A Hybrid Option Pricing Model Using a Neural Network for Estimating Volatility

Sunisa Amornwattana, David Enke, and Cihan H. Dagli
Department of Engineering Management and Systems Engineering Department
Laboratory for Investment and Financial Engineering
Smart Engineering Systems Lab
Center for Intelligent Systems
University of Missouri – Rolla
1870 Miner Circle
Rolla, MO 65409-0370
573-341-4565
E-mail: sunisa@umr.edu, enke@umr.edu, dagli@umr.edu

ABSTRACT

The Black-Scholes model is the standard approach used for pricing financial options. However, although being theoretically strong, option prices valued by the model often differ from the prices observed in the financial markets. This paper applies a hybrid neural network which preprocesses financial input data for improving the estimation of option market prices. This model is comprised of two parts. The first part is a neural network developed to estimate volatility. The second part is an additional neural network developed to value the difference between the Black-Scholes model results and the actual market option prices. The resulting option price is then a summation between the Black-Scholes model and the network response. The hybrid system with a neural network for estimating volatility provides better performance in terms of pricing accuracy than either the Black-Scholes model with historical volatility, or the Black-Scholes model with volatility valued by the neural network.

Keywords: Neural Networks, Option Pricing, Hybrid Model, Black-Scholes

INTRODUCTION

Option pricing has been of great interest for many years. Predicting the correct value of an option can be beneficial for both hedgers and speculators. The Black-Scholes and Merton (1973) model has become a common method to value option prices and has been honored with the award of the Nobel Prize in economics. Nonetheless, even though it has been widely accepted, actual market prices often differ from the value obtained from the Black-Scholes model. Part of this difficulty in predicting the proper price results since the closed-form solution of the Black-Scholes model was developed under assumptions that do not hold in the real world and the model has been shown to exhibit systematic biases from observed option prices (Rubinstein, 1985). The Black-Scholes formula is derived based on various assumptions, such as geometric motion of stock price movements, European style options, constant interest rate, continuous trading without dividends or taxes applied to the stocks, and a frictionless market.

The problems of the conventional model have motivated many researchers to develop option pricing models using a number of methodologies to provide a more accurate and realistic price calculation. Merton (1973) developed a model in which dividends are continuous and proportional to the price of the underlying asset for European options. The model has been extended and improved by calculating values for American call options on dividend paying stocks (Roll, 1977; Geske, 1979; and Whaley, 1981). In 1979, Cox, Ross and Rubinstein developed a binomial tree methodology for American option pricing which is a simple discrete-time model.

In the last decade, nonparametric approachs using neural networks have been studied for option pricing and volatility estimation, including studies by Hutchinson (et. al, 1994) and Hanke (1999). Chen and Lee (1997) and Chidambaran, et al. (1998) applied genetic algorithms as an alternative approach for option pricing. Salchenberger (1994) and Karaali et al. (1997) and Malliaris (1994) applied neural networks for volatility prediction benchmarked with conventional techniques in time series, digital signal processing, historical volatility, and implied volatility. Schittenkopf and Dorffner (2001) developed a method of risk-neutral density extraction from option prices based on mixture density networks.

A hybrid approach to estimate the residuals between the actual option price and the Black-Scholes model has also been analyzed by various researchers. Gultekin et al. (1982) developed a model to estimate the residuals by linear regression. Jacquier and Jarrow (1996) selected a Bayesian approach to model the residuals. Lajbcygier, et al. (1997) applied a neural network to estimate the residuals, showing statistically and economincally significant performance. Lajbcygier and Connor (1997) improved the hybrid model using bootstrap methods to reduce bias in the hybrid model.

In this paper, an alternative and new hybrid between the Black-Scholes model and a nonparametric pricing model is developed and comprised of two parts. The first neural network estimates the implied volatility for each option by using the parameters in the Black-Scholes model, excluding volatility. The result from the first part, along with the other necessary parameters for the Black-Scholes model, are inputed to the second part, which is comprised of the hybrid model and the Black-Scholes model. The neural network at this stage is applied to estimate the difference between the Black-Scholes option price and the actual market price. The option price is then derived from the summation of the Black-Scholes model and the second neural network model. The method is applied to data from traded options on stocks selected from different sectors for diversification, all of which are listed in the Dow Jones Industrial Index. Numerical results are presented that show a superior pricing accuracy for the new hybrid model.

The remainder of this paper is organized into seven sections. Section 1 provides a summary of general option knowledge, option terminology, option pricing, and the Black-Scholes model. Section 2 summarizes volatility estimation methods, along with the volatility smile. Section 3 describes the neural network architecture which is applied in this work. Section 4 presents the experiment details: the architectures of the three models developed, dataset information, and implementation. Section 5 reports the performance in terms of prediction accuracy. Finally, Section 6 concludes the paper with a discussion of the results and directions for future research.

1. OPTIONS

1.1 OPTION FUNDAMENTALS

An option is the right to take an action in the future without any obligation. Therefore, the option holder must pay some money for this previlage, often called the option price or premium. There are two basic types of options, call options and put options. A call option gives the holder the right to buy the underlying asset by a certain date, called maturity or expiration date, for a certain price, called exercise price or strike price. A put option gives the holder the right to sell the underlying asset by a certain date (maturity date, expiration date), for a certain price (exercise price, strike price). Most traded options expire on the third Friday of the expiration month.

Options can generally be divided into two types, American options and European options, based on the flexibility of when the option can be exercised. American options give the right to the holder to exercise the option any time up to the expiration date. European options give the right to the holder to exercise the option only on the expiration date. As a result of the added flexibility, American options are priced higher than European options, given that all other factors remain the same. In U.S. markets, most exchange traded options are American (with the exception of some index options), with each contract being an agreement to buy or sell 100 shares of the underlying security. Nonetheless, the option prices shown in the market are per share, therefore, the trader needs to pay 100 times the quoted price to control 100 shares.

Options are mentioned as in-the-money, at-the-money, or out-of-the-money based on their relationship between the strike price and current underlying primitive asset. An in-the-money option is an option that gives the holder a positive cash flow if it was exercised immediately. For example, an in-the-money option for a call is one where the strike price is less than the current underlying asset market price. On the other hand, an in-the-money option for a put is one where the option strike price is greater than the current underlying asset market price. An at-the-money option is an option that gives the holder a zero cash flow if it was immediately exercised – the strike price and market price are identical. Out-of-the money options give the holder a negative payoff cash flow if the option is exercised immediately.

Every option contract is comprised of two positions: the trader who buys the option has taken the long position, while the trader who sells or writes the option has taken the short position. The buyer must pay premium or option price to enter the option position, while the seller will receive the cash premium upfront, with potential liabilities afterward. The writer is often required to make a security deposit or post margin to guarantee their capability to respect their obligations to sell or buy the underlying item; in other words, to ensure that the writer can pay any loss that results from the position going against them. The profit and loss for the writer and buyer are the reverse of each other, minus any transaction costs.

The option payoff is usually calculated without including the initial cost of the option. The profit must then consider the cost to enter the long position, or the premium received from the short position. For instance, consider an option with an initial cost P, strike price X, and final price S of the underlying asset at maturity. Then, the profit and payoff from a long positon in a European call are max(S-X-P, -P) and max(S-X, 0), respectively (see Figure 1a). The profit and payoff for a short position in the European call are –max(S-X-P, -P) = min(P-X-S, P) and –max(S-X, 0) = min(X-S, 0), respectively (see Figure 1b).

(Please insert Figure 1 here)

The concept is the same for put options. The profit and payoff from a long position in a European put are max(X-S-P, -P) and max(X-S, 0), respectively (see Figure 1c), while the profit and payoff from a short position are –max(X-S-P, -P) or min(P-S-X, P), and –max(X-S, 0) or min(S-X, 0), respectively (see Figure 1d).

1.2 OPTION PRICING

Accurate valuation of options is a significance issue for investors and practitioners in the financial markets. Researchers have devoted time to developing methodologies to price not only traditional stock options, but also other option problems, such as determining the value of employee stock options (Core and Guay, 2002; Bens, et. al, 2002) and real options for project valuation. There are seven main factors that affect the premium of an option:

1. The current price of the underlying asset
2. The strike price of the option

3. The time remaining until expiration or maturity
4. The volatility of the underlying asset price
5. The risk-free interest rate
6. The dividends expected for the underlying asset or stock during the option life
7. The type of option (call, put, or exotic)

These factors are essential and play a role in the option price. Fundamentally, each of these factors can be directly observed from the market, with the exception of the volatility of the stock price, which must be estimated. Therefore, the option prices in the market are often undervalued or overvalued depending on individual trader volatility estimation. The methodology of volatility estimation will be discussed in the next section.

In general, the option price is comprised of two parts, the intrinsic value and the time value. The intrinsic value is definded as the maximum of zero and the value the option would have if it was exercised immediately (the difference between the strike price and the current stock price). The intrinsic value for call and put option are max(S-X, 0) and max(X-S, 0), respectively. The time value is derived from the possibility of future favorable movements in the stock price. For example, consider that the price of a six-month call option with a strike price of \$15 is \$3, while the price of the underlying stock price is currently \$17. This gives an intrinsic value of the call option at \$17 - \$15 = \$2, while the time value is \$3 - \$2 = \$1. Normally, the holder of an American option would not exercise the option if the holder does not intend to hold the stock, but rather sell the option back to the market. The trader will receive only the intrinsic value if the option is exercised. If the option is sold to close the position, the trader will receive both the intrinsic value and time value.

1.3 THE BLACK-SCHOLES MODEL

A major breakthrough in stock option pricing occurred with the Black-Scholes model (Black and Scholes, 1973), leading to tremendous growth and success in financial engineering. The Black-Scholes model is derived for valuing European options on non-dividend-paying stocks. It was developed and based on various assumptions and ideal conditions, as given below:

1. The risk-free interest rate is known and constant
2. The distribution of stock prices is lognormal with a constant variance rate of stock return
3. There are no dividends during the life of the derivative
4. The option is of the European type
5. There are no transactions costs
6. The source of funds is available at the risk-free interest rate
7. The short selling of securities is permitted

Under these assumptions, the option price based on the Black-Scholes model can be calculated from five parameters: volatility of the underlying asset (σ), exercise price of the option (X), price of the underlying asset (S_0), number of days until the expriation date (t), and the risk-free interest rate (r). The Black-Scholes formulas for valuing call and put options at time zero are shown in equations (1) and (2), respectively:

$$C = S_0 N(d_1) - Xe^{-rt} N(d_2) \tag{1}$$

$$P = Xe^{-rt} N(-d_2) - S_0 N(-d_1) \tag{2}$$

where

$$d_1 = \frac{\ln(S_0 / X) + (r + \sigma^2 / 2)T}{\sigma\sqrt{T}} \tag{3}$$

$$d_2 = \frac{\ln(S_0 / X) + (r - \sigma^2 / 2)T}{\sigma\sqrt{T}} \tag{4}$$

and C : European call option price P : European put option price

S_0 : Stock price at time zero $\quad$ X : Strike price

T : Time to maturity $\quad$ σ : Stock price volatility

r : Continuously compounded risk-free rate

$N(x)$: Cumulative probability distribution function for a standardized normal distribution

2. VOLATILITY

2.1 VOLATILTIY ESTIMATION

Volatility measures the amount by which an underlying asset is expected to fluctuate in a given period of time. This factor significantly impacts the price of the option and heavily contributes to the time value of the option. A greater expectation in volatility tends to increase option prices. When the volatility of a stock is high, there is an inclination for the market to drive the option price higher (Van Horne, J.C., 1969). In the market there are two widely accepted and standard methods to estimate volatility: historical volatility and implied volatility.

2.1.1 Historical Volatility

Historical Volatility (HV) estimates volatility using historical stock price data which is normally observed at fixed intervals of time, such as daily, weekly, or monthly. Historical volatility is regularly called statistical volatility and is calculated as the standard deviation of a stock's return over a fixed period of time, such as 30, 60, 90, 120, or 365 days. Stock return is often defined as the natural logarithm of the closing prices between each interval of time. The historical volatility and return can be calculated as shown in Equations 5 and 6.

$$X_i = \ln\left(\frac{P_i}{P_{i-1}}\right) \qquad (5)$$

$$HV = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Xi-\overline{X})^2} \qquad (6)$$

where X_i : Return at end of i^{th} interval

P_i : Stock close price at end of i^{th} interval

$n+1$: Number of observations or number of observed days for daily basis

The procecure for determining the appropriate number of observations is not simple. Normally, more observations leads to increased accuracy, however, volatility does change over time and data from deep in the past may not be relevant for predicting the future. A basic rule is to set *n* equal to the number of days to which the volatility is to be applied (Hull, 2003).

When historical volatility is high, the stock has previously been displaying increased movement in price. When it is low, the volatility implies quiet trading or low movement in price. Benchmarking with the historical volatility of other stocks allows you to compare the volatility of a stock with the general market, or the stocks within a specific sector. In the same way, benchmarking the volatility of a stock with its own historical volatility allows one to estimate the increased or decreased movement of volatility. For example, if the 15-day historical volatility of a stock is 10% and the 60-day historical volatility is 30%, this implies that the stock has had a recent sharp decrease in volatility.

2.1.2 Implied Volatility

Implied Volatility is a measure of market expectations regarding the asset's future volatility. It is derived from an option pricing model such as the Black-Scholes model by adding all known variables and market option price into the formula, and then calculating the only remaining unknown variable, volatility. This calculated volatility value is called implied volatility.

In order to determine at any given point in time whether implied volatility is relatively high or low, it is important to compare the current value of volatility to the levels that existed in the past. Each option contract has a different value of implied volatility which changes over time, even though the Black-Scholes model suggests that every option implies the same volatility for the same underlying. In the market, the option price of deep-in-the-money and deep-out-of-the-money options are relatively insensitive to volatility. Therefore, the implied volatilities derived from these options tend to lack reliability.

2.1.3 Volatility Smiles

Most derivative markets exhibit persistent patterns of implied volatility calculated from actual market option prices that vary over the strike price and time to maturity, instead of being constant as assumed by the Black-Scholes model. The relationship between the implied volatility of an option as a function of its strike price is known as a volatility smile. In some markets, such as currency options, the volatility smile shows that deep out-of-the-money and deep in-the-money options are priced by the market at prices that are higher than theoretically forecasted by the lognormally distributed Black-Scholes model (volatility smile is shown in Figure 2(a)). In other markets, such as for individual stocks and stock indices, the implied volatility pattern is normally different and often referred to as a volatility skew, such that the volatility tends to decrease as the strike price increases, as shown in Figure 2(b). The reason behind this phenomena can be explained by leverage. While an underlying asset price declines, the leverage of the underlying asset increases. Thus, the volatility of underlying asset increases as there is more possibility for the price to decrease. In the same manner, when an underlying asset price increases, the leverage of the underlying asset decreases as a result of decreasing volatility. For this situation, there is less possibility for the price to increase.

(Please insert Figure 2 here)

3. NEURAL NETWORK ARCHITECTURE

For this research the General Regression Neural Network (GRNN) (Specht, 1991) is applied. The GRNN architecture is chosen because it provides comparable accuracy to feedforward networks using the backpropagation learning algorithm with substantially less training time and convergence to the optimal surface. There are many advantages when using the GRNN as compared to other nonlinear regression methods. First, it obtains the estimate in a single pass, therefore, no iterative training is needed. Second, the estimate converges to the conditional mean regression surfaces and is bounded by the minimum and maximum of the observations. The estimate cannot converge to the poor outcomes corresponding to local minima of the error criterion. Lastly, the network also performs well under noisy environments when given sufficient data.

The GRNN is a supervised feedforward network, subsuming the radial basis-function method. It is based on nonlinear regression theory and computes its output using a variation of the “nearest-neighbor” approach. The forecast value for an input vector is a weighted average of the outputs in the training sample. The closer an input vector in the training sample, the larger the weight of its corresponding output. The GRNN is usually applied for function approximation.

The GRNN is a regression network utilizing a probabilistic model between the independent (input or x) and the dependent (output or y) variable. Assume that $f(x,y)$ represents the known joint continuous probability density function of x and y, where X is a measured value of the variable *x*. The expected conditional mean of *y* given *x* is shown in Equation 7. A probability estimator $\hat{f}(X,Y)$ is based on sample values X^i and Y^i of the ramdom variables *x* and *y*, as proposed by Parzen (1962) and shown in Equation 8. As a result, the expected value of y ($\hat{Y}(X)$) is shown in Equation 9.

$$E[y \mid X] = \frac{\int_{-\infty}^{\infty} y f(X,y)dy}{\int_{-\infty}^{\infty} f(X,y)dy} \tag{7}$$

$$\hat{f}(X,Y) = \frac{1}{(2\pi)^{(p+1)/2}\sigma^{(p+1)}} \cdot \frac{1}{n}\sum_{i=1}^{n}\{\exp\left[-\frac{(X-X^i)^T(X-X^i)}{2\sigma^2}\right] . \exp\left[-\frac{(Y-Y^i)^2}{2\sigma^2}\right]\} \tag{8}$$

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i \exp(-\frac{D_i^2}{2\sigma^2})}{\sum_{i=1}^{n} \exp(-\frac{D_i^2}{2\sigma^2})} \tag{9}$$

where $D_i^2 = (X - X_i)^T (X - X_i)$ and σ is a smoothing parameter of the radial basis function or kernel width. A small value of σ means that the output contribution is mainly derived from the nearby cells. A large value of σ means that the output is dervied from both the nearby and more distant cells. The kernel width should be smaller than the typical distance between the input vectors to fit the data very closely. The larger the kernel width, the smoother the function approximation.

The GRNN architecture is shown in Figure 3. The topology of a GRNN is comprised of four layers: input layer, pattern layer, summation layer, and output layer. The function of the input layer is to pass the input vector to all the units in the pattern layer. The neurons of pattern layer are created to compute the kernel function. The summation layer has two neurons to compute the numerator and denominator. The output layer provides a result by dividing the values of these two neurons.

(Please insert Figure 3 here)

4. EXPERIMENT

4.1 GENERAL HYBRID NEURAL NETWORK

A hybrid neural network model for options pricing was introduced by Lajbcygier et al. (1997). This approach is a combination of the standard Black-Scholes model and a neural network. The concept is to use the conventional Black-Scholes model as the core calculation and then apply a neural network to value the difference between the actual market price and the Black-Scholes model value. The inputs of this model are the risk-free interest rate (r), time to maturity (T), volatility (σ), and the ratio of the stock price and strike price (S/X). The volatility is obtained from the historical volatility that is calculated from past market data. The target is the difference between the Black-Scholes model and the ratio of call option price and strike price (C/X), as shown in Figure 4.

(Please insert Figure 4 here)

This methodology effectively learns the differences between the actual market price and the theoretically estimated value from the Black-Scholes model, and also allows the neural network to be trained on these differences.

4.2 BENCHMARK MODELS

For this research, two models are benchmarked with the proposed new hybrid model. The first model (HV-BS) is the simplest model since it is easy to set up and calculate, as shown in Figure 5. The volatility provided as input to the Black-Scholes

(BS) model is calculated by way of historical volatility (HV). Past prices (P) are also used as input. The prices are converted into a return format which is then used to calculate the standard deviation of the stock's return (σ). Next, the standard deviation of stock's return, along with the other input variables, including stock price (S), strike price (X), risk-free rate (r), and time to maturity (T) are provided as input to the Black-Scholes model to price the option. The disadvantage of this model is that there is merely one calculated volatility value for a single trading day. In fact, the option price for different strike price and maturity are valued by different implied volatilities, even for the same underlying asset. Not surprisingly, this model provides unsatisfactory results compared to the actual market price. Nonetheless, the reason to apply this basic model is to observe the rough capability of the standard Black-Scholes model and further investigate the possible development of an advanced alternative model.

(Please insert Figure 5 here)

The second benchmarked model incorporates the advantages of GRNN to capture the characteristic of each option volatility more precisely. Instead of single variable, the developed GRNN is multivariate with four inputs: stock price (S), strike price (X), risk-free rate (r), and time to maturity (T), which are the standard inputs for the Black-Scholes model. The GRNN is applied to generate volatility for every combination of inputs. Thus, the different strike prices, different times to maturity, and the value of the other inputs leads to different volatility values, regardless the same underlying asset. The generated volatility from the GRNN is combined with the four original inputs sent to the GRNN to provide five inputs to the Black-Scholes model for pricing the option. The benefit of the GRNN for this model is that the volatility for different strike prices and times to maturity are not constant, as highlighted and discussed in Section 2.1.3. The GRNN can provide different values of volatility for different combinations of input. It should provide an improvement in terms of pricing accuracy and allow the characteristics of volatility skew to be modeled.

(Please insert Figure 6 here)

4.3 NEW HYBRID NEURAL NETWORK MODEL

A hybrid neural network model was also developed and tested. For the model, the stock price and strike price are applied individually, instead of using their ratio as previously applied. Futhermore, the actual market option price replaces the ratio of the call option price and the strike price to obtain the actual difference between the Black-Scholes model and the market option price.

The major difference between the proposed model and the original model is an additional neural network module for improving the volatility estimation for feeding to the original hybrid model. Historical volatility estimates are typically poor since they tend to equally weigh each data point. In fact, the more recent data points will often provide more influence on the predicted volatility with regard to directing a more short-term volatility estimation. Several techniques are often applied to solve this problem, such as an exponentially weighted moving average (EWMA) (Robert, 1959; Hunter, 1986; Lucas, and Saccucci, 1990), autoregressive conditional heteroscedasticity (ARCH) proposed by Engle (1982), generalized autoregressive conditional heteroscedasticity (GARCH) (Bollerslev, 1986; Engle and Mezrich, 1996), and the sophisticated GARCH family models, such as EGARCH (Nelson, 1991). Some studies have shown that neural networks for estimating volatility outperform other commonly used techniques (Karaali et al., 1997; Refenes and Holt, 2001).

The new hybrid model (NN-HB) is comprised of two parts, as shown in Figure 7. The first part of the model applies a GRNN to predict volatility. This part has four inputs: stock price (S), strike price (X), time to maturity (T), and the risk-free interest rate (r). The implied volatility (σ) is the target output of this model. The system then combines the predicted volatility obtained from the first network with the first four parameters to provide five inputs to the second part of the model. The second part applies these inputs to obtain a value from the Black-Scholes model. The same five inputs are also applied to a second GRNN which is used to predict the price and to value the discrepancy of the Black-Scholes model. For the model, the difference between the Black-Scholes model and actual price are represented as targets. The estimated call or put option prices (C, P) are obtained from the combination of the Black-Scholes model and the second GRNN.

(Please insert Figure 7 here)

4.4 DATA SET

This research utilizes the call option data from five different primitive stock assets, including Coca-Cola (KO), McDonald’s Corporation (MCD), Boeing (BA), Citigroup (C), and International Business Machines (IBM) to train and test the model. These securities were chosen because of their liquidity and stability. These stocks, selected from different sectors for diversification, are all listed in the Dow Jones Industrial Average Index (DJIA) and often labeled as blue-chip stocks due to their relatively stable businesses and large capitalization. These stocks are also listed in the

Standard and Poor's 500 (S&P500). Table I provides the symbol, company name, sector, and industry of these stocks.

(Please insert Table I here)

The daily data of these stocks were either collected or calculated for a total of 27,496 records from July 1, 2002 to October 15, 2002. Even though the data seems like a short period of time, each stock offers various option types based on strike price and time to maturity for each trading day. The data were divided into two periods. The first period included the variable values for the three calendar month period from July 1, 2002 to September 30, 2002 for a total of 23,319 records. The second period contained the variable values for the fifthteen calendar day period from October 1, 2002 to October 15, 2002, for a total of 4,177 records. The first period was applied to train the network and the latter period was reserved for out-of-sample evaluation.

This study is focused on the closing prices for each day, which is used to compare the results of the system for pricing accuracy. Table II provides the number of training and testing records for each scenario. Each underlying asset is tested and trained under three situations, in-the-money, out-of-the-money, and all (in-the-money and out-of-the-money) to investigate their difference characteristic in terms of moneyness. This experiment focuses on the 90 day historical volatility since the networks are trained under a three-month period, allowing each model to be compared during the same time frame. The option data was gathered from http://www.ivolatility.com/, a paid subscribtion website. In addition, the risk-free rate information is gained from http://www.federalreserve.gov/. The volatility data used to train the network was derived from the past data using the implied volatility concept.

(Please insert Table II here)

4.5 IMPLEMENTATION

Data preparation is one of the significant steps necessary for obtaining a satisfying result when modeling a system. This is especially true for neural networks, where network training can be made more efficient if normalization steps are performed on the network inputs and targets. Normalization helps prevent attributes with initially large ranges from outweighing attributes with initially smaller ranges.

There are many normalization processes, such as min-max, decimal scaling, and the z-score method. Min-max normalizes data into a linear transformation which encounters an "out of bounds" problem in situations where the future input data falls beyond the original data range. The decimal scaling normalizes the data by moving the decimal point of values but still preserves most of the original character of the value. This study uses a z-score method to normalize the data set. The value for an input A is normalized based on the mean and standard deviation of A. A value v of A is normalized to v' by Equation (10).

$$v' = \frac{v - \overline{A}}{\sigma_A} \tag{10}$$

where $\overline{A}$: mean of input A v : original value

σ_A : standard deviation of input A v' : normalized value

After the normalization process, inputs and targets have zero mean and unity standard deviation. There are advantages and disadvantages among these normalization methods. However, the selected z-score method is best suited to this dataset since upper the limits (maximum) and lower limits (minimum) are unknown (Han and Kamber, 2001). After this preprocessing step and the training of the network, a reversing of the normalization is needed to transform the output back to the original data pattern.

There are several performance measurements that can be used for neural network testing. This research is focused on estimation accuracy. An accuracy measure is often defined in terms of the difference between the actual (desired) and the estimated value.

Let ($\hat{y}_1, \hat{y}_2, ..., \hat{y}_N$) denotes the estimated values and ($y_1, y_2, ..., y_N$) represents the actual values, where N is the sample size. The mean squared error (MSE) and mean absolute error (MAE) are applied to compare the performance, as shown in Equations 11 and 12, respectively.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \overline{y_i})^2 \qquad (11)$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}| y_i - \overline{y_i} | \qquad (12)$$

5. RESULTS

Once again, the research tested three models. The first model uses the 90 day Historical Volatility with the Black-Scholes model (HV-BS). The second model attempts to improve the volatility estimate by using a Neural Network to predict volatility before providing this value to the Black-Scholes model (NN-BS). The third model is the proposed new Hybrid model (NN-HB). The results of the MSE and MAE for the testing dataset of the all (in-the-money and out-of-the-money), in-the-money, and out-of-the-money options are shown in Tables III, IV, and V, respectively.

As seen from Table III, the MAE and MSE for each stock decreases as each system is modified and improved. For example, MAE of KO with HV-BS system (0.4329) is about three times higher than the MAE of NN-BS (0.1464), and about four times higher than the MAE of NN-HB (0.1025). Between the NN-BS (0.1464) and NN-

HB (0.1025) models, MAE decreases approximately 40 percent. The MSE result also follows the same direction as the MAE. The all (in-the-money and out-of-the-money), in-the-money, and out-of-the-money results converge to the same conclusion with decreasing error.

(Please insert Table III here)

(Please insert Table IV here)

(Please insert Table V here)

The results show that the Black-Scholes model with volatility calculated from the neural network estimates the option price significantly better than when historical volatility is used as an input to the Black-Scholes model. Futhermore, the neural network trained on the difference between the Black-Scholes model and the neural network result presents an improvement in option pricing estimation for the tested data set and demonstrates the usefulness of combining these two methods.

Normally for parametric modeling, when two samples need to be compared concerning their mean value for some variable of interest, a t-test is applied, whereas an analysis of vaiance is used for multiple samples. For nonparametric modeling, testing must be performed using different tools. In this paper, nonparametric tests are performed for the MAE and MSE results to test whether the three models (HV-BS, NN-BS, and NN-HB) provide statistically significant different results. These tests are based on the following scores: Wilcoxon (also known as the Kruskal-Wallis test), median, Van der Waeden, and Savage (each under the null hypothesis that there is no difference in location for MSE and MAE results among the three models). The p-values for each test of MAE and MSE results are shown in Tables VI and VII, respectively.

(Please insert Table VI here)

(Please insert Table VII here)

The p-values are less than 0.05 in all situations except the median scores for the in-the-money case. This leads to rejection of the null hypothesis at a 0.05 significant level. Nonetheless, in general these tests support the conclusion that there are differences in location for the MAE and MSE among the three models.

6. CONCLUSIONS AND FUTURE WORK

The research demonstrates the benefits of using neural networks for option pricing since neural networks can be used to learn from the past data, predict volatility, and account for non-ideal conditions that cause problems for the Black-Scholes model. The hybrid neural network has been shown to improve the performance of the Black-Scholes model to capture deviations hidden by the strict assumptions of the Black-Scholes model. The results here are highlighted by the dramatic and statistically significant decrease of the MAE and MSE results, at least for in- and out-of-the-money options. When combining a neural network for estimating volatility with the hybrid neural network, superior performance is achieved.

Further research should consider other potential influence variables, such as dividend payment. Multivariate nonstationary volatility estimation methodologies from the ARCH family, such as GARCH, should be considered for benchmarking during future investigations. For the option pricing model described for this research, the basic Black-Scholes model is applied. The various modified Black-Scholes models can also be considered for implementation with other statistical techniques to improve the option pricing results.

REFERENCES

Bens, D. A., Nagar, V., and Franco Wong, M.H. (2002) "Real Investment Implications of Employee Stock Option Exercises", *Journal of Accounting Research*, 40, No. 2, 359-393.

Black, F. and Scholes, M. (1973) "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-654.

Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Chen, S. and Lee, W. (1997) "Option Pricing with Genetic Algorithms: A Second Report", *Neural Network, International Conference*, 1, 21-25.

Chidambaran, N.K., Jevons, Lee, C.H., and Triguerors, Joaquin R. (1998) "An Adaptive Evolutionary Approach to Option Pricing Via Genetic Programming", *Genetic Programming: Proceedings of the third annual conference*, 38-41.

Core, J., and Guay, W. (2002) "Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility", *Journal of Accounting Research*, 40, No. 3, 613-630.

Cox, J., Ross, S.A., and Rubinstein, M. (1979) "Option Pricing: a Simplified Approach", *Journal of Financial Economics*, 7, 229-264.

Engle, R. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation", *Econometrica*, 50, 987-1008.

Engle, R., and Mezrich, J. (1996) "GARCH for Groups", *RISK*, August, 36-40.

Geske, R. (1979) "A Note on an Analytic Valuation Formula for Unprotected American Call Options on Stock with Known Dividends", *Journal of Financial Economics*, 7, 375-380.

Gultekin, N., Rogalski, R., and Tinic, S. (1982) “Option Pricing Model Estimates, Some Empirical Results”, *Financial Management*, Spring, 58-69.

Han, J. and Kamber, M. (2001) *Data Mining: Concepts and Techniques* (Morgan Kaufmann, San Francisco, CA).

Hanke, M. (1999) “Neural Networks Versus Black-Scholes: An Empirical Comparison of the Pricing Accuracy of two Fundamentally Different Option Pricing Methods”, *Journal of Computational Intelligent Finance*, 5, 26-34.

Hull, J.C. (2003) *Options, Futures, and Other Derivatives* (Prentice Hall, New Jersey).

Hunter, J. S. (1986) “The exponentially weighted moving average”, *Journal of Quality Technology*, 18, No. 4, 203-210.

Hutchinson, J.M., Lo, A. W. and Poggio, T. (1994) “A Nonparametric Approach to Pricing and Hedging Derivatives Securities via Learning Networks”, *Journal of Finance*, 49, 851-889.

Jacquier, E. and Jarrow, R. (1996) “Model Error in Contingent Claim Models (Dynamic Evaluation)” Rodney L. White Center for Financial Research Working Papers, *Wharton School Rodney L. White Center for Financial Research*, 7-96.

Karaali, O., Edelberg, W., and Higgin, J. (1997) “Modeling Volatility Derivatives Using Neural Networks”, *Computational Intelligence for Financial Engineering*, 280-286.

Lajbcygier, P. and Connor, J. (1997) “Improved Option Pricing Using Artificial Neural Networks and Bootstrap Methods”, *International Journal of Neural System*, 8, No. 4, 457-471.

Lajbcygier, P., Flitman, A., Swan, A., and Hyndman, R. (1997) “The Pricing and Trading of Options Using a Hybrid Neural Network Model with Historical Volatility”, *Neurovest Journal*, 5, No. 1, 27-41.

Lucas, J. M. and Saccucci, M. S. (1990) “Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements”, *Technometrics*, 32, No. 1, 1-29.

Malliaris, M. and Salchenberger, L. (1994) “Do-Ahead Replaces Run-Time: a Neural Network Forecasts Options Volatility”, *Artificial Intelligence for Applications, Proceeding of the Tenth Conferenc*, 480-481.

Merton, R. (1973) “Theory of Rational Option Pricing”, *Bell Journal of Economics*, 4, 141-183.

Nelson, D. (1991) “Conditional Heteroskedasticity in Asset Returns: A New Approach”, *Econometrica*, 59, 347-370

Parzen, E. (1962) “On Estimation of a Probability Density Fnction and Mode”, *Annals of Mathematical Statistics*, 33, 1065-1076.

Refenes, A.P. and Holt, W. (2001) “Forecasting Volatility with Neural Regression: A Contribution to Model Adequacy”, *IEEE Transactions on Neural Networks*, 12, No. 4.

Roberts, S. W. (1959) “Control Chart Tests Based on Geometric Moving Averages”, *Technometrics*, 1, 239 -250.

Roll, R. (1977) “An Analytical Formula for Unprotected American Call Options on Stocks with Known Dividends”, *Journal of Financial Economics*, 5, 251-258.

Rubinstein, M. (1985) “Nonparametric tests of alternative option pricing models using all reported trades ad quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978”, *Journal of Finance*, 40, 455-480.

Schittenkopf, C. and Dorffner, G. (2001) “Risk-Neutral Density Extraction from Option Prices: Improved Pricing with Mixture Density Networks”, *Neural Network, IEEE Transactions*, 12, 716-725.

Specht, D.F. (1991) “A General Regression Neural Network”, *IEEE Transactions on Neural Networks*, 2, 568-576.

Van Horne, James C. (1969) “Warrant Valuation in Relation to Volatility and Opportunity Costs”, *Industrial Management Review*, 10, 19-23.

Whaley, R. (1981) “On the Valuation of American Call Options on Stock with Known Dividends”, *Journal of Financial Economics*, 9, 207-211.

Table and Figure Name

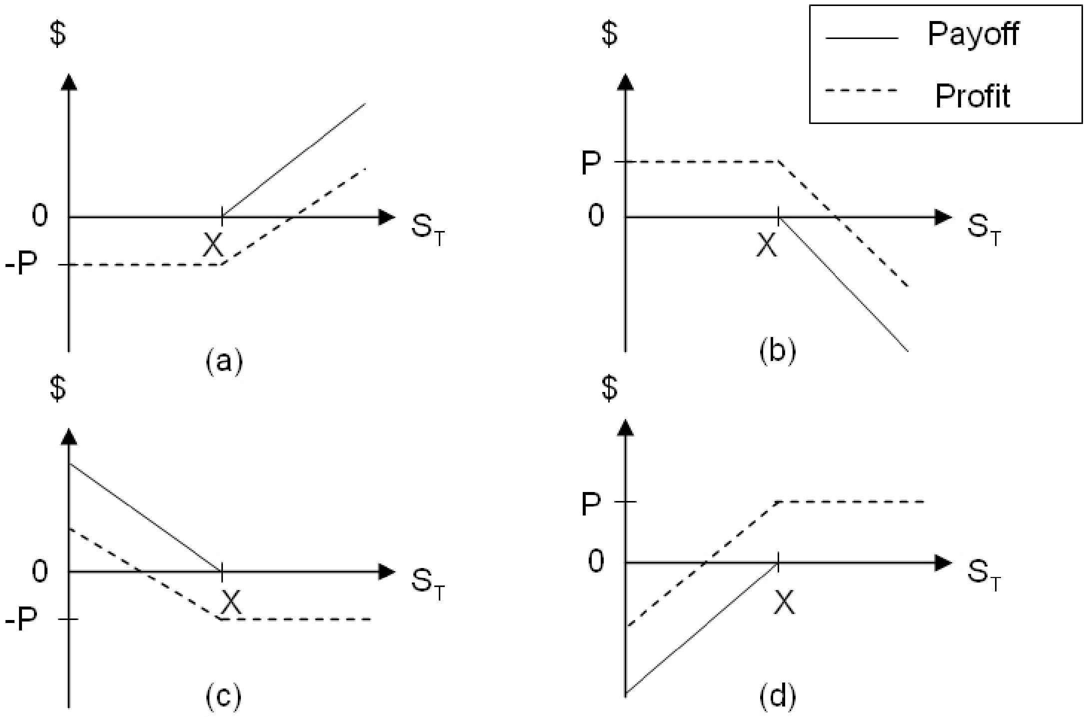

Figure 1

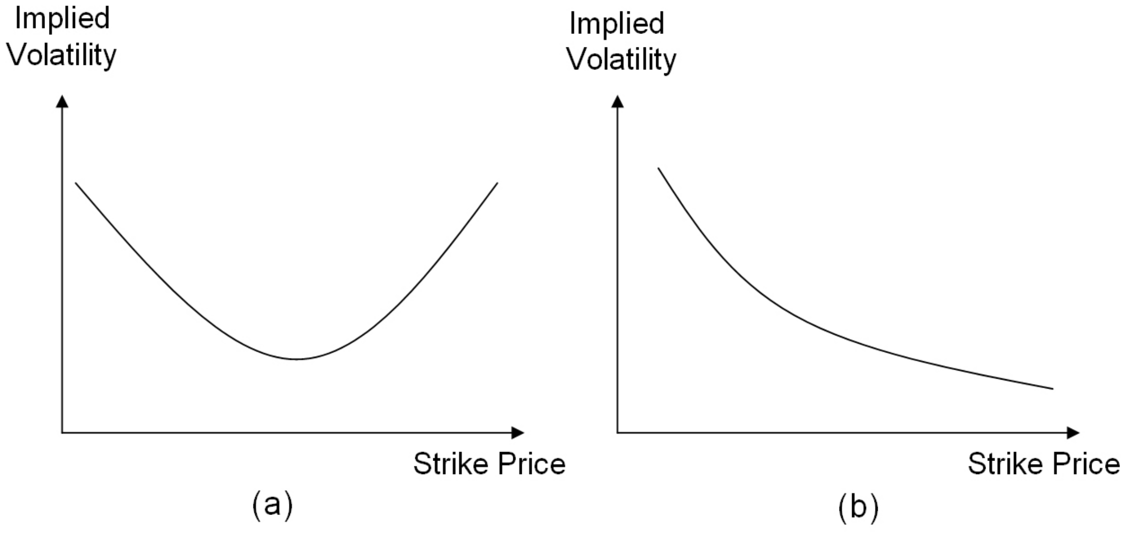
Implied
Volatility
Strike Price
(a)
Implied
Volatility
Strike Price
(b)

Figure 2

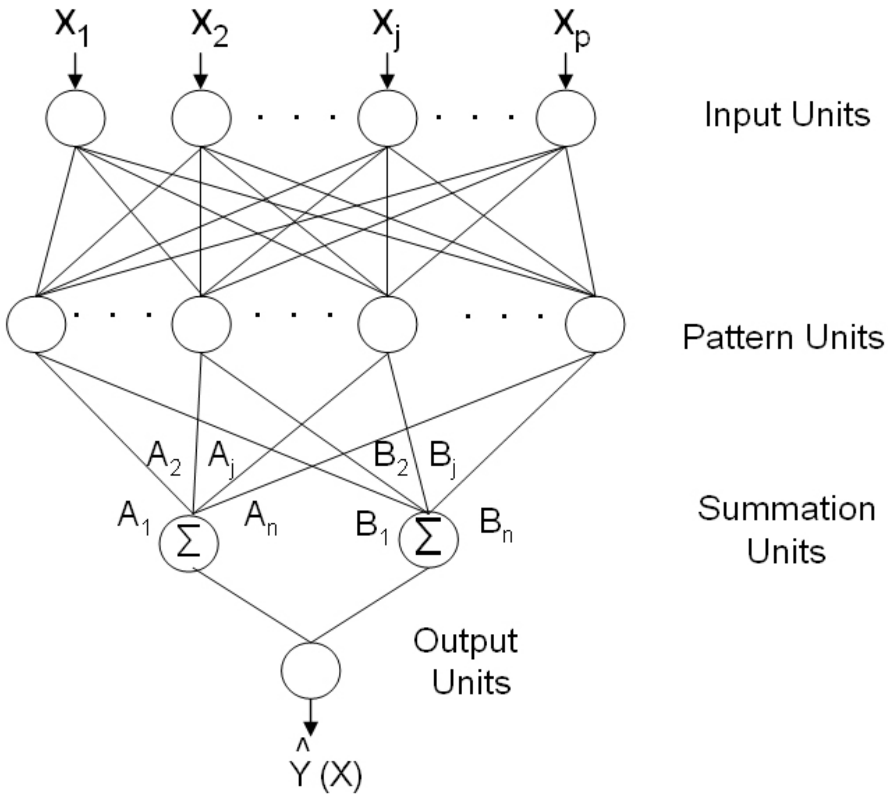
x1
x2
xj
xp
Input Units
Pattern Units
A2
Aj
B2
Bj
A1
An
B1
Bn
Summation Units
Output Units
Y (X)

Figure 3

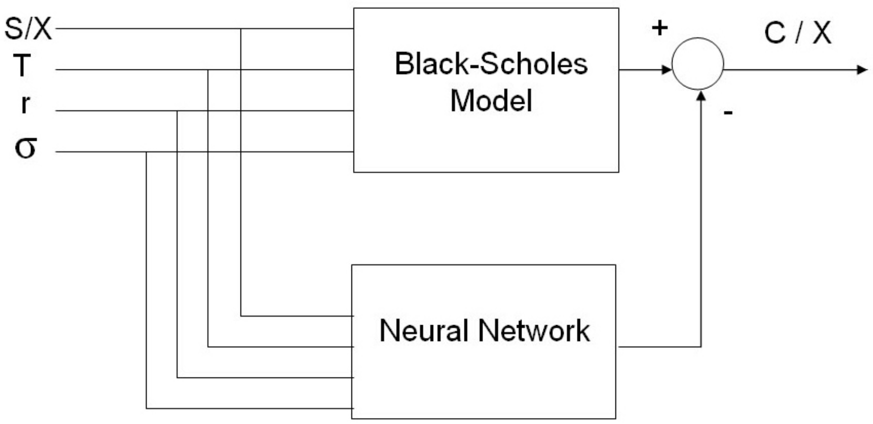

Figure 4

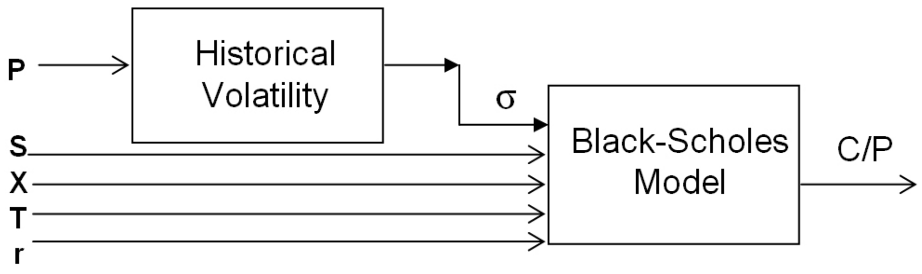

Figure 5

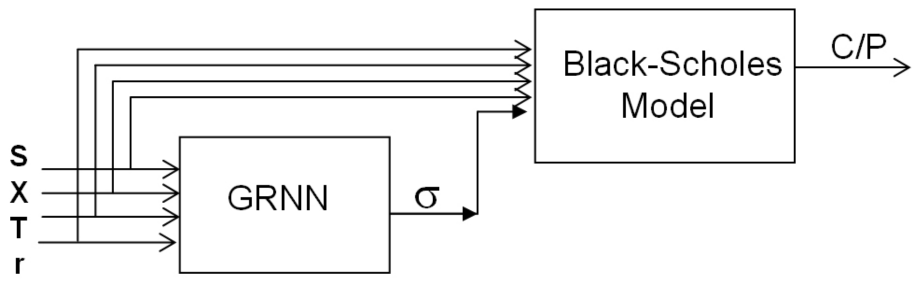

Figure 6

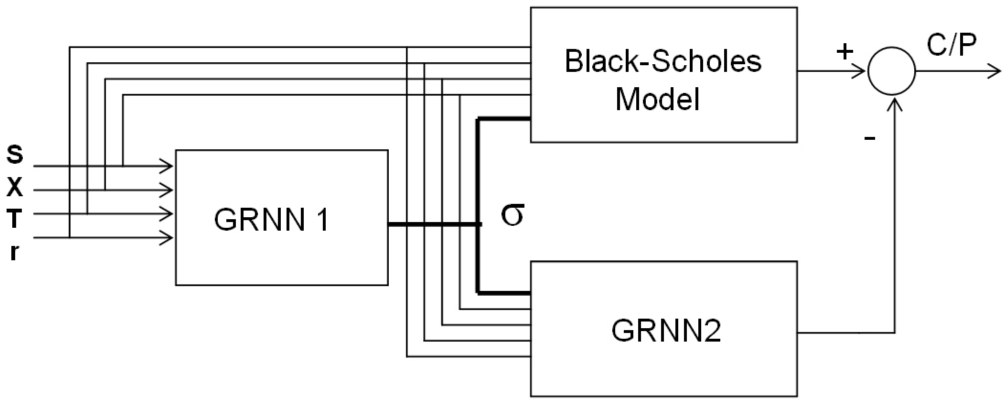

Figure 7

Symbol	Company Name	Sector	Industry
KO	Coca-Cola Co (Coke)	Consumer Non-Cyclical	Beverages (Non-Alcoholic)
MCD	McDonald’s Corp	Services	Restaurants
BA	The Boeing Co	Capital Goods	Aerospace & Defense
C	Citigroup Inc	Financial	Money Center Banks
IBM	International Business Machine Corp	Technology	Computer Hardware

Table I

Stock	All (in and out)		In-the-money		Out-of-the-money	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**
KO	3526	649	1676	345	1850	304
MCD	2678	579	915	164	1763	415
BA	4985	887	1755	194	3230	693
C	5678	944	2087	386	3591	558
IBM	6452	1118	2399	292	4053	826

Table II

Stock	**MAE**			**MSE**		
	HV-BS	**NN-BS**	**NN-HB**	**HV-BS**	**NN-BS**	**NN-HB**
KO	0.4329	0.1464	0.1025	0.3553	0.0388	0.0306
MCD	0.1846	0.0554	0.0329	0.0845	0.0049	0.0028
BA	0.2638	0.0812	0.0575	0.1695	0.0197	0.0057
C	1.1378	0.1644	0.0810	3.5470	0.0580	0.0150
IBM	0.7146	0.1451	0.0844	1.2099	0.0493	0.0184

Table III

Stock	**MAE**			**MSE**		
	HV-BS	**NN-BS**	**NN-HB**	**HV-BS**	**NN-BS**	**NN-HB**
KO	0.2328	0.2233	0.1427	0.1279	0.0993	0.0535
MCD	0.1132	0.0757	0.0479	0.0222	0.0089	0.0064
BA	0.3317	0.1805	0.0992	0.1756	0.0571	0.0262
C	0.9424	0.2332	0.0749	1.9558	0.1076	0.0128
IBM	0.5478	0.2120	0.0896	0.7930	0.0847	0.0213

Table IV

Stock	**MAE**			**MSE**		
	HV-BS	**NN-BS**	**NN-HB**	**HV-BS**	**NN-BS**	**NN-HB**
KO	0.4654	0.1252	0.1050	0.6134	0.0257	0.0229
MCD	0.2095	0.0482	0.0290	0.1092	0.0042	0.0020
BA	0.2996	0.0872	0.0834	0.2215	0.0156	0.0161
C	1.2343	0.1257	0.0959	4.6478	0.0319	0.0205
IBM	0.7494	0.1171	0.0782	1.3572	0.0329	0.0165

Table V

Data	**Wilcoxon Scores**	**Median Scores**	**Van der Waerden Scores**	**Savage Scores**
All (in & out)	0.0044	0.0035	0.0051	0.0055
In-the-money	0.0254	0.1512	0.0231	0.0300
Out-of-the-money	0.0039	0.0035	0.0048	0.0047

Table VI

Data	Wilcoxon Scores	Median Scores	Van der Waerden Scores	Savage Scores
All (in & out)	0.0039	0.0035	0.0046	0.0053
In-the-money	0.0259	0.1512	0.0231	0.0239
Out-of-the-money	0.0028	0.0035	00037	0.0029

Table VII