

An Economic Regression Model to Predict Market Movements

Timothy A. Smith^{#1}, Andrew Hawkins²

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

Abstract— In finance, multiple linear regression models are frequently used to determine the value of an asset based on its underlying traits. We built a regression model to predict the value of the S&P 500 based on economic indicators of gross domestic product, money supply, produce price and consumer price indices. Correlation between the error in this regression model and the S&P's volatility index (VIX) provides an efficient way to predict when large changes in the price of the S&P 500 may occur. As the true value of the S&P 500 deviates from the predicted value, obtained by the regression model, a growth in volatility can be seen that implies models like the Black-Scholes will be less reliable. During these periods of changing volatility we suggest that the user apply a regime switching approach and/or seek alternative prediction methods.

Keywords— Partial differential equations, regression analysis, stochastic, financial mathematics.
AMS classification: 35K10

I. INTRODUCTION

It has been observed that the commonly used stock market prediction models, such as the famous Black Scholes Stochastic Partial Differential Equation [1],

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

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 as

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

However, the performance in rapidly changing times of volatility performance has been debated. This model requires that the volatility σ be imputed in the same nature as the risk-free interest rate as a fixed real constant, which may not be a valid assumption.

Economic indicators allow predictions of the future performance of an economy to be drawn. In a broad sense, they consist of statistics that are derived from economic activity. They include measures of the unemployment rate, housing starts, inflation, and many others. Although there are many types of economic indicators we will focus on four major indicators, those being the consumer price index (CPI), producer price index (PPI), gross domestic product (GDP), and money supply (M). Prior study has been done to show that adding extra variables does not improve the model [3]. With historical values of these indicators, we are able to construct a linear regression model that can predict the value of the S&P 500 index. Further, we will examine when large errors in our model occur and correlate these errors with the VIX index.

A multiple linear regression model takes many independent variables and determines a linear relationship that best approximates each data point. In our case, the independent variables will be the CPI, PPI, GDP, and M; the dependent variable will be the actual value of the S&P 500. Large relative error in the regression model implies that the market has entered a period of increased volatility. Other models, such as the Black-Scholes, do not perform well during periods of changing volatility [2].

II. REGRESSION MODEL

After gathering monthly data, we can construct a MLR model from January 1990 to July 2013 by following a regression model of the form

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots$$

Because we want to compare the residues of the model with the VIX we must restrict the dates of analysis to the number of months of data we have for the VIX. The calculation methodology changed for the VIX in 1990 and, due to wanting to keep the data the same, we are restricted from going back further. A Matlab script was used to calculate the model and residuals. Figure 1 shows a plot of the actual S&P 500 and the predicted value of the S&P 500. The governing equation of the model over this time frame and the one shown in Figure 1 is

$$y_i = 1948.181 + 0.287x_1 + 16.581x_2 + 9.891x_3 + 0.074x_4.$$

In this model $i = 1, \dots, n$ corresponds with the month in question out of n total months, with x_1 =GDP and x_2 =PPI and x_3 =CPI and x_4 =M. This model was able to yield a R^2 value of 0.76, which shows a relatively strong relationship between the dependent variable and the independent variables.

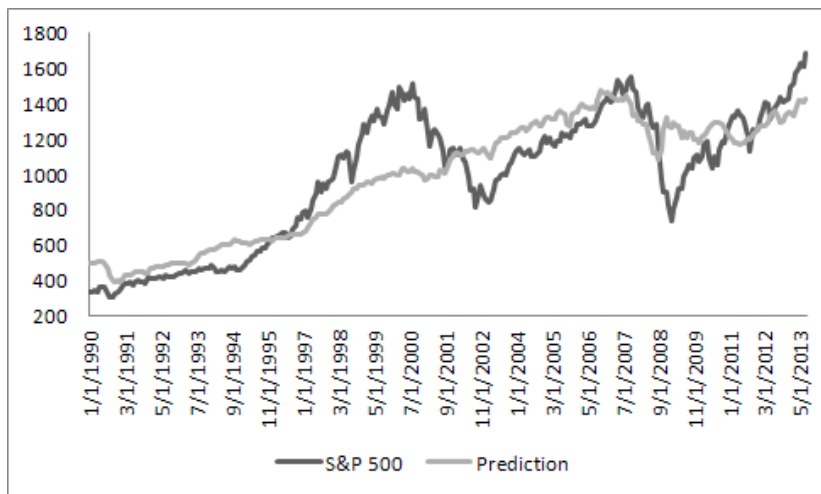


Fig. 1 This figure depicts the actual value of the S&P 500 along with the predicted value through the MLR model. As the error in the prediction becomes large we can correlate that error with the VIX.

One of the main points of this study is to determine if there is a significant correlation between the VIX and the residues in this model. We can determine whether a correlation between these two data sets exists by performing the usual hypothesis testing for regression. Doing so we obtain the test statistics as

$$t = \frac{0.44\sqrt{283-1}}{\sqrt{1-0.44^2}} \approx 8.$$

This hypothesis testing procedure also yielded a p-value of 7.18×10^{-15} . So either testing at the $\alpha = 0.05$ significance level, or interpreting such a small p-value we can state that there is significant evidence of a relationship between the residuals in our model and the VIX. This is a very practical result and, as previously mentioned, can be used as a trading strategy to hedge from significant market deviations (losses).

III. CONCLUSIONS

We have shown that, from the MLR model we created for the S&P 500, there is a correlation between the residuals and the VIX. It is important to state that this conclusion is dependent upon the many parameters contributing to our model. For example, if the range of data were to be shortened (less observations) or if other or fewer economic indicators were to be used, the conclusion that a correlation exists could not necessarily be drawn. So, for the specific model presented in this paper, a statistically significant correlation is found between

the residuals and the VIX. Due to this information we can conclude that the VIX is a good predictor of coming errors in the regression model, as well as other models. There is the possibility of an even greater correlation between the two studied sets than was measured. When the VIX takes relative maximum values it is almost always due to a significant downward slide in the S&P 500. If the predicted value is less than the actual S&P 500 value then a large drop in the actual S&P 500 would reduce the error yet increase the VIX. The final observation from this study is that, because of their correlation, large increases in the VIX are good predictors to when conventional modelling techniques, like MLR models, will incur error. During these periods we suggest using either a regime-switching approach or alternative models that account for large changes in volatility; moreover, simple conclusions can also be drawn to determine prime buy or sell points.

REFERENCES

- [1] Black, F. and Scholes, M. "*The Pricing of Options and Corporate Liabilities*," Journal of Political Economy 81 (3), 1973.
- [2] Moyaert, T. & Petitjean, M. "The performance of popular stochastic volatility option pricing models during the subprime crisis." Applied Financial Economics. 21(14), 2011.
- [3] Smith, T. "A regression model to investigate the performance of the Black-Sholes using macroeconomic predictors." International Journal of Mathematics Trends and Technology, 2013.