**Realistic Expectations: Implied Volatility and the S&P 500 Index**

**Abstract Draft:**

Realistic expectations of growth in the stock market are traditionally measured using the Capital Asset Pricing Model, which uses inputs for estimated growth. It’s commonly used by is often touted as having little correlation to the actual performance of a stock or index. This paper focuses on using the implied volatility in the derivatives market as an estimate for the future range of returns in the stock market index- the S&P 500. The results of this project concluded that implied volatility is correlated with the range of future returns and tends to overestimate the relationship of actual return.

**Background:**  
In modern portfolio theory, the relationship between risk and return in a stock or index is constantly being measured and updated against its peers and can be visually plotted as a point on a graph similar to Figure 1. Efficient stocks exist on the curve labeled as the efficient frontier- having the most efficient trade off of risk and investment volatility. This concept translates into a reality of mutual funds who are constantly adjusting their portfolios to make sure the risks they taking their portfolio are appropriate with the returns that they expect to receive (See Figure 1). The problem with this assessment is that the future is unpredictable, so expected volatility and returns doesn’t reflect actual performance. One proposed step is to use implied volatility in the over the counter derivatives market as a proxy for historical volatility. My research project attempts to asses if implied volatility can accurately predict a range of future outcomes.

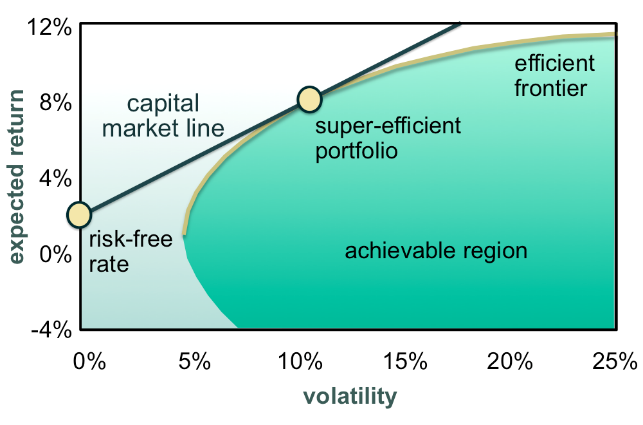


Figure 1- Visualization of the efficient frontier of stock return and volatility

The Black Scholes equation is one of the primary formulas used to derive the prices of put and call options. One of the explained variables that appears in the Black Sholes Model is implied volatility when the existing market price of an option is added into Black Scholes formula. To better understand what implied volatility is, let’s examine realized or historical volatility first. Historical Volatility is calculated from price changes in the underlying contract. It is the standard deviation of daily percentage change in the stock, often referred to as systemic risk. Implied volatility is a derived movement from the price of an option in the marketplace and represents a consensus of future realized volatility over the life of an option. Implied volatility is derived from the following relationships. There are five primary factors that help explain how these options are priced.

1) Underlying price of the stock or index

2) Strike price of the option.

3) Time until the Maturity of the Option

4) Interest Rate and Dividend Expecting to be paid

5) Implied Volatility-

It is important to note that implied volatility is calculated as a percentage of the underlying stock’s options prices. So there is a different implied volatility for the S&P 500, than Apple and a different one for Google, but they are all calculated exactly the same way. Implied volatility is designed to capture all of the expected movement, range and occurrences that might happen in the future of the stock or index. So events like earnings, product announcements, fear of expect takeovers are all priced into the implied volatility. Math equations derived from implied volatility indicate that we can use this concept to break out the expected movements of a stock. See figure 2- The price of the stock multiplied by the square root of the option time. This is later used in my project to derive implied volatility ranges at different time periods.

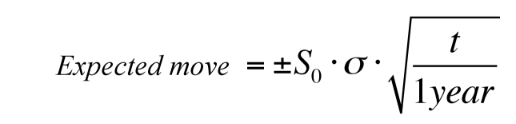


Figure 2- Calculation using Implied Standard Deviation to calculate an Estimated Move

**Dataset:**

The dataset used in this study contained 15 years of data in two datasets of historical price movement in the S&P 500, which is a stock market index, and the VIX Index, which measures the weighted 3 month implied volatility in options of the S&P500 and represents this data on an annual basis. The prices of these indices are supplied by the Chicago Mercantile Exchange and Chicago Board of Options Exchange and tracked on Yahoo Finance’s website. This data can be found at the following links: <http://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices> &

<http://finance.yahoo.com/q/hp?s=%5EVIX+Historical+Prices>.

I chose the S&P 500 and this associated measurement of implied volatility because the S&P 500 is the most liquid and frequently tracked index in the world and comprises of a diverse set of 500 actively traded US stocks. The VIX Index captures the implied volatility on the options in the S&P500, this is not an obscure index and is commonly quoted alongside the S&P 500. Its format represents the annual expected percentage move in the S&P500 based on the pricing of options in the derivatives marketplace. The dataset I used was from Yahoo Finance and contained 15 years of Historical data for the S&P 500 and the VIX Index.

|  |  |
| --- | --- |
| **Independent variables** | **Description** |
| Date | The date of trade and measurement in the index |
| Open Price | The starting price as of the open of the market |
| High Price | The highest price reached during the day |
| Low Price | The lowest price reached during the day |
| Close Price | The closing price as of 4 PM |
| Volume | Number of actively traded contracts in the index ( Not tracked for VIX) |
| Adj. Close Price | After the closing price of 4 PM, markets are halted from new trades, but existing trades are given 15 minutes to officially settle. This is the appropriate closing price to use because it captures all the data of one day. |

This data did contain some variables that required explanation. Specifically, volume represents the number of shares traded, this was only active for the S&P500 because the VIX is not traded as a stock. However, there are options on the VIX, there are VIX futures (both are traded on the CBOE) and there are exchange traded notes that attempt to represent the percentage change in VIX in shares of “stock” however due to advisory fees embedded into the price of these underlying stocks they are not reliable for this analysis. This dataset did not contain the specific variables needed for transformation, but did contain the components needed in order to calculate the required variables for analysis.

**Analysis:**

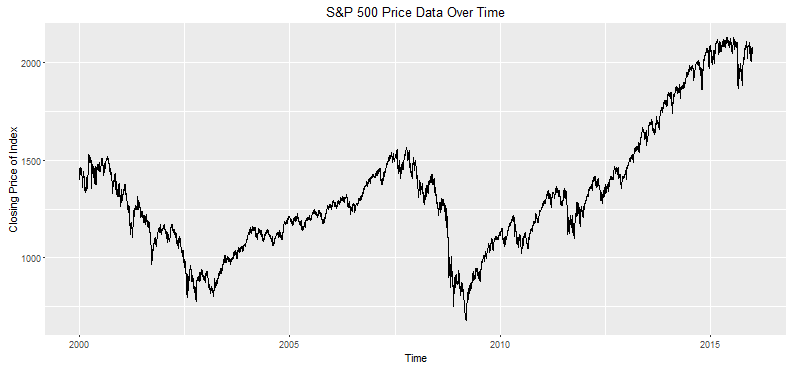
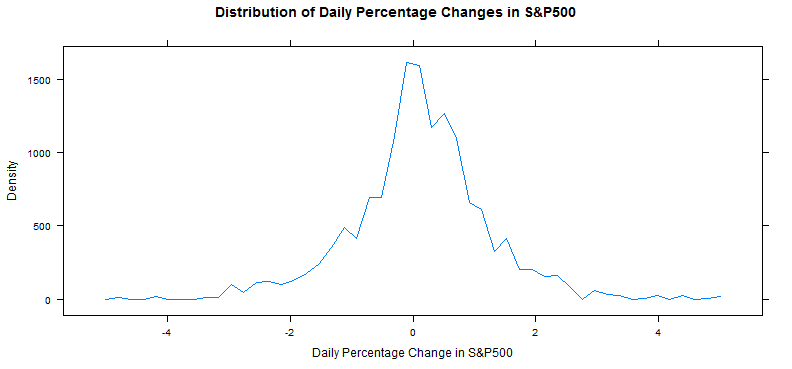
**Read-in and Examine Dataset:**

There were 4025 observations for each dataset. Although this data was relatively clean, I had to create a number of variables to calculate the measurements needed for analysis. Upon importing that data as a CSV file into R, I encountered two problems. First, the date classification was incorrect in both datasets was different and needed to be reformatted as a date variable. I merged these datasets together by date as a primary key, carefully indicating two different types of date formats. Second, I tested calculations on my time series data but noticed it was not being calculated appropriately, so I resorted the data, which resolved the problem. I checked for missing values and received in an indication that there were no specific missing values.

**Exploratory Data Analysis:**

My next step was to measure how many trading days there were in a year and month, to assess the reporting time periods of my analysis. My dataset had 255 observations a year and about 21 trading days in month. I used these estimates to calculate variables I use for later analysis.

The first step in data transformation of my new variables involved calculating the change in price of daily movement and adjusting for the total price of the index, or calculating the percentage change for both the S&P 500 and the VIX. I performed exploratory analysis of the change in the price of the variables and created some of the following visualizations. Figure 3 depicts the price of the S&P500 over time and Figure 4 shows the distribution of returns over this 15 year period. Using summarizing statistics of the daily percentage change, I identified there as a slight auto correlation due to some trendiness of my data in figure the. This prompted tests of autocorrelation in the stock market, see figure 5, indicating a low indication of data being auto correlated- which is in line with existing academic research.

Figure 3- Graph of the S&P500 from 1/1/00 – 12/31/15 Figure 4- Distribution of the % change in S&P500.

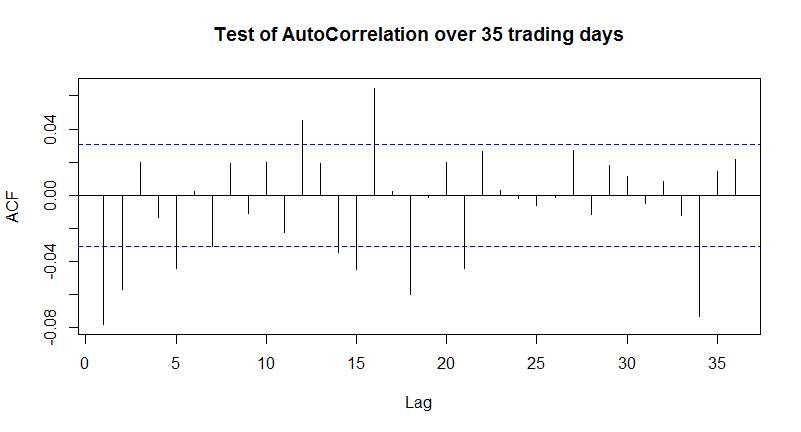


Figure 5- Lack of autocorrelation in the data- (Between -.08 and .4) indicate data is not likely correlated to previous period movement.

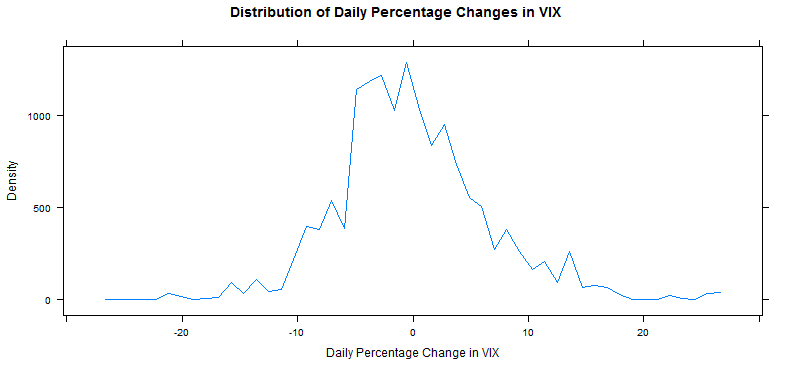
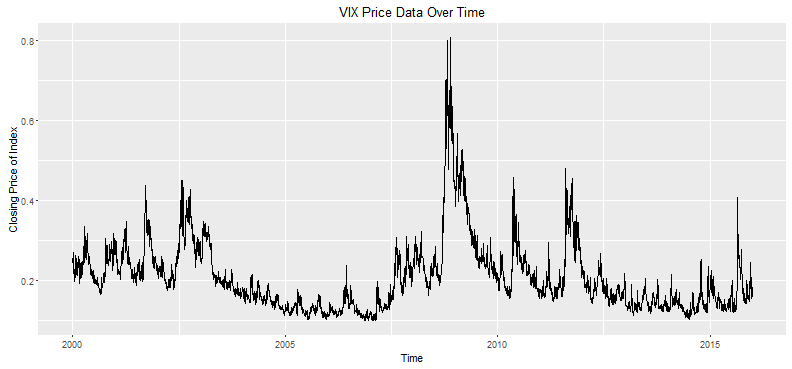


Figure 6- Graph of the VIX from 1/1/00 – 12/31/15 Figure 7- Distribution of the % change in the VIX

I performed the same analysis of the percentage change in the VIX as with the S&P 500. Again of my 4025 observations a normally distributed curve appeared. See figure 6 and 7 for the VIX Index over time and the distribution of daily returns in the same period. The large peaks in the trend of the data over time is due to extreme market sell-offs. The extreme values seen on the VIX are associated with sell-off in the 2008 financial crisis. What’s interesting about this is that implied volatility is a primary driver of the value of the option price, So during large sell offs or times of uncertainty, the option prices are expanding to account for a larger future trading range. It’s important to tie this concept down to prevent the misappropriation of cause and effect. We should approach this expecting that the underlying- the S&P500- changes and then drives a change of expected future price movement in the underlying option market and not the other way around. I ran a 3 month running correlation of the percentage daily change in the S&P 500 and VIX and plotted the relationship between these two variables against each other in Figure 8 and Figure 9.My descriptive model of the change in the VIX and the change in the S&P 500 indicated that price movement alone could explain 55% of the variance in the movement of the VIX.

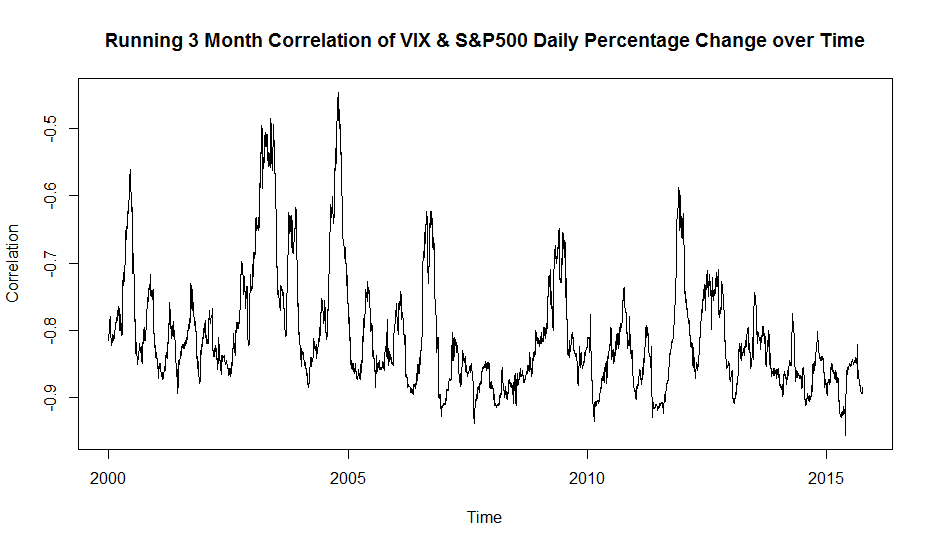
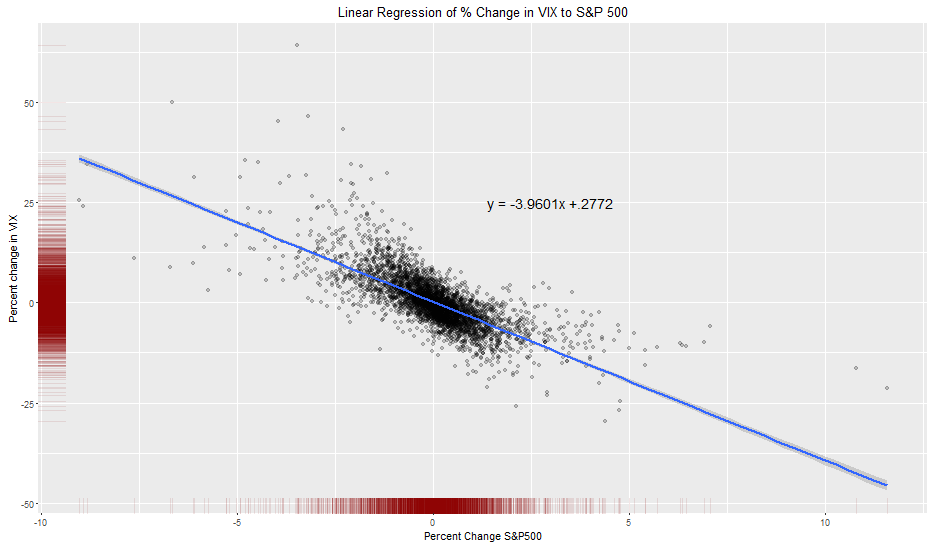


Figure 7- 3 Month Correlation of the % change in VIX Figure 8- Distribution of the % change in S&P500 and the S&P500 from 1/1/00 – 12/31/15 and VIX

Figures 7 and 8 act as plot points to contextualize the relationship of the stock market and the future range of motion predicted in the derivatives markets. This relationship is inversely correlated, and as depicted in figure 7 & 8, when the stock market increases the excepted future movement of the S&P 500 decreases and when the stock market decreases this expected movement increases.

**Analysis of Data Question:**

The next step in my exploration of the data was to calculate the actual implied percentage changes for each time period being tested and then calculate the actual- historical moves. I calculated there were 252 trading days in a year and on average 21 trading days in a month. I used this to extrapolate the time periods I wanted to examine, 2, 1 month, 2 month, 3 month, 6 month, 9 month and 1 year. The next step was to take the implied volatility – an annualized percentage figure and convert it to the respective time periods, so I took the implied volatility times the square root of my period times the number of trading days in a month divided by 252 trading days in a year for each of my desired outputs. Then I calculated the absolute percentage change in the indexes over the same period and lagged these results retroactively for comparability between implied and realized results. The reason why absolute realized returns were used is because implied volatility represents a positive or negative distribution range, and an absolute percentage can add greater clarity by just indicating if the return was inside or outside that expected range. As seen in Figure 9 and 10, the majority of realized returns were within the predicted distribution range for a 1 STD move. The black line indicates a line with the slope of one, so points below the line were under predicted by the implied volatility and points above exceeded a 1 STD prediction in implied volatility. These points show a correlation of .64. We would expect that 66% of the observations fall below the solid line and are thus within the 1STD predicted range of the Implied volatility. Instead we find, this number to be much higher, around 72-85% for each period.

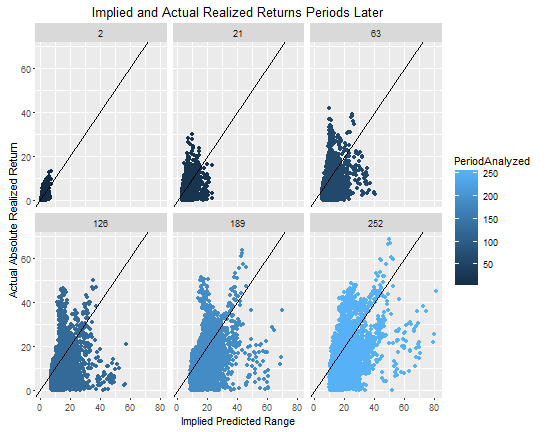
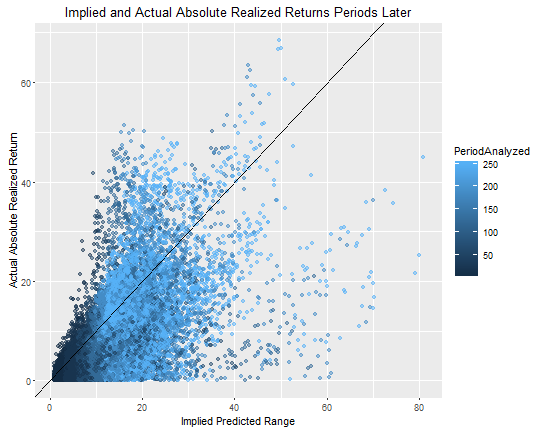


Figure 9: All future returns with past predicted ranges | Figure 10: All future returns and past predicted ranges

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Status** | **2 Days** | **1 Month** | **3 Month** | **6 Month** | **9 Month** | **12 Month** |
| Realized > Implied | 801 | 610 | 621 | 674 | 935 | 1061 |
| Realized < Implied | 3222 | 3394 | 3341 | 3225 | 2901 | 2712 |
| % Realized | 80.1% | 84.8% | 84.3% | 82.7% | 75.6% | 71.9% |

Figure 11- Expected 1 STD range would show 66% of occurrences within 1 STD range. Instead we see between 72% - 85%.

**Results and Discussion:**  
 This research shows a relationship does exist between realized and implied stock market returns and that implied volatility can be extrapolated to give context to market moves and that this extrapolated relationship is likely to be a high estimate because the implied volatility is overstating the risk in the prediction. How can this be used in the real world? Financial planners and financial media often talk about risk but never put a contextualized range regarding how much risk there is. Using implied volatility we can use this as a forward looking metric to measure risk in an index. Another way this research could be used, is in mutual funds. Most mutual funds are weighted to track the S&P 500 but do not have a liquid derivatives market to take implied volatility from. By beta weighting the correlated number of mutual fund shares held to the S&P 500 and multiplying by the Implied Volatility, one can put context around risk in mutual funds. Lastly, Efficient Market Hypothesis is an investment theory that states that it is impossible to beat the stock market because all information is priced into the stock at the current time. This research does not conflict with this theory, it could be that additional risk is priced into the implied volatility of the derivatives market in case for large outlier moves. To prove this, more research would be needed. One real world explanation this might help prove is why High Frequency Trader hedge their positions with options, it reduces directional risk in the stock and additional implied volatility might create an embedded edge in selling options against these positions.

**Appendix:**

Github Code: <https://github.com/TomHunter1/Springboard-Repository/blob/master/Final%20Project-%20Implied%20and%20Realized%20Volatility>

Presentation: <https://github.com/TomHunter1/Springboard-Repository/blob/master/Submission%20to%20Complete%20the%20Exploratory%20Portion%20of%20Capstone.pptx>