------------------------|-----|------------|---------|--------|-----------|
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Ratio** | **Prob>F** |
| Model      | 11   | 120.736   | 10.976  | 6.046  | 0.0001    |
| Error      | 24   | 43.5704  | 1.8154 |        |           |
| C. Total   | 35   | 164.307  |        |        |           |
Table 10- Corn Calls Means, Time to Maturity LS Means Differences

\[ \alpha=0.050 \quad t=2.0639 \]

<table>
<thead>
<tr>
<th>Level</th>
<th>Least Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIVE,L A</td>
<td>5.188033</td>
</tr>
<tr>
<td>BIV,L A B</td>
<td>4.739159</td>
</tr>
<tr>
<td>CCIV,L A B C</td>
<td>3.594917</td>
</tr>
<tr>
<td>ISD,L B C D</td>
<td>2.505385</td>
</tr>
<tr>
<td>LIIIV,L C D E</td>
<td>1.581165</td>
</tr>
<tr>
<td>ISD,G D E</td>
<td>1.308085</td>
</tr>
<tr>
<td>CCIV,G D E</td>
<td>0.810511</td>
</tr>
<tr>
<td>LIIIV,G D E</td>
<td>0.255919</td>
</tr>
<tr>
<td>DIVE,G D E</td>
<td>0.254714</td>
</tr>
<tr>
<td>BIV,G E</td>
<td>0.071553</td>
</tr>
<tr>
<td>CMIV,G E</td>
<td>-0.139067</td>
</tr>
<tr>
<td>CMIV,L E</td>
<td>-0.157021</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIIV represents the method provided by Li (2005)
*Levels not connected by same letter (A-E) are significantly different.

Surprisingly, these methods are Curtis and Carriker, Chargoy-Corona and Ibarra-Valdez and Bharadia et al. This result indicates that when analyzed together, the mean errors are significantly higher with a time to maturity of less than 20% of a year. It is also surprising that only three methods, rather than 5, have significantly different mean errors, as indicated by the graphs. Although Li (2005) and Bharadia et al. appear to have the exact same graph, when analyzed with each of the other methods, Bharadia et al. is significantly different between groups of time to maturity, where Li (2005) is not. Similarly, the methods developed by Curtis and Carriker, Chargoy-Corona and Ibarra-Valdez and Brenner and Subrahmanyam appear to have the very similar graphs; yet the Brenner and Subrahmanyam method is proven not to be significantly different for
maturities less than 20% versus maturities greater than 20% of a year. Therefore, these results indicate that the use of Brenner and Subrahmanyam, Li (2005) or Corrado and Miller will provide an approximation which is unaffected by time to maturity, when approximating for corn options. It is important to note that although the Brenner and Subrahmanyam method is unaffected by time to maturity, that it has been shown to be consistently less accurate than the other two methods.

The ANOVA results for the live cattle data set are shown in Table 11. The first result is that the null hypothesis is failed to be rejected, p-value=0.1071, denoting no significant difference in the mean percent errors of each of the methods. This means that no method is affected by time to maturity when analyzed together for the live cattle dataset.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>86.7652</td>
<td>7.88774</td>
<td>1.8172</td>
<td>0.1071</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>104.175</td>
<td>4.34061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>190.94</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In analyzing the ANOVA results for difference in mean variance there are further differences between the corn and live cattle data sets. For the corn data, a p-valued of 0.0005 leads to a rejection of the null hypothesis, or that there are differences in mean variance of the interaction between the methods and groups of time to maturity (Table 12). This result is confirmed in the effects test, with the interaction between methods and groups having an associated p-value of 0.0261 (Table 13). The corn results show that
Curtis and Carriker and Chargoy-Corona and Ibarra-Valdez approximations prove to have significantly different mean variances in percent error between the time to maturity groups, where each of the other methods are not significantly different (Table 14). This result confirms the lack in accuracy for Curtis and Carriker and Chargoy-Corona and Ibarra-Valdez as time to maturity is less than 20%. The ANOVA results for the live cattle data initially reject the null hypothesis, with a p-value of 0.0021 (Table 15). However, a p-value of 0.0755 from the effects test for the interaction of methods and groups leads to a failure to reject the null hypothesis. This indicates there are no significant differences in the mean variance of methods between the groups of time to maturity for live cattle, which confirms that the approximations are unaffected by time to maturity.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>9615961</td>
<td>874178</td>
<td>5.0446</td>
<td>0.0005</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>4158983</td>
<td>173291</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>1.4E+07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Nparm</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>5</td>
<td>5</td>
<td>3478359</td>
<td>4.0145</td>
<td>0.0087</td>
</tr>
<tr>
<td>Group</td>
<td>1</td>
<td>1</td>
<td>3433197</td>
<td>19.8117</td>
<td>0.0002</td>
</tr>
<tr>
<td>Method*Group</td>
<td>5</td>
<td>5</td>
<td>2704405</td>
<td>3.1212</td>
<td>0.0261</td>
</tr>
</tbody>
</table>
Table 14- Corn Calls Variance Time to Maturity LSMeans Differences

<table>
<thead>
<tr>
<th>Level</th>
<th>Least Sq Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIVE,L</td>
<td>1650.4981</td>
</tr>
<tr>
<td>CCIV,L</td>
<td>1053.2776</td>
</tr>
<tr>
<td>BSIV,L</td>
<td>783.4417</td>
</tr>
<tr>
<td>BIV,L</td>
<td>468.7876</td>
</tr>
<tr>
<td>CCIV,G</td>
<td>87.5986</td>
</tr>
<tr>
<td>DIVE,G</td>
<td>85.2569</td>
</tr>
<tr>
<td>BSIV,G</td>
<td>83.7485</td>
</tr>
<tr>
<td>LIIV,L</td>
<td>6.0396</td>
</tr>
<tr>
<td>CMIV,L</td>
<td>0.6451</td>
</tr>
<tr>
<td>BIV,G</td>
<td>0.1587</td>
</tr>
<tr>
<td>LIIV,G</td>
<td>0.1447</td>
</tr>
<tr>
<td>CMIV,G</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

*Levels not connected by same letter (A-D) are significantly different.

Table 15- Analysis of Variance, Live Cattle Calls Variance, Time to Maturity

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>1.2E+07</td>
<td>1100769</td>
<td>4.0175</td>
<td>0.0021</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>6575892</td>
<td>273996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>1.9E+07</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The third market condition to test method accuracy is the effect of changes in interest rates. Just as for moneyness and time to maturity, each of the approximation methods were potted with percent error versus interest rates. Figures 10 and 11 show the graphs for the corn data and live cattle data. By examining the graphs alone it is again easy to distinguish three different groups of methods. The first group is the Corrado and
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

Figure 10- Corn Calls Percent Errors and Interest Rates
DIVE represents the Direct Implied Volatility Estimate provided by Curtis and Carriker
ISD represents the Implied Standard Deviation method provided by Brenner and Subrahmanyam
CCIV represents the method provided by Chargoy-Corona and Ibarra-Valdez
CMIV represents the method provided by Corrado and Miller
BIV represents the method provided by Bharadia et al.
LIIV represents the method provided by Li (2005)

**Figure 11- Live Cattle Calls Percent Errors and Interest Rates**
Miller method which appears to essentially be a flat line, with only a few deviations to
the negative side of percent error. Next is the group of Bharadia et al. and Li (2005),
which are very similar and display what appear to be sporadic points of high positive
percent error over different interest rates. Finally, the graphs of the third group of Curtis
and Carriker, Brenner and Subrahmanyam and Chargoy-Corona and Ibarra-Valdez show
the same pattern of how different interest rates affect accuracy. By comparing each of
these plots, it appears that there is no affect on model accuracy as interest rates change;
therefore, the break to separate into two groups is placed at roughly the midpoint in
interest rates, or 5%.

Although it appears from these graphs that there is no change in accuracy given
different interest rates, it is necessary to confirm it. When analyzing the ANOVA results
for differences in means, the first observations are that the p-value of 0.0758 for corn and
p-value= 0.1852 for live cattle are both greater than the level of significance, 0.05 (Tables
16 and 17).

Table 16- Analysis of Variance, Corn Calls Means, Interest Rate

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>27.2367</td>
<td>2.47606</td>
<td>1.998</td>
<td>0.0758</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>29.7429</td>
<td>1.23929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>56.9796</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 17- Analysis of Variance, Live Cattle Calls Means, Interest Rate

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>16.8426</td>
<td>1.53114</td>
<td>1.5293</td>
<td>0.1852</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>24.0293</td>
<td>1.00122</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>40.8718</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This result indicates a failure to reject the null, that there are differences in the means between the two groups for all six of the approximation methods, for both the corn and live cattle data sets. The result proves that changes in interest rate have a negligible effect on the accuracy of all six approximations. When analyzing the ANOVA results for differences in mean variance, the initial p-values for corn and live cattle are 0.0004 and 0.0067, respectively (Tables 18 and 19). This result leads to a rejection of the null, that there are differences in the variances between the two groups of interest rates and the methods. However, p-values of 0.2645 for corn and 0.6078 for live cattle from the effects test prove a failure to reject the null that there are differences in mean variance for the interaction of methods and groups of interest rates. The low initial p-values from the ANOVA point to the strong differences in mean variance among the methods. Following the conclusion results from the graphical analysis, these statistical tests prove that accuracy is not significantly affected by different interest rates.
Table 18- Analysis of Variance, Corn Calls Variances, Interest Rate

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>390229</td>
<td>35475.4</td>
<td>5.1277</td>
<td>0.0004</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>166040</td>
<td>6918.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>556269</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 19- Analysis of Variance, Live Cattle Calls Variances, Interest Rate

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>409879</td>
<td>37261.7</td>
<td>3.3263</td>
<td>0.0067</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>268853</td>
<td>11202.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>35</td>
<td>678732</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Though all of these tests were conducted using nearest-to-the-money call options, the results are unchanged when averages of call and put options are considered. As mentioned in discussion of the Brenner and Subrahmanyam model, taking a straddle position will improve the accuracy of that particular model. Preliminary results suggest that this condition holds with the analysis done in this study.

With each of the methods for analyzing model accuracy presented here, there are clear and robust results which demonstrate that the Corrado and Miller model is the most accurate and will result in the best approximated value of implied volatility, followed by the Bharadia et al. and Li (2005) methods. The other three methods, Curtis and Carriker, Brenner and Subrahmanyam and Chagoý-Corona and Ibarra-Valdez are exceptional and accurate approximations; however the Corrado and Miller method consistently provides the closest value to the Black-Scholes implied volatility over various changing market conditions.
variables. Testing accuracy in the manner done in this study provides significant improvements to traditional measures of determining accuracy. In addition, the results have further reaching implications by providing evidence of accuracy tested across several variables of actual market data.
CHAPTER VI
SUMMARY AND CONCLUSION

Implied volatility provides information which is useful for not only investors, but farmers, producers, manufacturers and corporations. These market participants use implied volatility as a measure of price risk for hedging and speculation decisions. Because volatility is a constantly changing variable, there needs to be a simple and quick way to extract its value from the Black-Scholes option pricing model. Unfortunately, there is no closed form solution for the extraction of the implied volatility variable; therefore its value must be approximated. This study investigated the relative accuracy of six methods for approximating Black-Scholes implied volatility developed by Curtis and Carriker, Brenner and Subrahmanyam, Chargoy-Corona and Ibarra-Valdez, Bharadia et al., Li (2005) and Corrado and Miller. Each of these methods were tested and analyzed for accuracy using nearest to the money options over two data sets, corn and live cattle, spanning the years 1989 to 2008 and 1986 to 2008, respectively. This study focuses on accuracy for nearest-to-the-money options because the majority of traded options are concentrated at or near-the-money and several of the approximations were developed for at-the-money options. The aim of this study was to analyze the accuracy of these six methods using a variety of measures in order to determine which method most accurately approximates the Black-Scholes implied volatility.

Rather than following only the traditional measures of testing approximations for accuracy, this study considered several alternative ways for testing accuracy. In addition to analyzing mean errors and mean percent errors, other moments of the error
distributions such as variance and skewness were analyzed. Beyond this, measures of
goodness of fit, determined through an adjusted $R^2$, and accuracy over observed changes
in market variables, such as moneyness, time to maturity and interest rates, were
analyzed.

The error histograms provided the first comparison of methods for this study.
Both the corn and live cattle data sets revealed a clear distinction of three groups of
methods. The first group comprised of only the Corrado and Miller approximation. This
method was clearly the most accurate, followed by Bharadia et al. and Li (2005) in the
second group and finally the Curtis and Carriker, Brenner and Subrahmanyam, Chargoy-
Corona and Ibarra-Valdez methods in the third group.

The clear distinction of the three groups served as a starting point for comparison,
as well as an initial determination of accuracy. An adjusted $R^2$ value was found for each
approximation method to provide a standardized measure of accuracy both individually
and as a comparison to the other methods. This broke the methods into three distinctive
groups, identical to the ones found in the error histograms.

Next each of the approximation methods were tested for accuracy against the
three different market conditions of moneyness, time to maturity and changes in interest
rates. Analyzing approximation accuracy over these changing input variables was done
to ensure more robust and practical results. The three groups were still present, most
notably the difference between the Bharadia et al. and Li (2005) group, and the group
comprised of the Curtis and Carriker, Brenner and Subrahmanyam and Chargoy-Corona
and Ibarra-Valdez. The Corrado and Miller method proved to have no difference in
accuracy over any of the groups for each of the market conditions. This result is an astounding affirmation that the Corrado and Miller method for approximating implied volatility is not only a very close approximation to the true value, but that it is not affected by any change in market condition. Therefore, this approximation method should always be chosen given any possible market condition.

This study also demonstrated that although the Brenner and Subrahmanyam model is the starting point for many other approximation methods it ranks in the lowest accuracy group due to the assumptions of the authors which prevent the model from being accurate outside of exactly at-the-money options.

The methods based on the Brenner and Subrahmanyam method include Corrado and Miller, Bharadia et al. and Li (2005). When analyzing the groups of results, these three methods prove to be much more accurate than the method they stem from. There are several reasons for this; primarily that Brenner and Subrahmanyam assumes options which are exactly at-the-money. The underlying reason why the most accurate method, Corrado and Miller and the second group of methods Bharadia et al. and Li (2005) are proven to be most accurate is because they altered the Brenner and Subrahmanyam method to allow for near to the money options. By allowing for changes in option moneyness, most notably Li’s (2005) inclusion of a weighted moneyness variable, these methods are best for use with real market data.

The third group of methods, Brenner and Subrahmanyam, Curtis and Carriker and Chargoy-Corona and Ibarra-Valdez are not as accurate as the other methods for the same reason. Each of these methods was developed for at-the-money options, while the vast
majority of options are traded near-the-money, not at-the-money. These three methods became drastically less accurate with a marginal change in moneyness.

The study presented here clearly and accurately presents the most thorough study of the available approximation methods. It has been shown that with multiple comparisons of error, goodness of fit models and extensive statistical tests that the Corrado and Miller method stands out as the most accurate method for approximating implied volatility. Therefore, this method should be the primary method of approximation used for hedging. It is simple and can easily be calculated in spreadsheet form in order to make appropriate hedging decisions. This method is important because it will accurately provide a measure of price risk without the influence of moneyness, time to maturity or changes in interest rates; so that the most informed trading decision can be made.
APPENDICES
Appendix A

SAS Code Used to Merge Futures with Calls/Puts

data futures;
infile 'F:\New Folder\LC futures 1.9.10.csv' dlm=',' missover firstobs=2;
length date $ 10;
input Contract $ Date $ Settle;
run;
proc sort;
by Date Contract;
run;

Data Puts;
infile 'F:\New Folder\LC Puts 1.12.10.csv' dlm=',' missover firstobs=2;
Length date $ 10;
input Date $ Contract $ Strike Premium;
run;
Proc Sort;
By Date Contract;
run;

Data Combine;
Merge Puts Futures;
    by Date Contract;
diff=abs(strike-settle);
    if   diff eq .
    then delete;
run;

Data min;
set combine;
by date contract;
retain mindiff strikemin preatmin;
if   first.contract
then do;
    mindiff=diff;
    strikemin=strike;
    preatmin= premium;
    end;
else if   diff lt mindiff
then do;
    mindiff=diff;
    strikemin=strike;
preatmin= premium;
  end;
if last.contract then output;
run;
proc print data=min;
var mindiff strikemin date contract preatmin;
run; quit;
Appendix B

SAS Code Used to Merge Calls and Puts

data Calls;
infile 'C:\Users\Student\Documents\Implied Volatility\Data\LC\LC Calls Final.csv'
dlm=',' missover firstobs=2;
length date $10;
input Date $ Contract $ Settle StrikeC PremiumC Time;
run;

proc sort;
by Date  Contract ;
run;

data Puts;
infile 'C:\Users\Student\Documents\Implied Volatility\Data\LC\LC Puts Final.csv' dlm=','
missover firstobs=2;
Length date $10;
input Date $ Contract $ Settle StrikeP PremiumP Time;
run;

Proc Sort;
By Date  Contract;
run;

Data Combine;
Merge Calls Puts;
by Date  Contract ;
run;

proc print data=Combine;
var Date  Contract  Settle StrikeC StrikeP PremiumC PremiumP Time;
run;
quit;
Appendix C

SAS Code Used to Find a Benchmark Black-Scholes Implied Volatility for Call Options

data A;
length date $10;
infile "F:\New Folder\C5 1.21.10.csv" dlm=',' firstobs=2;
input date $ s x c p t r100 r ;
run;
proc sort; by date;

data A1;
set A;
do iv=.001 to .4 by .000001;
iv2=iv*iv;
d1=(log(s/x) + (iv2/2)*t)/(iv*sqrt(t));
d2=(log(s/x) - (iv2/2)*t)/(iv*sqrt(t));
cdf1=cdf('NORMAL',d1,0,1);
cdf2=cdf('NORMAL',d2,0,1);
cc=exp(-r*t)*((s*cdf1) - (x*cdf2));
diffc=abs(c-cc);
if diffc<.001 then output;
end;
proc sort; by diffc;
proc print data=A1 (obs=6) ;
  var date  s x c p t r100 r cc diffc iv   ;
run; quit;
proc sort; by date diffc ;
run;

Data mins (keep= date diffc pc iv);
Set B;
by date;
if first.date;
run;

Data C;
length date $10;
infile "E:\New Folder\Corn IV Merge 1.25.10.csv" dlm=',' firstobs=2;
input date $ contract $ s x c p t r100 r ;
run;
Proc Sort data=C ;
  by date;
run;
Data combine;
Merge mins C;
   By date;
run;

Proc sort Data=combine;
   by date;
run; quit;
Appendix D

SAS Code Used to Find a Benchmark Black-Scholes Implied Volatility for Put Options

data A;
length date $10;
infile "U:\Corn IV put data.csv" dlm=',' firstobs=2;
input date $ s x c p t r ;
run;
proc sort; by date;
data A1;
set A;
do iv=.001 to .9 by .0000001;
iv2=iv*iv;
d1=(log(s/x) + (iv2/2)*t)/(iv*sqrt(t));
d2=(log(s/x) - (iv2/2)*t)/(iv*sqrt(t));
cdf1=cdf('NORMAL',d1,0,1);
cdf2=cdf('NORMAL',d2,0,1);
pp=exp(-r*t)*((x*(-cdf2)) - (s*(-cdf1)));
diffp=abs(p-pp);
if diffp<.001 then output;
end;
proc sort; by date diffp ;
run; quit;

Data mins (keep= date diffp pp iv);
Set A1;
by date;
if first.date;
run;

Data C;
length date $10;
infile "U:\Corn IV put data.csv" dlm=',' firstobs=2;
input date $ s x c p t r ;
run;

Proc Sort data=C ;
    by date;
run;

Data combine;
Merge mins C;
    By date;
run;

Proc sort Data=combine;
  by date;
run; quit;
REFERENCES


