Analysis of Market Volatility

Ariel Khait

Fall 2024

Table of Contents

ANALYSIS OF VIX AND VOX INDEXES **2**
- WHAT IS IMPLIED VOLATILITY 2
- EVOLUTION OF VIX VS. VXO 2
- HISTORICAL DISTRIBUTION OF VIX VS. VXO 3
- ROLLING WINDOW COMPUTATIONS 5
- CHANGE IN METHODOLOGY 6
- COVID-19 PANDEMIC 7

PREDICTABILITY OF THE VIX **8**
- LITERATURE REVIEW OF EFFICIENT MARKET HYPOTHESIS 8
- ARMA SPECIFICATION ON LOGGED VALUES 8
- ARMA SPECIFICATION ON LOGGED DIFFERENCES 10
- APPLYING EFFICIENT MARKET HYPOTHESIS TO VIX 12
- RESIDUAL ANALYSIS 13
- PREDICTABILITY PUZZLE 14

REALIZED VS. IMPLIED VOLATILITY **17**
- REALIZED VOLATILITY 17
- SPX RANDOM WALK 17
- VOLATILITY SERIES ANALYSIS 18
- SPECTRAL DECOMPOSITION 19
- VECTOR AUTOREGRESSION MODEL 20
- CAUSALITY LINKS 20

CONCLUSION **21**

REFERENCES **22**

APPENDIX **23**

Analysis of VIX and VOX Indexes

What is Implied Volatility

Implied Volatility is the market's forecast of the underlying equity or index's future volatility, derived from the prices of options contracts that expire at a future date. Using the Black-Sholes model to price options, an options price depends on its strike price, time to maturity, underlying asset price, risk free rate, and volatility. This is considered arbitrage free pricing since it is a risk neutral way to price options. To calculate the value of a European option, all components of Black-Sholes model are known to the trader today with the expectation of volatility. With this in mind, we can use market prices of options in combination with the other components of the Black-Sholes model to determine an implied volatility.

The VIX is the 30-day implied volatility of the S&P500 index created by the Chicago Board Options Exchange (CBOE). Calculated using out-of-the-money, and in-the-money options with differing strike prices and maturities, the VIX is representative of the markets outlook on future volatility of the S&P 500. This is often known as the “fear” index since times of high volatility are known for their fear of market collapse meaning a high VIX value indicated to trader a large amount of “fear” in the market caused by some type of shock.

Its precursor, the VXO [also calculated by the CBOE] was calculated using at-the-money options on the S&P100. In 2003, CBOE switched over their calculation from strictly at-the-money options of the top 100 US companies [S&P100] to both out-of-the-money options and in-the-money options on the top 500 US companies [S&P500].

Evolution of VIX vs. VXO

Figure 1 is the VIX index from 2004 until the present day and Figure 2 is the VOX calculated from 2004 until CBOE stop calculating the index in September of 2021. We first observe that Figures 1 and 2 are very similar, in shape and structure as they are both calculated on options where the underlying assets are high market cap companies that usually behave similarly. It is important to note that there are large spikes in 2008, and 2020 corresponding to the housing bubble and the COVID-19 crisis respectively. These were times of large uncertainty in the economy and high volatility as seen by both indexes.

The only noticable differences we see between the two figures are the noticably consistent higher values of the VXO. This can be attributed to the less diversity in the index, looking at only a subset of options for the top 100 US companies as compared to a wide variety of options of the top 500 US companies.

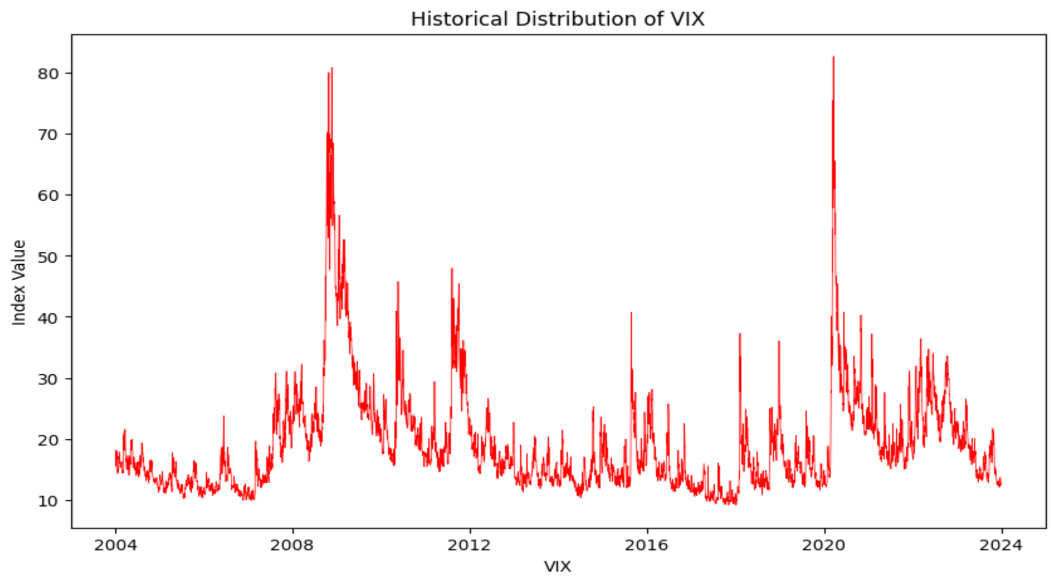


Figure 1(VIX)

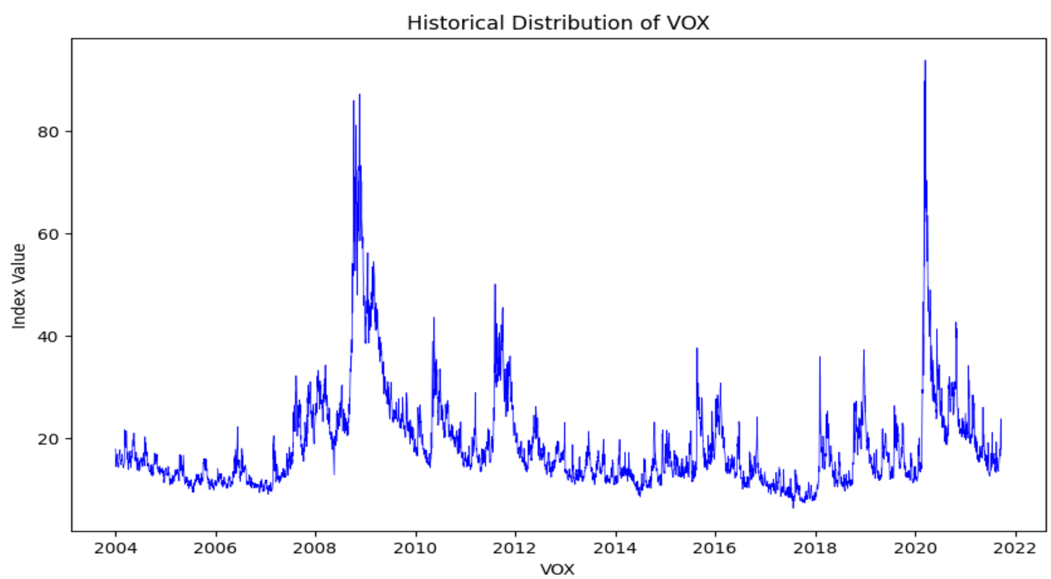


Figure 2 (VOX)

Historical Distribution of VIX vs. VXO

Firstly, when looking at the historical distributions of both the VIX and VXO, we must determine if they are stationary processes. Using an augmented Dickey-Fuller (ADF) test, we can determine if both processes have a unit root. The null hypothesis is that both processes have a unit root. Using the ADF test on both the VIX and VOX, we get an ADF of statistic of -5.422 and -4.684 with p values of 3.04e-06 and 9.01e-05 respectively. This means we can reject the null hypothesis and conclude both processes are stationary.

Calculating the moments of each distribution we get to Table 1:

	Mean	Variance	Skew	Kurtosis
VIX	*19.15*	*76.77*	*2.53*	*9.471*
VOX	*18.59*	*94.92*	*2.76*	*10.91*

Examining Table 1, we immediately notice the high skew and kurtosis in the data. Considering the standard normal values of skew and kurtosis are 1 and 3 respectively, we predict this distribution to be heavily skewed to the left with fat tails. Comparing this with the kernel estimate of our distribution in Figures 3,4 we see such a pattern appear. Furthermore, the average value of both indices around 18.75 with a higher variance as expected from the VOX due to its diversification compared to the VIX.

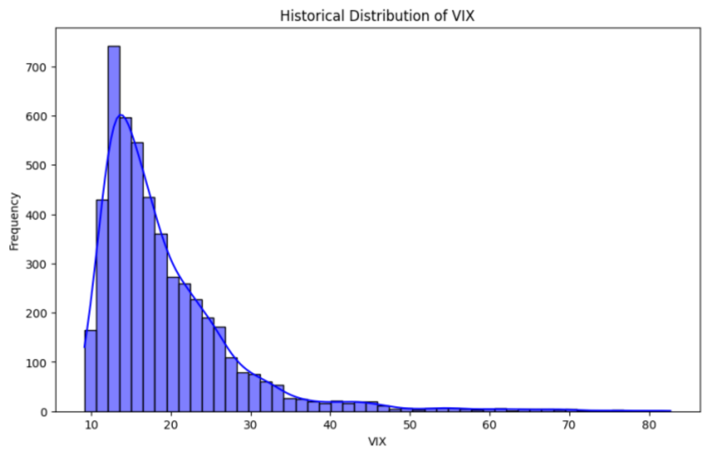


Figure 3: VIX Historical Distribution

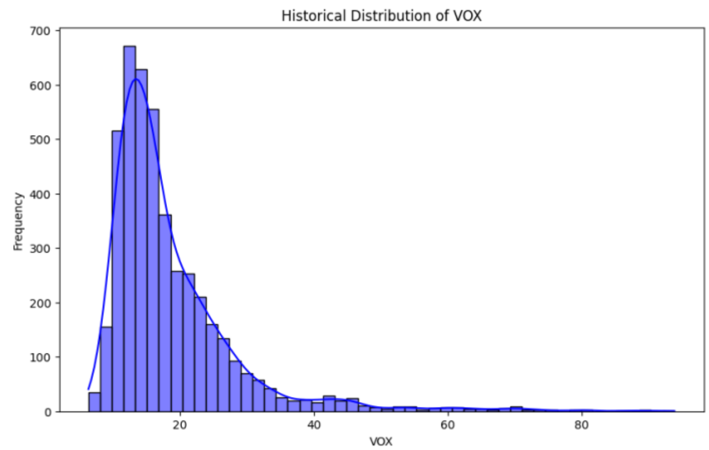


Figure 4: VOX Historical Distribution

Figures 5 through 8 represent the different Q-Q plots for both distributions, it's clear that the distributions are not Gaussian but instead exhibit characteristics of a lognormal distribution. Neither of the Q-Q plots aligns along the 45-degree line expected for a Gaussian distribution. Instead, there is a noticeable curvature in the middle, with significant deviations at both ends. This curvature indicates fat tails, meaning the distributions have high kurtosis and exhibit skewness as seen in Table 1.

Looking at Figures 5 through 8, we see that both the VIX and VOX data better fit the 45-degree line in Figures 6 and 8 as compared to Figures 5 and 7. This further suggest that both time series follow a lognormal distribution with one fat tail towards the top right of the plot rather than fat tails on both sides of the 45-degree line.

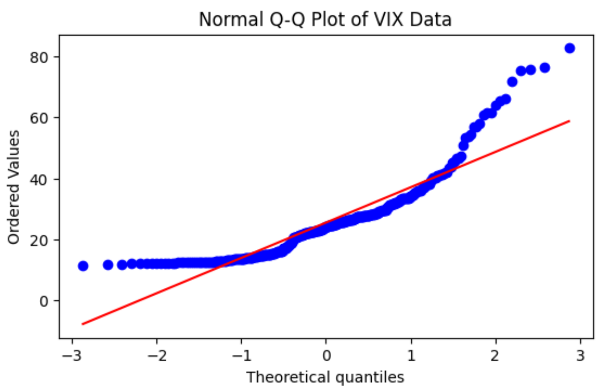


Figure 5: Normal Q-Q Plot of VIX

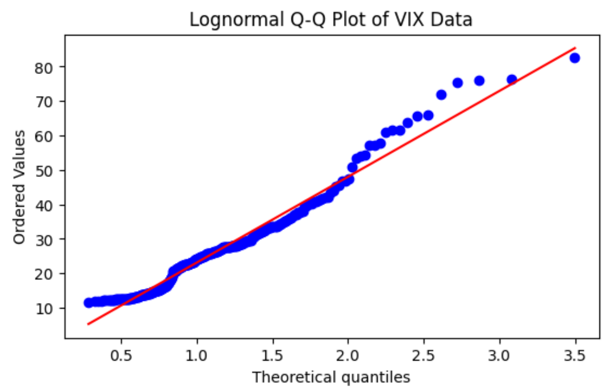


Figure 6: Lognormal Q-Q Plot of VIX

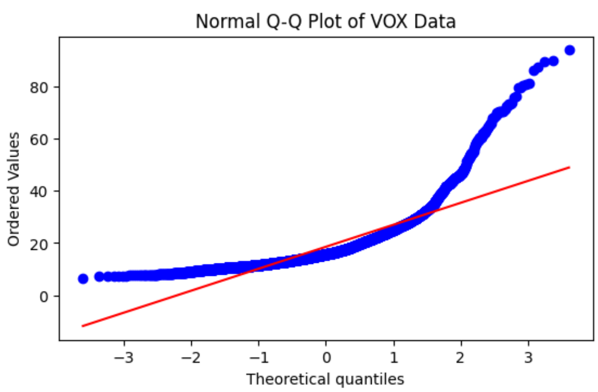


Figure 7: Normal Q-Q Plot of VOX

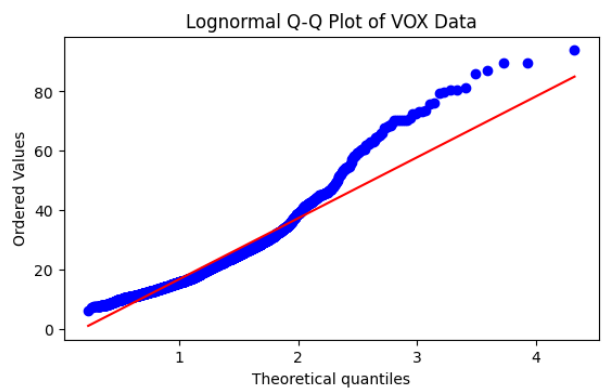


Figure 8: Lognormal Q-Q Plot of VOX

Rolling Window Computations

The market for the volatility of volatility consists of volatility derivatives, such as VIX futures and VIX options. These derivatives do not have a stock or index as their underlying asset; instead, they are based on the implied volatility of the market, specifically the volatility of the S&P 500 index as measured by the VIX. This allows traders to speculate on or hedge against changes in market sentiment rather than price movements of a specific index or stock. Investors use these products to manage risk in periods of high uncertainty, such as during economic downturns or significant market events, where volatility tends to spike. We see this in Figures 9 through 12 where during the 2010 depression, the rolling standard deviation had a very similar shape to the rolling median and VaRs. This allows traders to trade options on the VIX and VOX to hedge themselves against market downturns such as the 2008 housing market bubble. Furthermore, we see more recently the same activity where investors could use the variance of the VIX and VOX indexes to hedge against the COVID-19 crisis.

In general, we see that the rolling variance spikes earlier than the rolling mean and VaR’s. This implies that the rolling variance of the VIX can be used as an early indicator of a market collapse like we saw in 2008 and like we are now seeing after COVID-19.

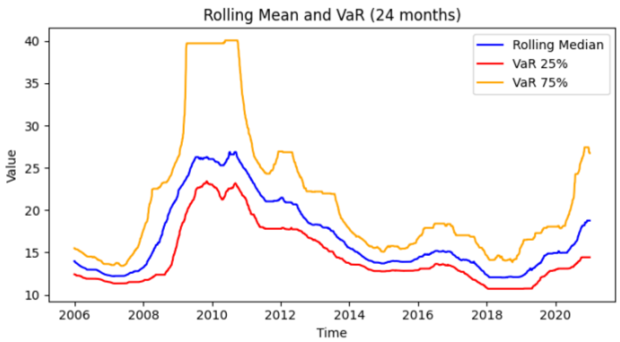


Figure 9: VIX Rolling Mean and VaR

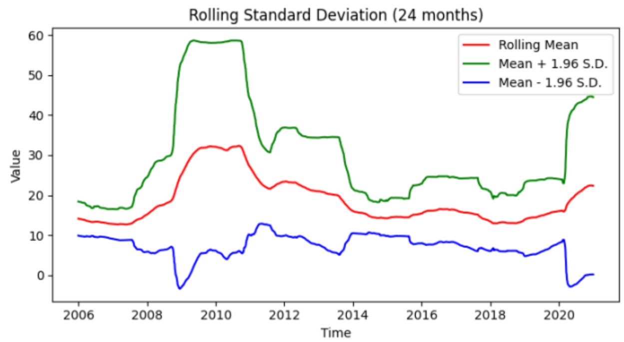


Figure 10: VIX Rolling Variance

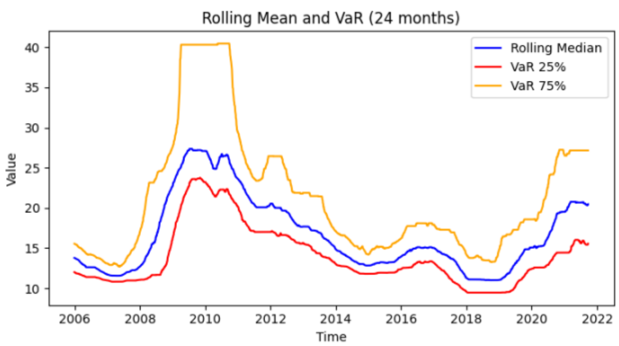


Figure 11: VOX Rolling Mean and VaR

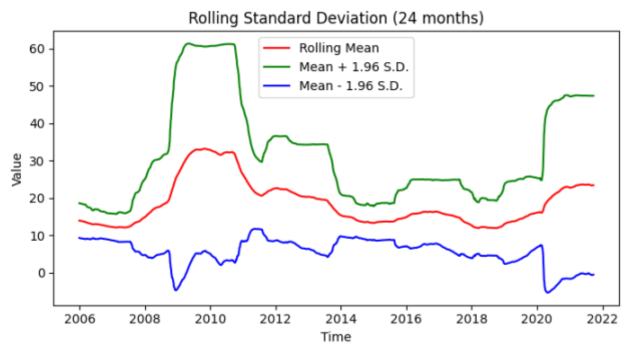


Figure 12: VOX Rolling Variance

Change In Methodology

Since the change in methodology, the VIX has become more comprehensive, covering options for 500 of the largest-cap companies in the S&P 500, rather than focusing solely on at-the-money options for the top 100 companies (S&P 100). The VXO is still calculated because it provides traders with insight into the sentiment difference between the "too big to fail" companies in the S&P 100 and the broader market represented by the other large-cap companies in the S&P 500.

This diversification is useful during market downturns. In times of crisis, the Federal Reserve may view the bankruptcy of any S&P 100 company as detrimental to the economy and consequently, it is more inclined to act if the situation arises. This safety net results in less implied volatility for these companies. However, the 400 companies in the S&P 500 but not in the S&P 100 may not receive the same federal support. As a result, the fear of collapse could be higher for these companies, leading to a divergence between the values of the VIX and the VXO. This divergence helps traders gauge the severity of implied volatility between the top 100 companies and the remaining 400.

COVID-19 Pandemic

Focusing on the COVID-19 pandemic, we notice that the rolling mean and VaR's begin to increase towards the beginning of the pandemic but more importantly, the rolling variance of the VIX index blows up from a value of around 10 to 125. This is very similar to the 2008 crisis when the rolling variance shot up while the rolling mean and VaR's gradually increased. If we use the variance of the VIX as an indicator for panic in the market leading to an eventual depression, we can then quantify the COVID-19 pandemic one such potential event.

Predictability of the VIX

Literature Review of Efficient Market Hypothesis

According to Clark (1973), if we let X_t denote the price of an equity at time t, examining the data shows that X_t exhibits a random walk, which is characterized by the equation $X_t = X_{t-1} + \varepsilon_t$ where ε_t is independently distributed random variable of price adjustments. The efficient market hypothesis then claims:

$$\mathbb{E}(X_{t+1} | X_t) = X_t + \mathbb{E}(\varepsilon_{t+1}) = \mathbb{E}(X_{t+1} | X_t, X_{t-1}, X_{t-2}, \dots)$$

When predicting X_{t+1} using the current price of an asset (X_t), it is as effective as using the entire history of prices for that asset, a property known as the Markov process. Furthermore, by examining past data, Clark postulates that changes in asset prices, although independent, are not normally distributed but instead leptokurtic. He attributes this to independent but non-finite variances in ε_{t+1}, which form a leptokurtic distribution. This distribution is then transferred to changes in asset prices through independent information shocks.

ARMA Specification on Logged Values

We begin by analyzing the series $y_t = \ln(VIX_t)$, representing the log returns of the VIX index, over the period from 2004 to the present, as shown in Figure 13. Next, we examine the Autocorrelation Function (ACF) for up to 300 lags, as illustrated in Figure 14, trying to estimate an autoregressive model for y_t.

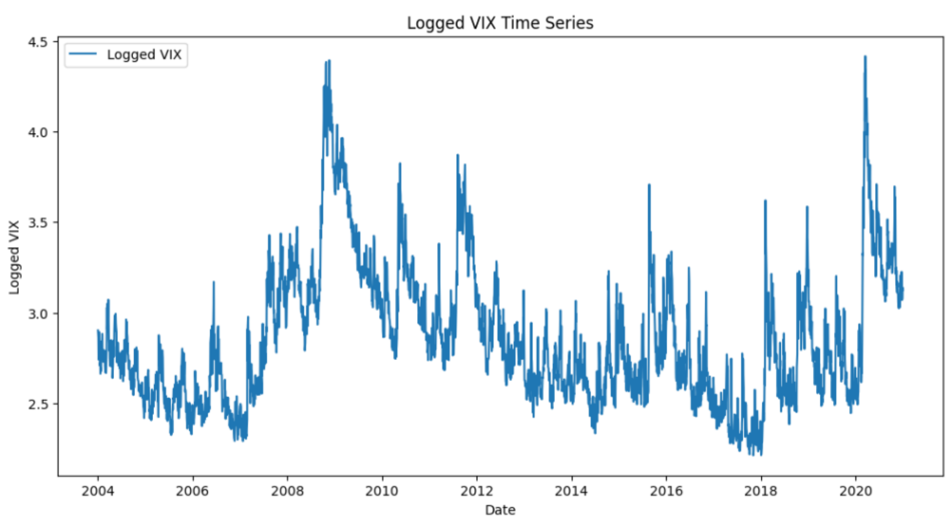


Figure 13 (Logged VIX Returns)

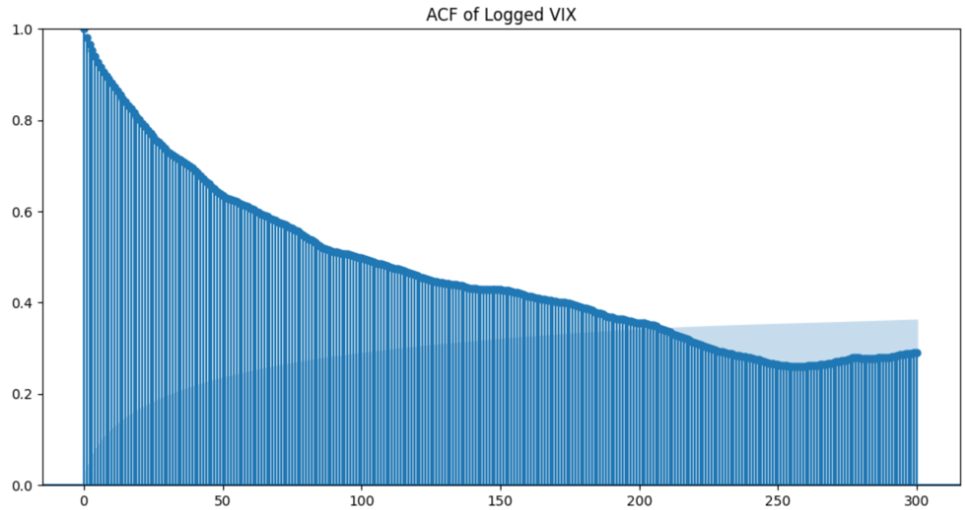


Figure 14 (ACF of Logged VIX Returns)

Upon examining the autocorrelation function (ACF) of the logged VIX returns shown in Figure 14, we observe that autocorrelation remains present up to approximately the 210th lag, at which point it becomes statistically insignificant (indicated by the light blue shaded confidence interval). Initially, the autocorrelation is close to 1.0, gradually decaying to around 0.2 over the course of approximately 260 lags, which corresponds to a full trading year.

To fit an autoregressive model to this data, we employed a grid search algorithm to identify the model with the lowest Akaike Information Criterion (AIC). This analysis points to the ARMA(2,1) model as the optimal choice for modeling y_t, with the corresponding coefficients shown in Table 2 below:

	Coefficient	S.E.	P Value
ϕ_1	*1.7048*	*(0.044)*	*0.000*
ϕ_2	*-0.7074*	*(0.043)*	*0.000*
θ_1	*-0.8018*	*(0.038)*	*0.000*

Revisiting the ADF test from Section 1.2 of our analysis, we rejected the possibility of a unit root in the VIX data. When applying the ADF test to the logged data, we obtained an ADF statistic of -4.52 with a p-value of 0.00018, allowing us to reject the null hypothesis that the logged data has a unit root (i.e., the logged data is stationary). This result contrasts with our ARMA(2,1) model, where the sum of the coefficients $\phi_1 + \phi_2 = 0.9974$, is very close to 1, suggesting non-stationarity.

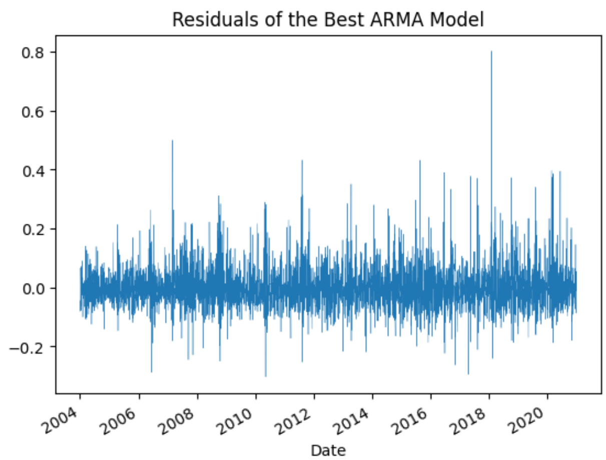


Figure 15 (Residuals of ARMA(2,1) Model)

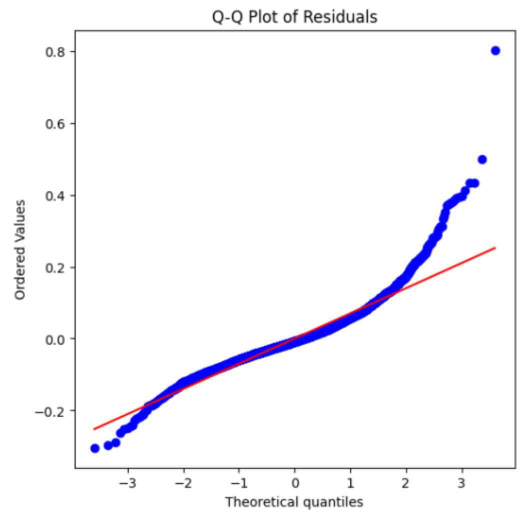


Figure 16 (Q-Q Plot of Residuals)

Focusing on the residuals of the model (Figure 15), we observe that they appear to resemble a white noise process, as one would expect from an ARMA model. However, a closer examination of the Q-Q plot in Figure 16 reveals that the residuals do not follow a normal distribution. Additionally, the results of the Ljung-Box Test yield a p-statistic of 0.31 implying series correlation amongst the residuals and the Jarque-Bera test yields a significant statistic of11592.8996, allow us to reject the null hypothesis of normality. These findings suggest that the model may be mis-specified.

ARMA Specification on Logged Differences

We now shift the focus of our analysis to the change in the logged VIX data: $\Delta y_t = y_t - y_{t-1} = \ln(VIX_t) - \ln(VIX_{t-1})$. After applying this transformation to the data, we observe that it appears to be stationary, as shown in Figure 17. The autocorrelation function (ACF) of this differenced data differs significantly from that of the original logged data. In Figure 18, we see that only the first two lag exhibits significant autocorrelation, suggesting an AR(2) process contrary to y_t model where the ACF was significant until the 210th lag.

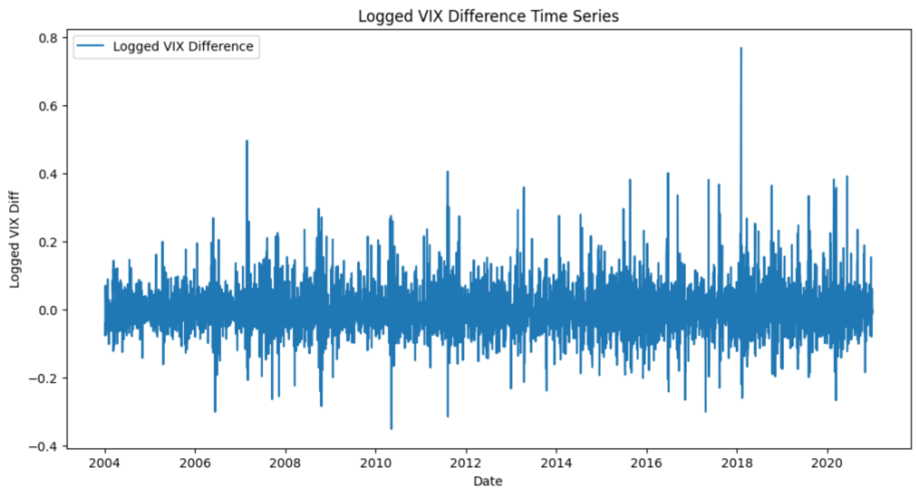


Figure 17 (Logged VIX Difference Over Time)

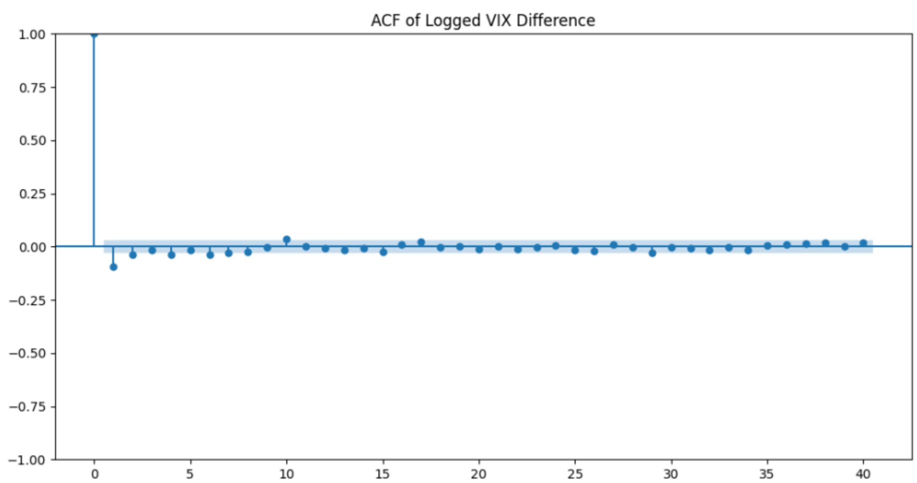
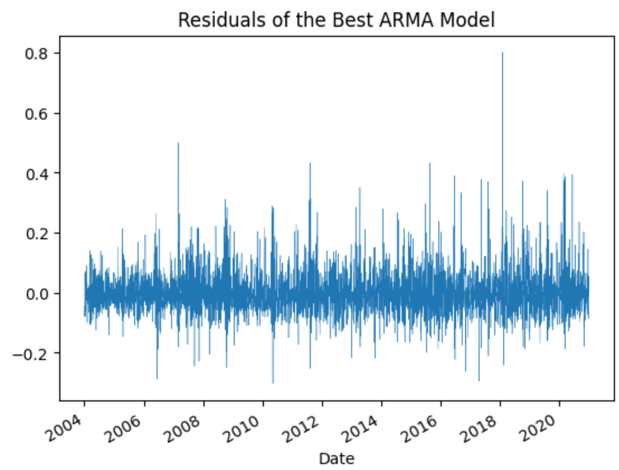


Figure 18 (ACF of Logged VIX Differenced Data)

Again, using the AIC to find the best fitting model, we find that the ARMA(3,1) with parameters seen in Table 3 below, best fit the data.

	Coefficient	S.E.	P-Value
ϕ_1	*0.8768*	*(0.011)*	*0.000*
ϕ_2	*0.0454*	*(0.014)*	*0.002*
ϕ_3	*0.0293*	*(0.011)*	*0.011*
θ_1	*-0.9861*	*(0.005)*	*0.000*

We find that the ADF statistic is a significant -23.4, once again rejecting the null hypothesis that the data is non-stationary. From Table 3, the sum of the coefficients is still slightly lower than 1, but higher than in the y_t model. Continuing with our comparison, we examine the residuals and the resulting Q-Q plot, as shown in Figures 19 and 20.



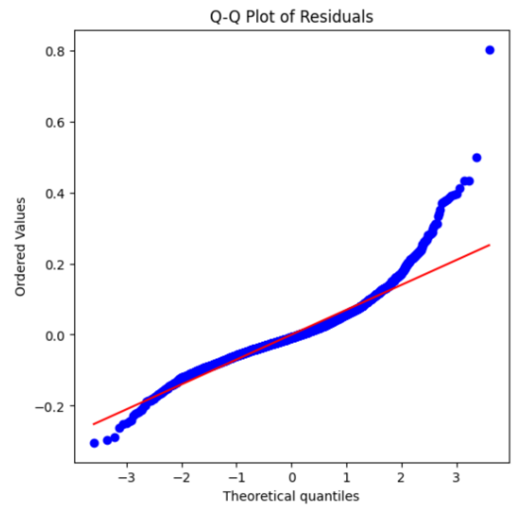


Figure 19 (Residuals of ARMA(3,1) Model) Figure 20 (Q-Q Plot of Residuals)

Similarly to the y_t model, the Δy_t model appears to have residuals that are random, as shown in Figure 19. However, this assumption is disproven upon closer inspection. The Q-Q plot reveals that the residuals are not normally distributed, and the Ljung-Box and Jarque-Bera tests further confirm that the residuals are both serially correlated and not normally distributed. Overall, this ARMA model for Δy_t can be seen as an integrated version of the y_t model. However, due to the serial correlation present in the residuals, the ARMA model does not adequately fit our data, so it seems that the VIX does not exhibit a linear structure, making the ARMA model unsuitable for modeling the VIX.

Applying Efficient Market Hypothesis to VIX

The Efficient Market Hypothesis (EMH) states that prices follow a random walk with no long-term “steady-state” or average level. Applying the EMH to our analysis of the VIX implies that its returns should exhibit a random walk, represented by the following equation:

$$y_t = \alpha + y_{t-1} + \varepsilon$$

where y_t is the log returns of the VIX. Furthermore, notice that by subtracting y_{t-1} from both sides and substituting for Δy_t we get:

$$\Delta y_t = \alpha + \varepsilon.$$

This intuitively means the difference in log returns resembles a random walk ε up to a drift of α. Turning our attention back to the VIX data, we conduct a regression to test for the presence of the Efficient Market Hypothesis (EMH):

$$\Delta y_t = \alpha + \beta_1 \Delta y_{t-1} + \beta_2\, \Delta y_{t-1}^2 + \varepsilon$$

Table 4	α	β_1	β_2	R^2
Reg 1: Δy_t	***6.641e-05**** *(0.001)*	***-0.950**** *(0.015)*	-	***0.009***
Reg 2: Δy_t	***0.0020*** *(0.001)*	-	***-0.3647**** *(0.071)*	***0.006***
Reg 3: Δy_t	***0.0014*** *(0.001)*	***-0.0769**** *(0.016)*	***-0.2369**** *(0.076)*	***0.011***

****: Statistical Significance at a 5% level*
***: Statistical Significance at a 10% Level*

Upon examining the regression results, we initially expected that both $\beta_{1,2}$ would equal zero to confirm the Efficient Market Hypothesis (EMH) in our data. However, contrary to this expectation, both $\beta_{1,2}$ are statistically significant at the 5% level across all three regressions. Focusing on Regression 3, we observe a significant, non-zero correlation between both the lagged differenced log returns (Δy_t), its previous iteration (Δy_{t-1}) and the squared lagged differenced log returns (Δy_{t-1}^2), directly contradicting the EMH as we defined it.

Furthermore, the significant negative correlation between the current change in log returns (Δy_t) and its lagged value (Δy_{t-1}) suggests that the VIX can serve as a hedge against short-term market volatility, as discussed in Chapter 1. Additionally, the strong negative correlation between Δy_{t-1}^2 and Δy_t, in Regression 3 indicates that the VIX exhibits short-term predictability based on previous levels of volatility. This result further refutes the EMH in the short term, implying that VIX movements can be anticipated to some extent using past volatility patterns.

Residual Analysis

Turning our attention to the residuals of our regression, we plot the historical distribution (Figure 20) and associated Q-Q plot (Figure 21) to determine normality.

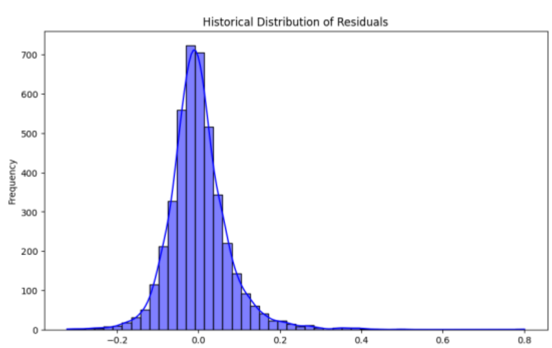


Figure 20 (Historical Distribution of Residuals)

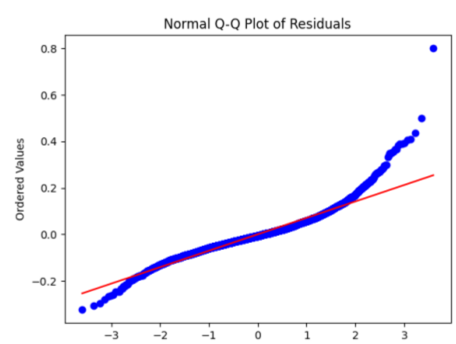


Figure 21 (Normal Q-Q Plot of Residuals)

Based on the Efficient Market Hypothesis (EMH), we assume that our model is influenced by random white noise with a constant drift. In Figure 20, the residuals visually appear to follow a Gaussian distribution, as expected. However, a closer examination in Figure 21 reveals that the distribution actually exhibits fat tails, or high kurtosis. This observation is confirmed by the moments presented in Table 5:

	Mean	Variance	Skew	Kurtosis
Residuals	-1.34e-18	0.005	1.177	7.198

Table 5 reveals a mean that is close to zero, aligning with what we would expect in a Gaussian distribution. However, the variance is smaller than the Gaussian benchmark of 1, indicating that the data is more tightly clustered around the mean than anticipated. Examining the third and fourth moments, we observe a slight positive skew and a kurtosis significantly higher than what would be expected in a normal distribution. These findings raise questions about potential model misspecifications, such as the assumption that each residual is independently drawn. To investigate further, we turn to the ACF of the residuals and residuals squared (Figure 22) with this potential issue in mind.

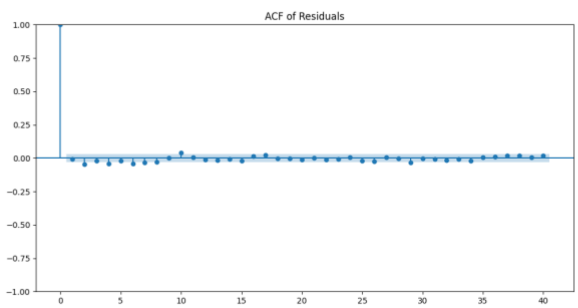


Figure 22 (ACF of Residuals)

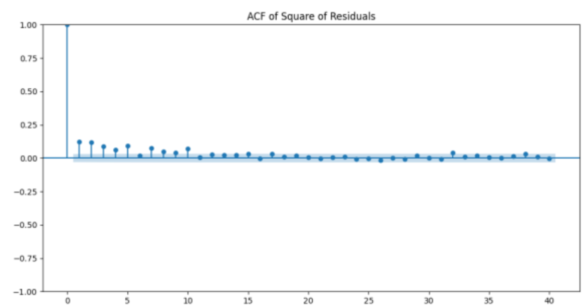


Figure 23 (ACF of Residuals Squared)

In Figure 22, we observe a significant autocorrelation at lag 1, while the remaining lags fail to reject the null hypothesis of zero autocorrelation. This figure indicates a short-term autocorrelation in the residuals, leading to serial correlation between consecutive residual draws. Figure 23 further shows a significant autocorrelation for approximately the first 10 lags, suggesting short-term correlation in the volatility of the residuals. Overall, our regression residuals do not follow a Gaussian distribution, and the presence of serial correlation contradicts the Efficient Market Hypothesis (EMH) in the short term.

Predictability Puzzle

Extending our analysis in the previous section, we look to at the predictability of the VIX over h periods. We firstly define $\Delta y_{t+1,t+h} = \Delta y_{t+1} + \Delta y_{t+2} + \cdots + \Delta y_{t+h} = \sum_{i=1}^{h} \Delta y_{t+i}$ as the sum of the logged returns h periods into the future with Δy_t defined as it has been in previous chapters. Secondly, we define $\Delta y_{t-k+1,t} = \Delta y_t + \Delta y_{t-1} + \cdots + \Delta y_{t-k+1} =$

$\sum_{i=1}^{k} \Delta y_{t-i+1}$ as the sum of logged returns h periods into the past, where once again, Δy_t is defined similarly to previous chapters. We then define the regression:

$$\Delta y_{t+1,t+h} = \beta_h \Delta y_{t-k+1,t} + w_{t,h}$$

Let *h* represent the time horizon at which we aim to predict the VIX, using the sum of logged returns from *h* periods ago along a white noise component. We focus specifically on β_h, which reflects the relationship between past returns over *h* periods and future returns over the same period. To extend our analysis, we examine the R^2 of our regression, which represents the variance in the sum of future logged returns explained by the sum of past logged returns over *h* periods. In this context, we interpret R^2 as a measure of predictability.

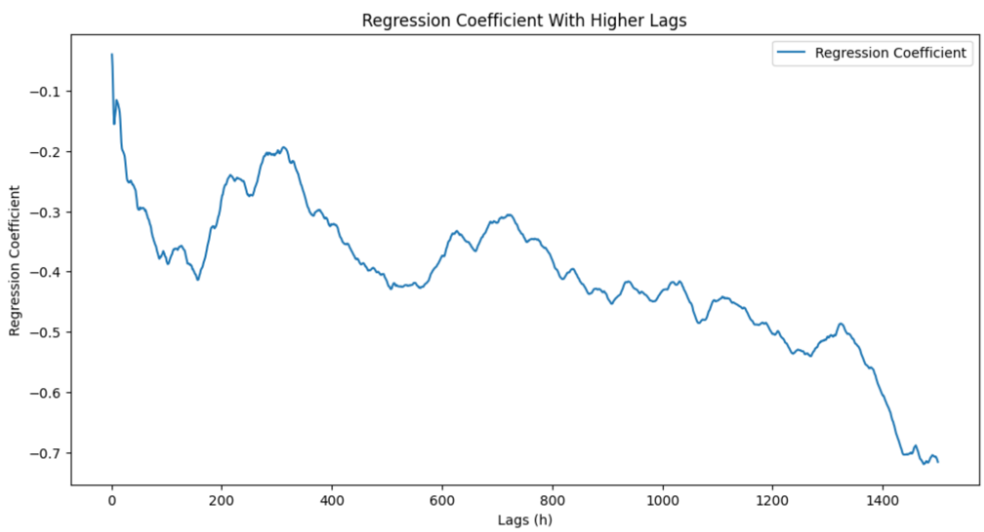


Figure 24 (Regression Coefficient)

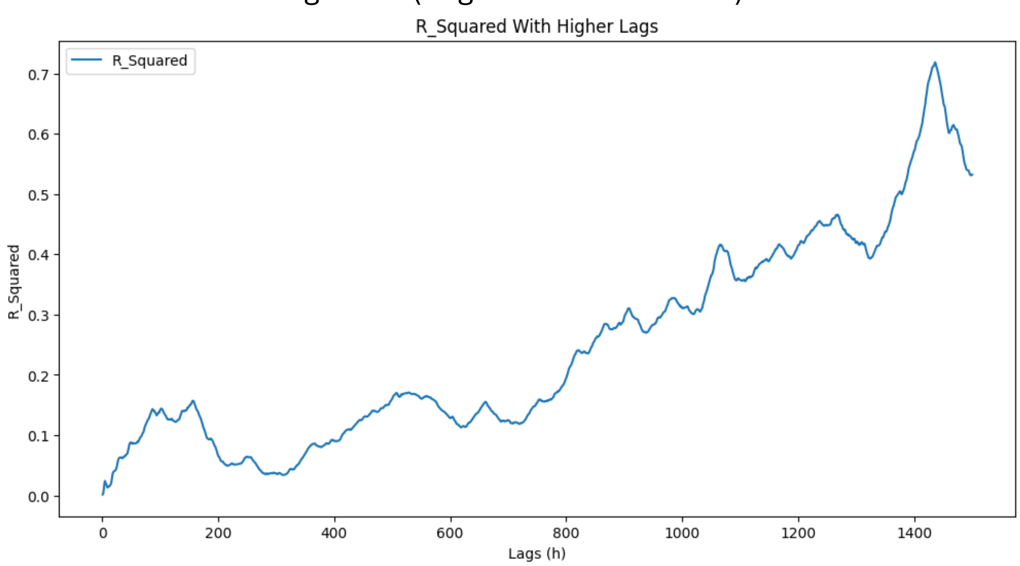


Figure 25 (R Squared of Regression)

Figures 24 and 25 reveal a trend that contradicts the Efficient Market Hypothesis (EMH), showing that predictability increases up to around the 1400th lag, with the regression coefficient also trending downward. This implies that the VIX becomes more predictable as the forecast horizon extends, although it is important to note that the number of observations decreases significantly with the number of lags. This finding mirrors the conclusion of Fama and French (1988), who observed that 'autocorrelations become negative for 2-year returns and reach their lowest points for 3-5-year returns' when predicting stock returns.

Realized vs. Implied Volatility

Realized Volatility

We compute two measures of realized volatility. Let the return (in percentage terms) of the Standards & Poor's Index (SPX) be computed as such:

$$y_t = \frac{SPX_t - SPX_{t-1}}{SPX_{t-1}} \times 100$$

Furthermore, we compute the realized short-term volatility (v_1) as:

$$v_1 = y_t^2$$

We then compute a 25-day moving average of realized volatility (v_2) as such:

$$v_2 = \frac{1}{25}\sum_{i=1}^{25} y_t^2$$

SPX Random Walk

We start by testing the stationarity of the y_t time series using the Augmented Dickey-Fuller (ADF) test. The result yielded an ADF statistic of -5.42, which is lower than the 1% critical value of -3.43. This indicates that the series is stationary, with a p-value close to zero.

Next, we examine Figures 26 and 27, which display the returns and autocorrelation function (ACF) of our SPX time series. The daily mean return is 0.0385%, which translates to an annualized return of 13.63% since 1990. Additionally, the ACF shows a sharp drop after the first lag, suggesting that the series behaves like a Markov chain, where all relevant information is captured by the first lag.

Given the stationarity of the SPX returns and the significant autocorrelation at the first lag, I believe the Efficient Market Hypothesis (EMH) holds for the S&P500 returns.

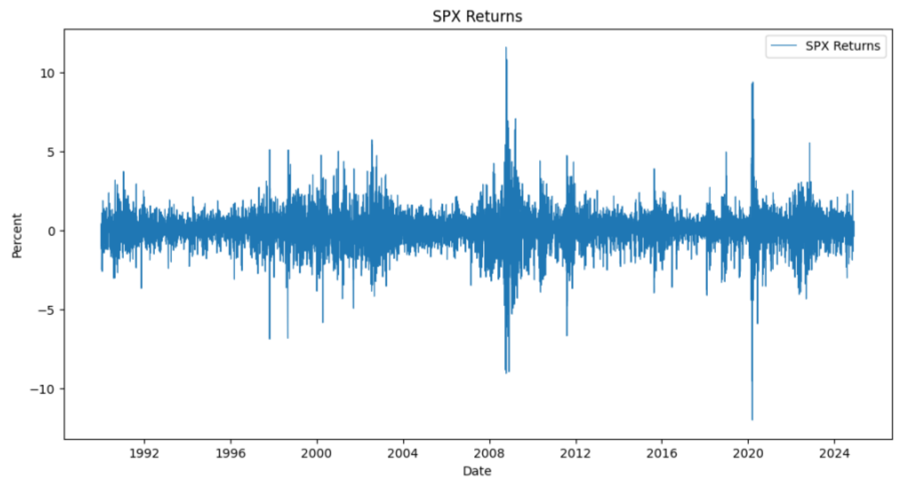


Figure 26 (SPX Returns)

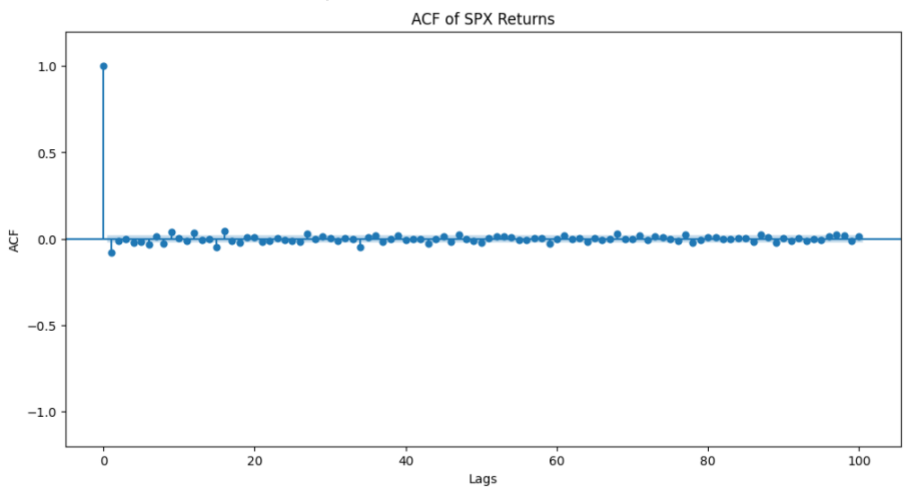


Figure 27 (ACF of SPX Returns)

Volatility Series Analysis

Looking at the previously constructed series, we calculate their first four moments. We note that because we conducted our analysis in percentage terms, we must convert back to conduct a spectral decomposition. The first four moments are as follows:

Moments	Mean	Variance	Skew	Kurtosis
VIX	19.15	76.79	2.54	9.48
v_1	0. 000145	3.13×10^{-7}	13.50	249.88
v_2	0.000145	9.84×10^{-8}	6.26	44.57

We further calculate their first and second cross moments:

$$\Sigma = \begin{pmatrix} \sigma_{VIX,VIX} & \sigma_{v_1,VIX} & \sigma_{v_2,VIX} \\ \sigma_{VIX,v_1} & \sigma_{v_1,v_1} & \sigma_{v_2,v_1} \\ \sigma_{VIX,v_2} & \sigma_{v_1,v_2} & \sigma_{v_2 v_2} \end{pmatrix} = \begin{pmatrix} 76.79 & 0.002519 & 0.002165 \\ 0.002519 & 3.13 \times 10^{-7} & 8.66 \times 10^{-8} \\ 0.002165 & 8.66 \times 10^{-8} & 9.84 \times 10^{-8} \end{pmatrix}$$

Examining Figure 28, we observe that all the time series exhibit spikes at similar periods. Notably, this pattern emerges during the 2008 market crash, the 2010 European Sovereign Debt Crisis and its lingering effects into 2012, and the COVID-19 pandemic in 2020. Additionally, we find that the 25-day rolling average presents a 'smoother' version of the realized volatility, effectively dampening the extreme spikes observed in the v_1 series. This makes it more comparable to the VIX index, as it averages out short-term fluctuations.

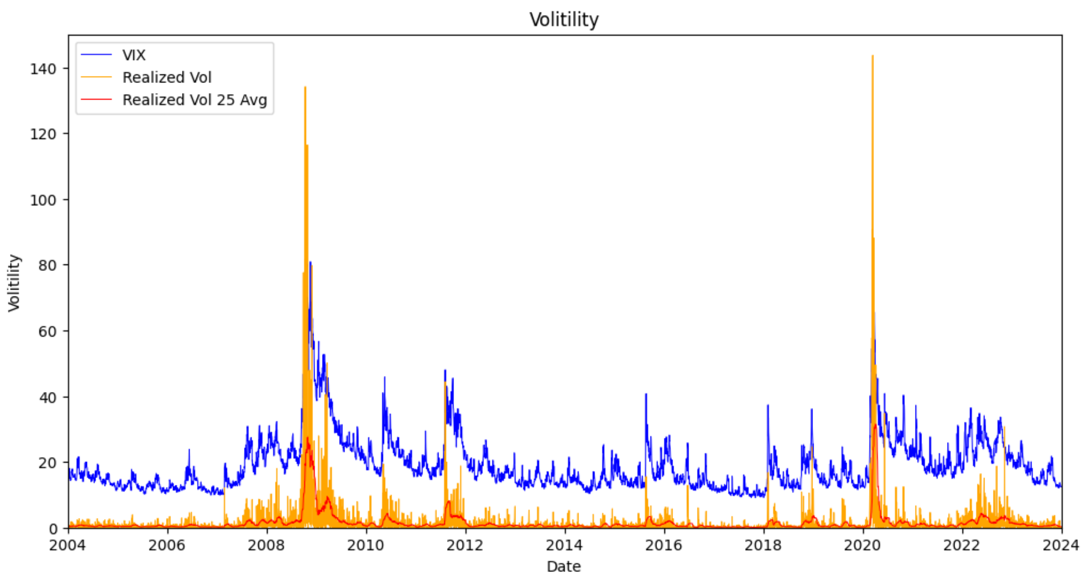


Figure 28 (Volatility Time Series)

Spectral Decomposition

Solving for the eigenvalues by way of spectral decomposition, we define Q and Λ as $\Sigma = Q\Lambda Q^{-1}$ where Q and Λ are both 3 by 3 matrices of eigenvectors and eigenvalues respectively. Doing the decomposition yields us:

$$Q = \begin{pmatrix} -0.999 & -3.5 \times 10^{-5} & -2.55 \times 10^{-5} \\ -3.28 \times 10^{-5} & 0.97 & -8.02 \times 10^{-2} \\ -2.82 \times 10^{-5} & -8.02 \times 10^{-2} & 0.97 \end{pmatrix} \Lambda = \begin{pmatrix} 76.79 & 0 & 0 \\ 0 & 2.32 \times 10^{-7} & 0 \\ 0 & 0 & 3.6 \times 10^{-8} \end{pmatrix}$$

We notice that the lowest eigenvalue is 3.6×10^{-8} which is close but not equal to 0. This means we do not have an arbitrage opportunity such that buying and selling our portfolio would yield to a risk-free return higher than that of federal bonds.

Vector Autoregression Model

We estimate a VAR(1) model with the following specifications:

$$Y_t = \begin{bmatrix} VIX_t \\ v_{1,t} \\ v_{2,t} \end{bmatrix} \text{ and } Y_t = c + AY_{t-1} + \varepsilon_t$$

Which results as follows:

$$\hat{c} = \begin{pmatrix} 0.63 \\ -0.000311 \\ -0.000013 \end{pmatrix} \quad \hat{A} = \begin{pmatrix} 0.96 & -24.75 & 547.86 \\ 2.15 \times 10^{-5} & 0.1 & 0.21 \\ 8.31 \times 10^{-7} & 0.01 & 0.97 \end{pmatrix}$$

Doing the spectral decomposition on the residuals of the VAR(1) model then yields:

$$\hat{Q} = \begin{pmatrix} -0.99 & -6.22 \times 10^{-5} & 2.83 \times 10^{-7} \\ -6.21 \times 10^{-5} & 0.99 & -4.34 \times 10^{-2} \\ -2.42 \times 10^{-6} & -4.34 \times 10^{-2} & 0.99 \end{pmatrix} \widehat{\Lambda} = \begin{pmatrix} 3.40 & 0 & 0 \\ 0 & 2.23 \times 10^{-7} & 0 \\ 0 & 0 & 3.64 \times 10^{-10} \end{pmatrix}$$

We observe an eigenvalue close to zero, specifically 3.64×10^{-10}. However, since this is a spectral decomposition of the residuals, it does not allow us to construct an arbitrage portfolio. To construct an arbitrage portfolio, we would need a combination of assets that can produce a risk-free asset (eigenvalue of exactly 0) with a return different from the prevailing risk-free rate. In such a case, an arbitrage opportunity arises by creating a costless risk-free position through shorting one component (either the risk-free asset or the portfolio) while longing the other.

An eigenvalue of exactly 0 in this model indicates that the residuals associated with the eigenvector corresponding to this eigenvalue exhibit no variation rather than a combination of assets yielding a risk-free return. This eliminates the stochastic component, allowing future values to be predicted with certainty, assuming prices remain unchanged. However, in our case, the eigenvalue, while close to zero, is not exactly zero. This suggests there remains a stochastic element to future volatility, and no perfectly deterministic trend can be identified or exploited. Therefore, there is no arbitrage opportunity in this context.

Causality Links

Finally, we examine the causal relationships between implied volatility and realized volatility. Through Granger Causality tests, we observe that the VIX time series Grangercauses both v_1 and v_2. However, the two realized volatility time series (v_1 and v_2) only Granger-cause each other. This relationship is further illustrated in Figure 29, where we analyze the residuals from the VAR(1) model. The residuals exhibit instantaneous causality, highlighting the difficulty in accurately predicting any of the volatility time series due to this link in the residual terms. These findings underscore the key advantage of implied volatility over realized volatility: its superior predictive power for future volatility.

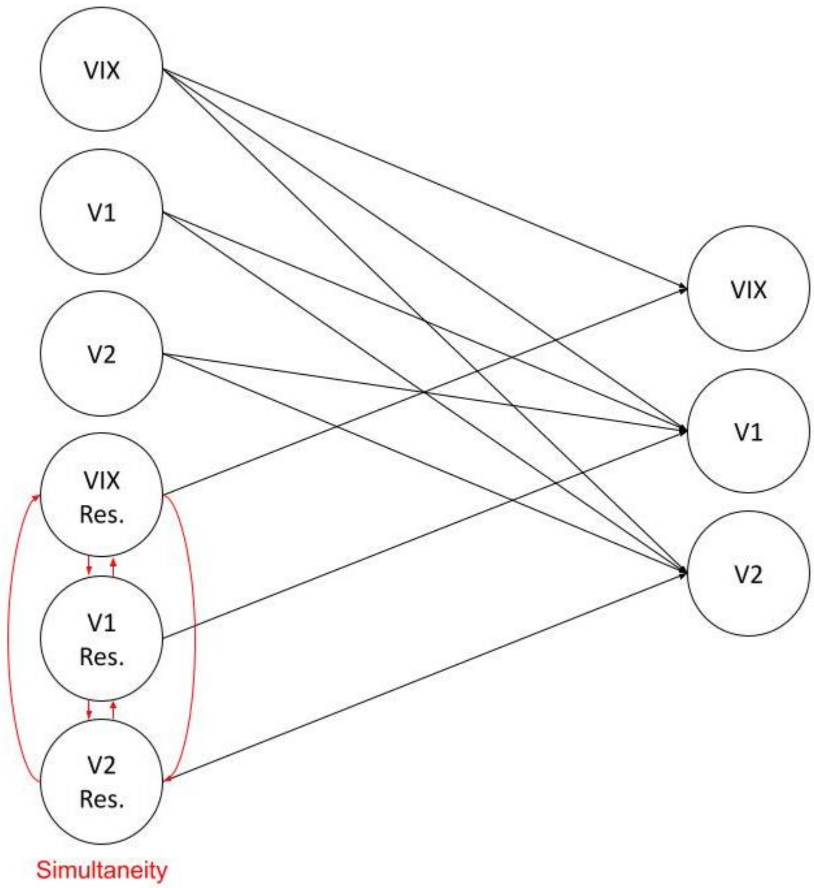


Figure 29 (Causality)

Conclusion

In conclusion, implied volatility and realized volatility are fundamentally different concepts. Implied volatility is derived from options prices, and under the assumption of the Efficient Market Hypothesis (EMH), it reflects the market's consensus on future volatility, incorporating all available information. In contrast, realized volatility is based solely on historical asset price movements and measures the actual past volatility. As a result, implied volatility is considered to have greater predictive power for future price movements, as it accounts for market expectations and current information, while realized volatility is limited to past data and price behavior.

References

Cboe Exchange, Inc. (2019). CBOE Volatility Index®. In *White Paper*. https://www.sfu.ca/~poitras/419_VIX.pdf

Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, *41*(1), 135. https://doi.org/10.2307/1913889

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, *96*(2), 246–273. https://doi.org/10.1086/261535

Saha, A., Malkiel, B. G., & Rinaudo, A. (2018). Has the VIX index been manipulated? *Journal of Asset Management*, *20*(1), 1–14. https://doi.org/10.1057/s41260-018-00102-4

Wikipedia contributors. (2024, September 3). *Implied volatility*. Wikipedia. https://en.wikipedia.org/wiki/Implied_volatility

Appendix

Code Available by Request