A Regression Model to Investigate the Performance of Black-Scholes using Macroeconomic Predictors

Timothy A. Smith[#1], Ersoy Subasi[*2], Aliraza M Rattansi[#3]

[#] *Embry Riddle Aeronautical University*
600 S. Clyde Morris Blvd. Daytona Beach, Fl 32114 U.S.A

Abstract— **As it is well known an option is defined as the right to buy sell a certain asset, thus, one can look at the purchase of an option as a bet on the financial instrument under consideration. Now while the evaluation of options is a completely different mathematical topic than the prediction of future stock prices, there is some relationship between the two. It is worthy to note that henceforth we will only consider options that have a given fixed expiration time T, i.e. we restrict the discussion to the so called European options. Now, for a simple illustration of the relationship between true stock prices and options let us consider the following situation: if at the begging of January the S&P index is valued at $1277 and then at the end of December of the same year the price of the index became $1400 then the fair price of the option to buy this in January would be $123, assuming T=1. This is a fair price because if the holder purchases the option for $122 or less then he or she gains while if the holder purchases the option for $124 or more then the bank wins while $123 is neutral for both parties. As one can see from this simple illustration predicting the fair price of an option is directly related to predicting the value of the stock price in a future time T.**

In recent years great work has been put into both developing and applying the famous Black Scholes model [1], which is used to predict the fair price of a European option. However, in this model the volatility, σ , is required to be taken as constant and that is often not a valid assumption. It has been illustrated that during time periods of rapidly changing volatility the Black Scholes model does not perform well, however, in times of non rapidly changing volatility the Black Scholes model performs extremely well. An underlying issue is answering the question of what exactly the volatility of the market is? Moreover, many people have defined measures, but many of them are not applicable. In this paper we investigate the S&P 500 index and the commonly noted volatility for this index, VIX, which is not applicable to use within the Black Scholes as volatility. This is due to the fact that the VIX itself is an index which is speculated upon; in fact the VIX itself is both traded and sold in many ways. This is clearly not usable for the σ input as the foundations of the Black Scholes model, consistent with most mathematical models, are based on modelling a phenomenon based on logical behaviour as opposed to human erratic behaviour. A good example of this is how the VIX or S&P may change rapidly due to one word a single news reporter says, regardless if it was a truthful statement or not.

In this study we seek to remedy this by using a simple comparison technique to identify in real time when the Black Scholes model truly is violating the assumptions of constant volatility with volatility defined more along the lines as market deviation from reality. To do this we create a linear regression model to predict the value of the S&P 500 using input values of macroeconomic measures such as GDP (Gross Domestic Product),M (total money aggregate), PPI (Producer Price). A correlation is proven here empirically and it is suggested that one can use the regression model as a red flag to caution the user in real time that at those points Black Scholes performance is suspect, with high or rapidly changing volatility implied. At these times one should use the Black Scholes with extreme caution, perhaps a regime switching approach, or look to alternate prediction methods often just common sense possibly just seeking secure assets.

Keywords— **Partial differential equations, parabolic, stochastic, financial mathematics. AMS classification: 35K10**

I. Introduction

As it is well known [2] the commonly used stock market prediction models, such as the famous Black Scholes Stochastic Partial Differential Equation,

$$\frac{\partial X}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 X}{\partial S^2} + rS\frac{\partial X}{\partial S} - rX = 0$$

used to predict the price of a derivative X as a function of time and initial stock price have demonstrated some limitation during events of recent years. For example, in non-rapidly changing times of volatility, the famous Black Scholes Formula, using the standard notation for the normal distribution $N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\tau^2/2} d\tau$ along with the call option with a maturity date T, strike price k and risk free interest rate r, will accurately predict the fair price of a stock given its price today as x_0 through the formula

$$x_0 N\left(\frac{ln\left(\frac{x_0}{k}\right)+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)-ke^{-rT}N\left(\frac{ln\left(\frac{x_0}{k}\right)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

But the performance in rapidly changing times of volatility their performance has been debated. This model requires that the volatility σ be imputed in the same nature as the riske-free interest rate as a fixed real constant. This assumption may not be practical as while interest rates realistically do change slowly with time the volatility often changed drastically with time, and this has been the center of much debate as of late with many suggestions being made. In what follows we do not attempt to remodel nor create any new stock market prediction formula, but rather we pose a method to better predict when the volatility is truly out of line.

As it is well known in late 2008 significant nearly unprecedented behaviour occurred in financial markets, a 50% crash was followed by a full recovery within less than half a decade. This behaviour was not quite in line with principles as the peak of the foreclosure crisis began several years earlier in 2006 yet the market still continues to grow. If the market was truly following fundamental economic principles one would have expected the market to cool in 2006 then perhaps lightly decline in 2007/8 as opposed to the actual events that happened in 2007/8 of a steady strong increase followed by a sharp correction of nearly 50% in \ just over a year. The correction is shown in Table 1.

TABLE I

S&P 500 & VIX VALUES FOR 2008; MAY – DEC 1ST TRADING DAY MONTH

S&P500	VIX
1409	18.88
1385	17.83
1285	23.65
1260	23.59
1278	21.99
1161	39.81
966	59.85
816	68.51

One can demonstrate that on either sides of this bubble the Black Scholes model does not accurately predict the future value of S&P 500. We believe that this is not due to flaws in the model, but rather that it comes from the issue of the difficultly to truly measure volatility using current measures. For example the VIX should have been high at the beginning of the year when the S&P was at 1447, but rather the VIX did not jump until late in the year which was too late and well after the market began to fall. We seek a model to identify these red flag times by looking at how far away the index is pulling away from a linear predictor built on macroeconomic predictors, with the underlying hypothesis being that the economy will always drive the market and win in the long run with market regressing back to it.

For illustration let us perform a calculation mid-year with : x_0 =1285 as the initial value, a minimal interest of r=0.05 and T=0.5 for a 6 month investment with k=1349 for a gain equivalent to double the fixed rate computed using simple interest on the given time period. Doing so we obtain we obtain the fair price

$$x_0 N\left(\frac{ln\left(\frac{x_0}{k}\right)+\left(r+\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)-ke^{-rT}N\left(\frac{ln\left(\frac{x_0}{k}\right)+\left(r-\frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\right)$$

Now, the issue is what should be inputted for volatility, σ? We perform these calculations three times: the first being the value of volatility taken as exactly the value of the VIX at the same time period from which we obtained x_0. Then the second being the value of volatility drastically reduced to a minimum value of, say 5%, and then the final calculations with volatility being drastically increased to a maximum value of, say 40% . With routine computations one can see that the different scenarios yield extremely different outputs (117,26,203) . As previously noted it is not feasible to use any VIX index as that is a human driven measure. So we seek an alternate approach by constructing a long term regression model with the hypothesis that if the value is outside of a certain range (i.e. error tolerance on $y - \hat{y}$) then the volatility should be adjusted accordingly. Namely, if the residuals are minimal one would expect the second computation to be appropriate; on the other hand if the residuals are major the third computation would be appropriate.

II. STATEMENT OF REGRESSION AND COMPUTATIONS

As it is well known various organizations, both privative banks and government institutions, provide many indicators that explain the current state of various components of the economy. If we consider these as variables x_1,x_2,x_3... one can hypothesis that for a given large market stock that if the market is following fundamentals the value of this stock y should follow a simple linear regression line with the multiple input variables

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots$$

For our immediate purposes we choose only the input variables x_1 = PPI producer price index and x_2= GDP gross domestic product and x_3 = CPI consumer price index and x_4 = M the total money aggregate with y as the value of the S&P 500 index at the same time of the reporting values x_1,x_2,x_3 and x_4 were from. We obtained approximately 30 years of monthly data for all of these values; it is worthy to note that some figures were only published quarterly so we just repeated those as the same value for each month within the quarter to complete the data. Our analyses show that this model does not have common hypothesis violations that arise such as collinearity. Hence, the addition of any other variable would not improve the model. It is worthy to note that other work has been done to justify a similar model in the past [4] has been used by a large trading firm.

For investigation, we split our model into two segments: the first being 1984-1996 which we dubbed as the stable market times and the second being 1996-2012 for the chaotic times. For the later we obtained

$$y_i = 864.873663 + 240.7970562 x_{1i} + 637.5742501 x_{2i} + 162.0869638 x_{3i} - 707.0030155 x_{4i}$$

through the time range i=1 being Jan 1996 up to i=192 being Jan 2012. The input variables x_1 through x_4 along with the predictor y were normalized with the standard normal transformation of $(x_{ji}-\bar{x}_j) / s_j$ where s_j is the standard deviation of the variable under consideration. We consider the precent differences between y and ŷ where a sample is given from 2008 with the VIX values at the same time, as shown in Table II.

TABLE III

S&P 500 & Ŷ REGRESSION PREDICTED VALUES FOR 2008

y	ŷ	& change	VIX
1447	1275	-13	23.17
1395	1259	-10	24.02
1331	1256	-6	26.54
1370	1248	-9	22.68
1409	1261	-11	18.88
1385	1264	-9	17.83
1285	1266	-2	23.65
1260	1256	0	22.57
1278	1237	-3	21.99
1161	1184	2	39.81
966	1146	17	59.89
816	1073	27	68.51

We then consider the difference between the predicted Black Scholes and true nature of fair price (S&P in 1 year future minus current S&P) in the same time periods, as shown in Table III

TABLE IIIII

BLACK SCHOLES PREDICTED VALUES FOR 2008, JAN-MAR & OCT-DEC

(BS)-((t+12)S&P – $x_0)_{\sigma=0.05}$	**(BS)-((t+12)S&P – $x_0)_{\sigma=0.40}$**	**VIX**
-508.1	-313	23.17
-562.7	-474.8	24.02
-623.7	-444.4	26.54
|	|	|
-124.6	32.8	39.81
102.5	235	59.89
292.9	357.6	68.51

As one can see the results are as expected. The value in the first three months of the year, in either case, is extremely negative which implies that the market was overvalued and nowhere to go but down. On the other hand the value in the last three months of the year, in either case, is extremely positive which implies that the market was undervalued and nowhere to go but up. However, one must remember that these calculation cannot be done in real time as it is not possible to know the last value in the computation (value of the S&P in one year future time). However, a correlation is noted between this and the output of the regression line: the first few months it's residual as a precent difference was above 10% negative while in the latter few months it's residual as a precent difference was above 10% positive, hence empirically proving the hypothesis.

III. Acknowledgements and Data Note

The author's would like to thank Aliraza for his time running the various codes and analysing the data. We truly thank you for taking the time to do all of this as a side project for just fun while also working full time on completing your Master's Degree. Again we thank you and wish you the best with your future studies and career, shot for the stars!

A Matlab code used to analyse the data and compute Black Scholes is below which can be easily adjusted for further investigation. The regression model was built directly from the full data set in SPSS and analysis was done in excel type spread sheets. The data and or code is available upon request through email smitht1 at erau dot edu.

```
clear all
        clc
        p = xlsread(Quaterly.xlsx)
        [n,m]=size(p)
        Xo=p(:,1)
        T=p(:,5)
        K=p(:,3)
        r=p(:,7)
        Sigma=p(:,6)

for i=1:size(Xo)

    A(i,1)=(log((Xo(i,1))/(K(i,1))))+((r(i,1))+(0.5*(Sigma(i,1))*(Sigma(i,1)))*(T(i,1)))/((Sigma(i,1))*(sqrt(T(i,1))))

    B(i,1)=(log((Xo(i,1))/(K(i,1))))+((r(i,1))-(0.5*(Sigma(i,1))*(Sigma(i,1)))*(T(i,1)))/((Sigma(i,1))*(sqrt(T(i,1))))

    C(i,1)=normcdf((A(i,1)),0,1)

    D(i,1)=normcdf((B(i,1)),0,1)

    E(i,1)=((Xo(i,1))*(C(i,1)))

    F(i,1)=((K(i,1))*(exp((-r(i,1)*T(i,1))))*(D(i,1)))

    FuturePrice(i,1)=E(i,1)-F(i,1)

end
```

References

[1] Black, Fisher and Scholes, Myron. "*The Pricing of Options and Corporate Liabilities,*" Journal of Political Economy 81 (3), 1973.
[2] Stewart, Ian. "*The Mathematical equation that caused the banks to crash*," The Observer 2/12/12, 2012.
[3] Park, Sam. "*Reducing the Noise in Forecasting the SP 500*," Wentworth, 2005.