Using Multiple Linear Regression to Estimate Volatility in the Stock Market

Alex J. Caligiuri, Embry-Riddle Aeronautical University '18

Introduction:
- Currently no great model for volatility to forecast stock market performance
- VIX- Chicago Board Options Exchange Implied Volatility Index used today
- VIX back-solves the Black Scholes model for volatility using current options price
- Fails to show any warning to potential crash
- Created multiple linear regression model for S&P 500 price
- Deterministic variables were GDP, Money Supply, and Unemployment Index

Materials and Methods:
- First task- update current data set and add any possible new variables
- Originally 8 independent variables before inter-correlation analysis
- Check for cross correlation between independent variables, then run regression
- Removed least deterministic variable for each analysis until desired tolerance was reached
- Final regression model compared with S&P 500 from 1991-2018 (Figure 1)
- New volatility models created using residuals, compared with VIX (Figure 2)
- Input new volatilities and VIX into Black-Scholes formula in MATLAB for 1-Year S&P 500 options, compared current price + option premium with future S&P price (Figure 3)
- Black-Scholes Inputs- current S&P 500 price as spot price, 10-Year Treasury yield as risk-free rate, 1 year as time until expiry, and nearest tenth dollar for the S&P as strike price.

Results and Key Takeaways:
- None of the three volatilities could accurately predict S&P 500 one year in advance by adding the option premium to the current S&P 500 price.
- VIX was always a large value, new models for volatility only large when an "unexpected" recession based off of poor fundamentals in the economy was imminent, like 2002 crash.
- Volatility model spikes for new models are explained through it's calculation using the differences between model S&P 500 and actual S&P 500 values
- Residuals not large for 2008 because market moved downwards with the economy
- While new volatility models are not perfect, can be used to de-risk portfolio before financial crash

Acknowledgements:
- Dr. Timothy A. Smith, Department of Mathematics, Embry-Riddle Aeronautical University
- Dr. Jayendra Gokhale, Department of Finance and Economics, Embry-Riddle Aeronautical University

Further Information: Contact Alex J. Caligiuri, Caligiua97@gmail.com

Abstract:
This project entails an in-depth analysis on the current mathematical methods used to calculate volatility in the stock market like the Black-Scholes Stochastic Partial Differential Equation (PDE). Expanding upon the faculty mentor's previously published work, a new method for calculating volatility using multiple linear regression on key macroeconomic factors was tested in predicting the price for the Standard and Poor's 500 Index (S&P 500) for a given month. The results from the new volatility model were then compared with the well-known Chicago Board Options Exchange Volatility Index (VIX) since the VIX's creation in 1991. While the new models for volatility could not be used to adequately price S&P 500 options for the future, it performed similar to VIX over time, and while the value for VIX was always high, the values for new volatility only spiked significantly when there was a potential crash in the market caused by weak fundamentals, like the dot com crash in 2001. With this information in hand, one could use a disparity between the regression model and the actual S&P 500 value to forecast a future crash caused by poor fundamentals in the market.

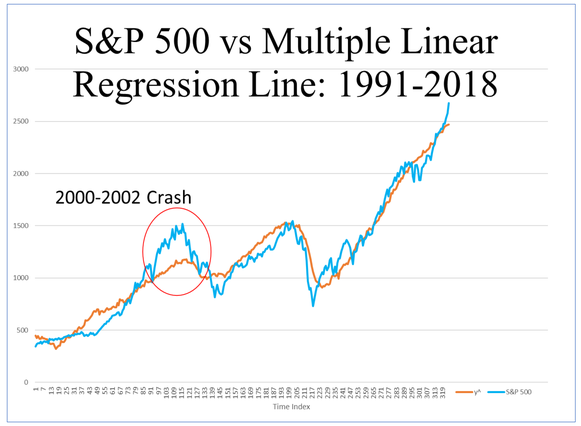

Figure 1. *S&P 500 vs Multiple Linear Regression Line: 1991-2018*

Regression Line Equation and Statistics

Model S&P 500 = (198.84*GDP) + (298.79*Money Supply) - (210.38*Unemployment Index) +1188.90

Correlation Coefficient (R) = 0.9694
Coefficient of Determination (R-Squared) = 0.9397
F-Score = 1664.048
F-Significance = 8.734E -195

Note: All data was normalized before regression.

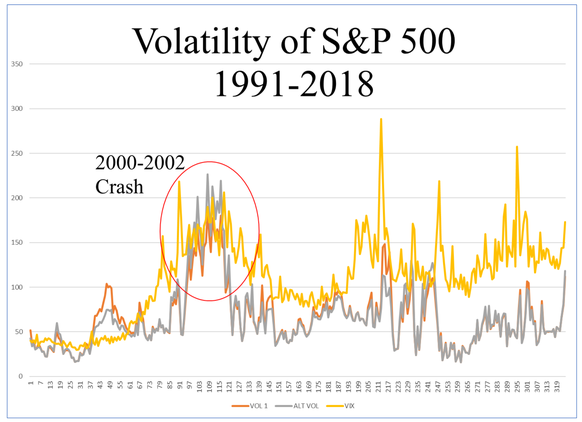

Figure 2. *Volatility of S&P 500 1991-2018.*

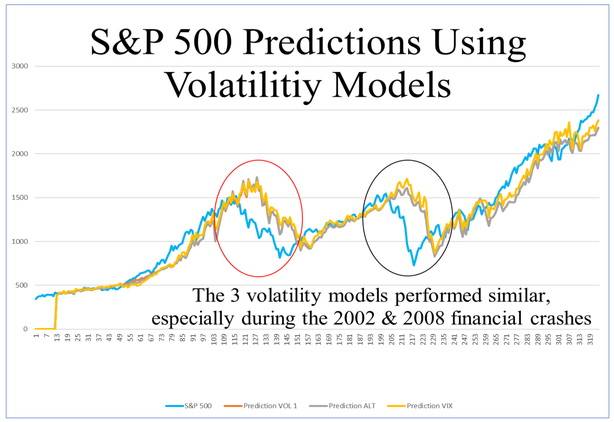

Figure 3. *S&P 500 Predictions Using Volatility Models.*