I. Abstract

This project explores a hybrid approach to options trading by combining forward-looking and backward-looking volatility measures to generate predictive trading signals. Specifically, we integrate implied volatility (IV) derived from the Black-Scholes-Merton (BSM) model with historical volatility forecasts produced by univariate time series models, particularly GARCH(1,1) with a skewed Student-t distribution. We use Apple Inc. (AAPL) and Walmart Inc. (WMT) as representative assets, growth and defensive stocks, respectively. We first back out BSM-implied volatilities from market option prices. We then develop a series of linear and nonlinear regressions to model the relationship between BSM IV and GARCH-based forecasts. The ultimate goal is to leverage deviations between observed and forecasted IVs to produce actionable trading signals within a simulated market framework. By forecasting IVs using a GARCH-based model and constructing option prices through the Black-Scholes formula, the strategy aims to replicate realistic trading conditions. These signals are then backtested under risk-managed scenarios, including stop-loss constraints, to assess profitability and stability.

II. Data Collection and Challenges

To support our analysis over the period December 13, 2024, to April 11, 2025, we collected a comprehensive dataset spanning both market prices and macroeconomic indicators. This included the adjusted closing prices of AAPL and WMT stock, which served as the underlying asset for both the Black-Scholes and univariate volatility models. For the options data, we focused on at-the-money (ATM) options, selecting strike prices approximately 5% above and below the stock's closing price to closely resemble realistic trading conditions. In addition, we gathered macro-financial variables, specifically VIX closing prices and CPI, believed to influence implied volatility but not captured directly by the Black-Scholes model. These variables were incorporated into our enhanced regression model to understand if they explain deviations in implied volatility derived from BSM and to improve the predictive performance of our volatility-based trading signals.

One of the challenges we encountered during data collection was sourcing historical options data for contracts that had already expired over a longer time horizon. This was essential for improving the accuracy and robustness of our analysis. However, Bloomberg only provided expired options data up to December 2024, limiting our ability to include a broader set of observations and thereby reducing the size of our dataset. Additionally, we faced difficulties in obtaining complete macroeconomic data, specifically for CPI, which are released on a monthly basis. As a result, we were only able to collect this data up to March 2025, leaving a gap for April 2025, which slightly constrained the completeness of our macroeconomic variable analysis.

III. Implied Volatility from Black-Scholes

To calculate the implied volatility (IV) for both call and put options on Walmart and Apple for our specific options expiration dates, we implemented the Black-Scholes model in Python and used a numerical root-finding method (brentq from scipy.optimize) to solve for volatility. The process began by loading cleaned historical options data, including bid and ask prices, spot prices, strike prices, expiration dates, and risk-free rates. Midpoint prices were computed from bid-ask spreads to approximate market value.

For each option, we calculated the time to maturity in years based on the expiration and current dates. The core of the process involved defining the Black-Scholes pricing formulas for calls and puts and then iteratively solving for the implied volatility that equates the model price to the market price. This was done separately for each row in the dataset. The resulting IVs were then stored and exported to a new Excel file, allowing for further analysis and integration into volatility modeling.

IV. Preliminary Time Series Analysis

In order to build an accurate regression model that forecasts implied volatility in the Black-Scholes model, we implemented a systematic GARCH modeling approach for one growth stock and one defensive stock, Apple (AAPL) and Walmart (WMT), respectively. This approach was taken to examine the results using two assets with different volatility characteristics. Our preliminary analysis aimed to identify the most appropriate univariate volatility models for each asset by evaluating a range of ARCH and GARCH specifications under varying distribution assumptions. The chosen models are used to forecast volatility for the underlying assets (AAPL and WMT) on the days that the historical options for these same assets have expired. Subsequently, the forecast from the chosen model also becomes a parameter in the following linear regression analysis.

We start our analysis by computing the daily log returns from the adjusted closing prices for both AAPL and WMT. To assess volatility clustering and the presence of ARCH effects, we examined the autocorrelation (ACF) and partial autocorrelation (PACF) plots of the squared returns. In both cases, the Ljung-Box test rejected the null hypothesis of no autocorrelation in the squared returns, supporting the use of ARCH-type models, therefore there is autocorrelation.

- For AAPL, the PACF of squared returns cuts off at lag 9, from our observations. Though we could have also used lag 4 as there is a sharp cut off there, we decide to overspecify since AAPL is a more volatile stock, to capture hidden dynamics and test significance through lags (Figure 1).
- For WMT, the PACF cuts off at lag 1 (Figure 2).

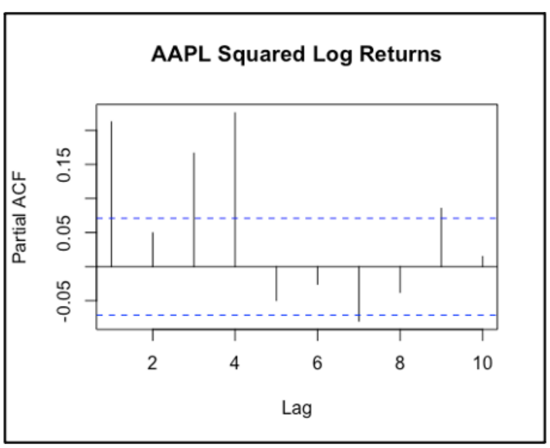


Figure 1

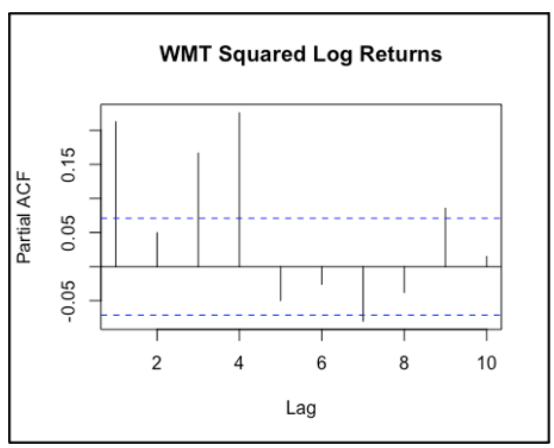


Figure 2

Based on the PACF, we initially fit an ARCH(9) model for AAPL. However, several coefficients were statistically insignificant, prompting a reduction to an ARCH(1) model. For

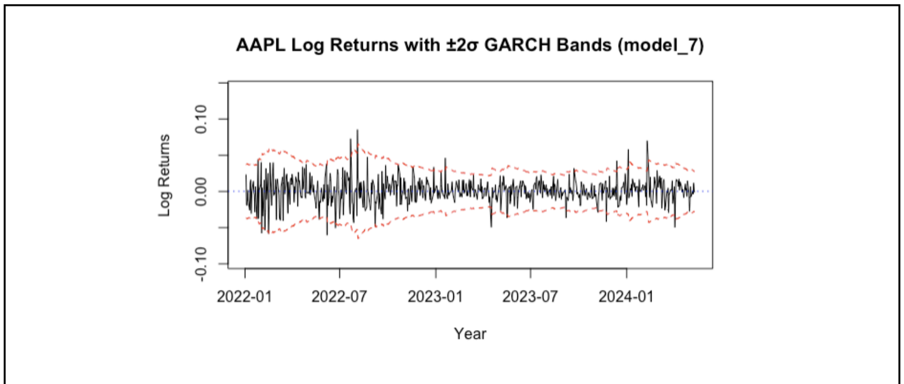
WMT, we fit an ARCH(1). While both ARCH(1) models yielded statistically significant coefficients, residual diagnostics such as the Ljung-Box test on squared residuals implied that there was autocorrelation in the conditional variance of the ARCH(1) models. Additionally, for both AAPL and WMT, the Jarque-Bera and Shapiro-Wilk tests rejected the null hypothesis of normality, indicating that residuals exhibited skewness and heavy tails. These results highlighted the need for more flexible error distributions.

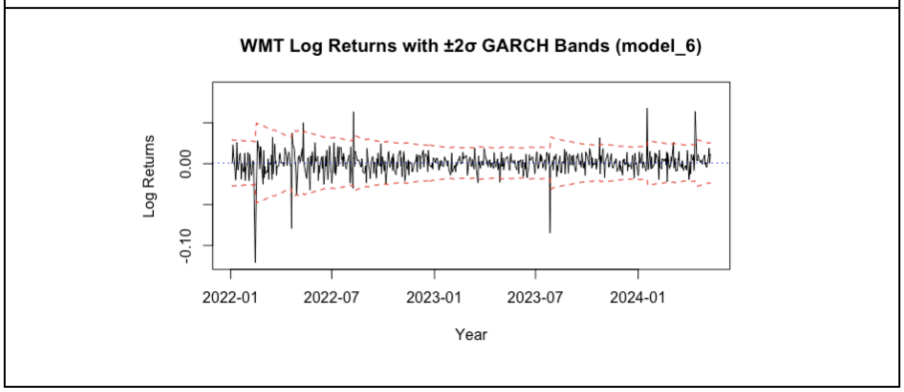
Thus, we changed distribution assumptions and ran an ARCH(1) under Student t-distribution and under a Skewed Student t-distribution (SSTD). The Skewed Student t-distribution improved fit and delivered the lowest AIC values for both stocks under the ARCH framework. More specifically, for WMT, the ARCH(1) under an SSTD effectively captured volatility dynamics and passed all residual diagnostics checks. On the other hand, the same model for AAPL showed lingering autocorrelation. This motivated a transition to a GARCH(1,1) specification to explore whether a more flexible model could capture volatility dynamics better and improve overall fit.

Next, we estimated a GARCH(1,1) model under a normal, t, and skewed t-distributions. For both assets, all estimated coefficients were statistically significant. Moreover, the ACF and PACF plots of standardized and squared residuals confirmed that autocorrelation was eliminated. All GARCH(1,1) models passed their residual diagnostics. Thus, residuals behaved as white noise and volatility clustering was captured. The models were then evaluated using the Akaike Information Criterion (AIC), with the GARCH(1,1) under a SSTD yielding the lowest AIC values. Below is a table with a summary of AIC comparisons:

Model Specification	AAPL AIC	WMT AIC
ARCH(1), normal	-5.259	-5.712
ARCH(1), t	-5.352	-6.042
ARCH(1), skewed t	-5.352	-6.044
GARCH(1,1), normal	-5.315	-5.746
GARCH(1,1), t	-5.408	-6.077
GARCH(1,1), skewed t	-5.097	-6.080

For the final models, we looked at the ACF and PACF plots of residuals and squared residuals, which showed no significant lags. Additionally, a time series plot of conditional standard deviations revealed the expected clustering behavior, and the Q-Q plots of standardized residuals closely matched the theoretical quantiles from the SSTDs. We also visualized GARCH $\pm 2\sigma$ volatility bands overlaid on the return series (shown below) to confirm how well each model captured volatility dynamics.





Lastly, using the GARCH(1,1)-SSTD models, we computed the one-step-ahead volatility forecasts both manually and via the predict() function. Results were highly consistent and validated the model's stability for forward-looking volatility estimation. Hence, based on the PACF-guided order selection, AIC comparisons, statistical significance, and residual diagnostics, we conclude that the GARCH(1,1)-SSTD assumptions is the best-fitting model for both AAPL and WMT. These models will serve as the volatility forecasting engines for our broader implied volatility regression and option pricing analyses.

V. Forecasting Using GARCH(1,1)-SSTD

The next step in the experiment is to forecast the volatility of AAPL and WMT stock returns for our option expiration dates. To achieve this, a rolling forecasting method is employed using the GARCH(1,1)-SSTD model — a widely used econometric model for capturing time-varying volatility in financial markets.

For each forecast date, the approach begins by collecting the previous two years of AAPL's and WMT's daily return data. This rolling window ensures that each volatility prediction is informed by the most recent market behavior leading up to that date, rather than relying on the full historical dataset, which may include outdated or irrelevant information.

The GARCH(1,1)-SSTD model is then applied to this two-year data window. In this model, current volatility is assumed to depend on two key factors: the magnitude of the most recent return shock (squared return) and the previous period's forecasted volatility. This structure captures the phenomenon of volatility clustering, where large market movements are typically followed by more large movements, and calm periods tend to persist.

Once the GARCH model is fitted to the training data, it is used to produce a one-step-ahead forecast of volatility — that is, the expected volatility for the next trading day, which aligns with the option expiration date. This process is repeated for each of the specified dates, yielding a series of volatility predictions that are customized for the market conditions preceding each date.

This method is particularly effective in the context of options, where implied volatility is a critical component of pricing models. By using a GARCH-based forecast that dynamically adjusts to market conditions, the predicted volatilities provide a more realistic and responsive input for option valuation compared to static or long-term average estimates.

VI. Linear Regression Methodology

Once we had compiled the data for expired options ranging from December 2024 to April 2025 and extracted implied volatilities from market option prices using the Black-Scholes-Merton (BSM) model, we aligned these with the one-step-ahead volatility forecasts produced by our selected GARCH models. Specifically, we computed the average of the BSM-implied volatilities for both calls and puts on each trading day and matched them to the GARCH-forecasted volatility for that same day. This produced a clean, time-aligned dataset for calls and puts for our regression analysis.

Our initial modeling approach was to regress BSM-implied volatility on GARCH-implied volatility separately for call and put options, using a simple linear regression framework:

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \varepsilon$$

However, in many cases, this model yielded statistically insignificant coefficients and low explanatory power. We first attempted to do our analysis based on a rolling regression, but due to the small number of observations in each rolling window, many of our coefficients were insignificant. Thus, we decided to proceed with the entire sample size. Nevertheless, to improve model performance and recover statistically meaningful relationships, we systematically enhanced the model through a series of extensions aimed at increasing flexibility and accounting for non-linearities or omitted macro factors.

To increase model flexibility, we first introduced a polynomial term to the regression. This specification allows us to capture curvature in the relationship between GARCH volatility and BSM-implied volatility, which is plausible given the convexity in option pricing and potential regime shifts in market volatility.

We also experimented with log-transforming the explanatory variable, IV_GARCH, to linearize exponential or multiplicative relationships and stabilize variance. In certain financial settings, proportional changes in volatility may impact options pricing more than absolute changes, making log transformations theoretically appropriate.

Additionally, we fit a generalized linear model (GLM) with a Gaussian family and identity link to test whether relaxing distributional assumptions could yield a better fit. Although the GLM does not change the functional form of the regression, it improves robustness to non-normality and provides a flexible framework for future modeling enhancements.

If these transformations and re-specifications did not lead to statistically significant predictors, we expanded the information set by introducing macroeconomic and market-wide variables. We first check the residuals of these explanatory variables and they are non-linear, so we apply the transformations applied above to IV_GARCH to these as well. We added the VIX, which captures market-wide expectations of volatility and may supplement stock-specific GARCH forecasts by explaining systematic volatility shocks. We also incorporated the Consumer Price Index (CPI) to proxy for inflation trends, which can affect investor risk appetite, expected future interest rates, and ultimately, implied volatilities in options pricing. As a note, we checked the residuals of the other two independent variables, VIX and CPI. As a note, for polynomial logarithmic transformation linear regression, we will perform logarithmic transformation on both of these variables if or when they are included in the regression since their residuals are non linear.

Each model, if it produced significant parameters, was then evaluated using the classical Gauss-Markov assumptions. We tested for homoscedasticity using the Breusch-Pagan test, for residual autocorrelation using the Durbin-Watson test, and for model stability using the Chow test at key breakpoints. If a model passed all diagnostic tests and its coefficients were statistically significant, we considered it valid for use in trading strategy development. Otherwise, the model was discarded as unreliable. This systematic and layered modeling approach ensures that only robust, well-specified models guide our volatility-based trading signals. By progressively refining model structure and inputs from linear forms to polynomial and macro-augmented regressions, we aim to capture the true underlying relationship between forecasted and market-implied volatility.

Below are all the models that were attempted and their results:

a. *WMT CALLS*

i. *Simple Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch, data = calls_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.12109 -0.10888 -0.09631  0.07271  0.35932

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1923     0.1393   1.381    0.188
IV_Garch     -4.1436     9.9142  -0.418    0.682

Residual standard error: 0.1748 on 15 degrees of freedom
Multiple R-squared:  0.01151,	Adjusted R-squared:  -0.05439
F-statistic: 0.1747 on 1 and 15 DF,  p-value: 0.6819
```

The coefficients are statistically insignificant, and the F-statistic ($p = 0.6819$) shows the overall model does not do a good job at explaining the dependent variable. Furthermore, $R^2 = 0.01151$ and Adjusted $R^2 = -0.05439$, confirming poor fit.

ii. *Logarithmic Transformation Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \varepsilon$$

```
> summary(lm_log)

Call:
lm(formula = IV_BSM ~ log_IV_Garch, data = calls_df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.11897 -0.10724 -0.09467  0.07374  0.36030 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.09329    0.66716  -0.140    0.891
log_IV_Garch -0.05284    0.15285  -0.346    0.734

Residual standard error: 0.1751 on 15 degrees of freedom
Multiple R-squared:  0.007903,	Adjusted R-squared:  -0.05824 
F-statistic: 0.1195 on 1 and 15 DF,  p-value: 0.7344
```

Here the coefficients are statistically insignificant, and the F-statistic ($p = 0.7344$) shows the overall model lacks explanatory power. Furthermore, $R^2 = 0.007903$ and Adjusted $R^2 = -0.05824$, confirming this model is also not a good fit.

iii. Quadratic Model Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \varepsilon$$

```
> summary(call_model_quad)

Call:
lm(formula = IV_BSM ~ IV_Garch + I(IV_Garch^2), data = calls_df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.13717 -0.11774 -0.09159  0.08659  0.37384 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)     -0.0807     0.5123  -0.158    0.877
IV_Garch        32.6459    67.0769   0.487    0.634
I(IV_Garch^2) -1111.1742  2002.6241  -0.555    0.588

Residual standard error: 0.1789 on 14 degrees of freedom
Multiple R-squared:  0.03278,	Adjusted R-squared:  -0.1054 
F-statistic: 0.2372 on 2 and 14 DF,  p-value: 0.7919
```

None of the coefficients are statistically significant ($p > 0.05$), and the F-statistic ($p = 0.7919$) shows the overall model lacks explanatory power. Also, the $R^2 = 0.03278$ and adjusted $R^2 = -0.1054$, confirming poor fit and possible overfitting with the quadratic term.

iv. GLM (Generalized Linear Modeling) Regression

```
> summary(glm_call)

Call:
glm(formula = IV_BSM ~ IV_Garch, family = gaussian(link = "identity"),
    data = calls_df)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1923     0.1393   1.381    0.188
IV_Garch     -4.1436     9.9142  -0.418    0.682

(Dispersion parameter for gaussian family taken to be 0.03053988)

    Null deviance: 0.46343  on 16  degrees of freedom
Residual deviance: 0.45810  on 15  degrees of freedom
AIC: -7.1921

Number of Fisher Scoring iterations: 2
```

In conclusion, across all models, IV_GARCH shows no statistically significant effect on IV_BSM. While classical Gauss-Markov assumptions largely hold, the low R^2 and poor residual normality suggest these models are inadequate in capturing the underlying volatility dynamics. Alternative features or nonlinear frameworks may be necessary.

b. *WMT CALLS + VIX*

i. *Simple Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times VIX + \varepsilon$$

```
> summary(lm_model_vix)

Call:
lm(formula = IV_BSM ~ IV_Garch + VIX, data = df_calls_vix)

Residuals:
     Min       1Q   Median       3Q      Max
-0.16478 -0.10905 -0.08018  0.11443  0.39397

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.05983    0.38169  -0.157    0.878
IV_Garch     3.43881   30.33304   0.113    0.912
VIX          0.00953    0.03398   0.281    0.784

Residual standard error: 0.1863 on 12 degrees of freedom
Multiple R-squared:  0.03315,	Adjusted R-squared:  -0.128
F-statistic: 0.2057 on 2 and 12 DF,  p-value: 0.8169
```

The intercept estimate is -0.05983 ($p = 0.878$), and the coefficients for IV_GARCH and VIX are 3.43881 ($p = 0.912$) and 0.00953 ($p = 0.784$), respectively. We can see that these results are not statistically significant. The R-squared value is very low (3.32%), indicating that the model explains minimal variation in IV_BSM.

ii. *Logarithmic Transformation Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \beta_2 \times log(VIX) + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ log_IV_Garch + log_VIX, data = df_calls_vix)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.17557 -0.11066 -0.08746  0.12430  0.38240 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.45343    3.43417   0.132    0.897
log_IV_Garch  0.09964    0.40611   0.245    0.810
log_VIX       0.04887    0.63055   0.078    0.940

Residual standard error: 0.1873 on 12 degrees of freedom
Multiple R-squared:  0.0227,	Adjusted R-squared:  -0.1402 
F-statistic: 0.1394 on 2 and 12 DF,  p-value: 0.8713
```

The intercept estimate is 0.45343 ($p = 0.897$), and the coefficients are 0.09964 ($p = 0.810$) and 0.04887 ($p = 0.940$), respectively. These results are clearly not statistically significant, as all p-values are well above conventional thresholds. The R-squared value is extremely low at 2.3%, indicating that the model explains virtually none of the variation in IV_BSM. Additionally, the adjusted R-squared is negative (-0.1402), suggesting that the model performs worse than a simple mean model. The overall F-statistic is 0.1394 with a p-value of 0.8713, further confirming that the model lacks explanatory power and is not statistically meaningful.

iii. Polynomial Model Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \beta_3 \times VIX + \beta_4 \times (VIX)^2 + \varepsilon$$

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.23609 -0.06888 -0.03249  0.04881  0.24650 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)  
(Intercept)  6.146e+00  2.485e+00   2.474   0.0329 *
IV_Garch    -1.044e+02  3.117e+02  -0.335   0.7446  
VIX         -5.922e-01  2.499e-01  -2.370   0.0393 *
IV_Garch_sq  3.561e+03  1.123e+04   0.317   0.7577  
VIX_sq       1.632e-02  6.636e-03   2.460   0.0337 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1548 on 10 degrees of freedom
Multiple R-squared:  0.444,	Adjusted R-squared:  0.2215 
F-statistic: 1.996 on 4 and 10 DF,  p-value: 0.1712
```

The results show that while the coefficients for IV_Garch and IV_Garch² are statistically insignificant ($p = 0.7446$ and 0.7577, respectively), both the linear and squared terms of VIX are significant at the 5% level ($p = 0.0393$ and 0.0337). This indicates that VIX has both a direct and nonlinear influence on implied volatility. The model explains approximately 44.4% of the variation in IV_BSM ($R^2 = 0.444$), with an adjusted R^2 of 0.2215. The model as a whole is insignificant since the overall F-statistic (1.996, $p = 0.1712$) has a p-value greater than 0.05.

iv. GLM (Generalized Linear Modeling) Regression

```
Call:
glm(formula = IV_BSM ~ IV_Garch + VIX, family = gaussian(link = "identity"),
    data = df_calls_vix)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.05983    0.38169  -0.157    0.878
IV_Garch     3.43881   30.33304   0.113    0.912
VIX          0.00953    0.03398   0.281    0.784

(Dispersion parameter for gaussian family taken to be 0.03470416)

    Null deviance: 0.43073  on 14  degrees of freedom
Residual deviance: 0.41645  on 12  degrees of freedom
AIC: -3.1924

Number of Fisher Scoring iterations: 2
```

The Generalized Linear Model (GLM) regression results indicate that the model provides minimal explanatory power for the Black-Scholes implied volatility of WMT call options. None of the predictors—IV_GARCH, or the VIX—are statistically significant, as reflected by their high p-values (all above 0.05).

c. WMT CALLS + VIX + CPI

i. Simple Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times CPI + \beta_3 \times VIX + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch + VIX + CPI, data = df_calls_cpi)

Residuals:
     Min       1Q   Median       3Q      Max
-0.18581 -0.08597 -0.06329  0.06470  0.43939

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 23.11965   24.28081   0.952    0.361
IV_Garch    14.48625   32.57008   0.445    0.665
VIX          0.00363    0.03466   0.105    0.918
CPI         -0.07271    0.07615  -0.955    0.360

Residual standard error: 0.187 on 11 degrees of freedom
Multiple R-squared:  0.1071,    Adjusted R-squared:  -0.1364
F-statistic:  0.44 on 3 and 11 DF,  p-value: 0.729
```

All predictors in the model are statistically insignificant, with high p-values: 0.361 for the intercept, 0.665 for IV_GARCH, 0.360 for CPI, and 0.918 for VIX. The overall model fit is extremely weak, with an R^2 value of only 0.1071, indicating that just about 11% of the variability in implied volatility is explained by the included predictors. The F-statistic (0.44, p = 0.729) further confirms that the model as a whole is not significant.

ii. Logarithmic Transformation Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \beta_2 \times log(CPI) + \beta_3 \times log(VIX) + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ log_IV_Garch + log_VIX + log_CPI, data = df_calls_cpi)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.19612 -0.09170 -0.06859  0.06487  0.42744 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  139.1646   142.4850   0.977    0.350
log_IV_Garch   0.2603     0.4392   0.593    0.565
log_VIX       -0.0797     0.6456  -0.123    0.904
log_CPI      -23.8698    24.5121  -0.974    0.351

Residual standard error: 0.1877 on 11 degrees of freedom
Multiple R-squared:  0.1003,	Adjusted R-squared:  -0.1451 
F-statistic: 0.4086 on 3 and 11 DF,  p-value: 0.75
```

All coefficients are statistically insignificant, with p-values of 0.350 for the intercept, 0.565 for log IV_GARCH, 0.351 for log CPI, and 0.904 for log VIX. The model shows a very weak fit, as reflected by an R^2 of 0.1003 and an adjusted R^2 of -0.1451, indicating that the model explains less variation in IV_BSM than a simple mean model would. The F-statistic (0.4086, p = 0.75) further confirms that the model as a whole is not statistically significant.

iii. Polynomial Model Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \beta_3 \times VIX + \beta_4 \times (VIX)^2 + \beta_5 \times CPI + \beta_6 \times (CPI)^2 + \varepsilon$$

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.15154 -0.05341 -0.02454  0.04892  0.21788 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)   
(Intercept)  1.553e+04  7.443e+03   2.087  0.07036 . 
IV_Garch     2.459e+01  2.544e+02   0.097  0.92538   
VIX         -6.854e-01  2.035e-01  -3.368  0.00981 **
CPI         -9.735e+01  4.671e+01  -2.084  0.07068 . 
IV_Garch_sq -7.749e+02  9.122e+03  -0.085  0.93439   
VIX_sq       1.860e-02  5.382e-03   3.456  0.00862 **
CPI_sq       1.526e-01  7.329e-02   2.082  0.07090 . 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1234 on 8 degrees of freedom
Multiple R-squared:  0.717,	Adjusted R-squared:  0.5048 
F-statistic: 3.378 on 6 and 8 DF,  p-value: 0.05778
```

The p-values of all the other variables except VIX and VIX_sq are above the conventional significance thresholds ($p<0.05$) and statistically insignificant. The overall model is not significant since the p-value of the F-statistic is 0.05778 and IV_GARCH is not significant with

IV_BSM, so the polynomial model fails to capture the relationship between the predictors and implied volatility.

iv. *GLM (Generalized Linear Modeling) Regression*

```
Call:
glm(formula = IV_BSM ~ IV_Garch + VIX + CPI, family = gaussian(link = "identity"),
    data = df_calls_cpi)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 23.11965   24.28081   0.952    0.361
IV_Garch    14.48625   32.57008   0.445    0.665
VIX          0.00363    0.03466   0.105    0.918
CPI         -0.07271    0.07615  -0.955    0.360

(Dispersion parameter for gaussian family taken to be 0.03496181)

    Null deviance: 0.43073  on 14  degrees of freedom
Residual deviance: 0.38458  on 11  degrees of freedom
AIC: -2.3867

Number of Fisher Scoring iterations: 2
```

The Generalized Linear Model (GLM) regression results indicate that the model provides minimal explanatory power for the Black-Scholes implied volatility of WMT call options. None of the predictors—IV_GARCH, the Consumer Price Index (CPI), or the VIX—are statistically significant, as reflected by their high p-values (all above 0.05).

d. *WMT PUTS*

i. *Simple Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch, data = puts_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.15388 -0.10123 -0.04344  0.09210  0.26523

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.0812     0.1153   0.704    0.492
IV_Garch      6.8683     8.2102   0.837    0.416

Residual standard error: 0.1447 on 15 degrees of freedom
Multiple R-squared:  0.04458,	Adjusted R-squared:  -0.01912
F-statistic: 0.6998 on 1 and 15 DF,  p-value: 0.416
```

Both the intercept and slope are statistically insignificant, with p-values of 0.492 and 0.416 respectively. The R^2 value is very low (≈ 4.5%), suggesting that IV_GARCH alone poorly explains the variation in IV_BSM. These findings imply that a more flexible modeling approach, such as a quadratic regression, GLM, or using non-parametric techniques, may better capture the relationship between IV_BSM and IV_GARCH.

ii. Polynomial Model Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch + I(IV_Garch^2), data = puts_df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.12284 -0.03684 -0.02347  0.01929  0.22949 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)   
(Intercept)      1.1357     0.3128   3.631  0.00273 **
IV_Garch      -135.2223    40.9572  -3.302  0.00525 **
I(IV_Garch^2) 4291.6362  1222.8030   3.510  0.00347 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1093 on 14 degrees of freedom
Multiple R-squared:  0.4918,    Adjusted R-squared:  0.4191 
F-statistic: 6.773 on 2 and 14 DF,  p-value: 0.00876
```

All coefficients are statistically significant at the 1% level, indicating a strong non-linear relationship between GARCH volatility and BSM implied volatility. The adjusted R^2 of 0.4191 reflects a substantial improvement in explanatory power compared to the linear model. We will proceed with the Gauss Markov assumptions to check whether the OLS estimator is the best linear unbiased estimator.

iii. Polynomial Model Linear Regression: Investigation of the Gauss Markov Assumptions

- Linearity in terms of parameters – even though the relationship between IV_BSM and IV_GARCH is nonlinear, the model is linear in the parameters.
- Multicollinearity – Upon conducting regression analysis, we get a valid output. Hence there is no perfect multicollinearity between variables.
- Independence & Randomness of Error terms –
  The performed Durbin-Watson test for positive autocorrelation within this model, we consider $\varepsilon_i = a\varepsilon_{i-1} + W_i$, where a is a deterministic factor and $\{W_i\}$ is a stochastic process, $i = 1, .., 17$. Within this model, we test our hypothesis:

  $$H_0 : \alpha = 0 \text{ (no autocorrelation)}$$

  $$H_1 : \alpha > 0 \text{ (positive autocorrelation)}$$

  The test yielded a Durbin-Watson statistic of 2.6327 and a p-value of 0.8246. At a significance level of $\alpha = 0.05$, we fail to reject the null hypothesis that the residuals are uncorrelated. This implies that there is no statistically significant evidence of positive autocorrelation in the model residuals.
- Zero Conditional Mean – We use residuals to estimate errors. As observed from the sum of residuals being equal to 0, there is zero conditional mean.

- Homoscedasticity – For formal analysis, we will perform the Breusch-Pagan test within this model. Consider an auxiliary regression: $\varepsilon^2 = b_0 + b_1 x_1 + b_2 x_2$, we test our hypothesis:

$$H_0 : b_1 = b_2 = 0$$

$$H_1 : \textit{at least one } b_i, \textit{ where } i = 1, 2, \neq 0 \quad \text{(heteroscedastic)}$$

The test yields a p-value of 0.2678, which is greater than 0.05, therefore there is enough evidence to NOT reject the null hypothesis of constant variance. This suggests that there is not enough evidence of heteroskedasticity.

- Normality of Error terms – We pursue the Shapiro-Wilk, so we test our hypothesis:

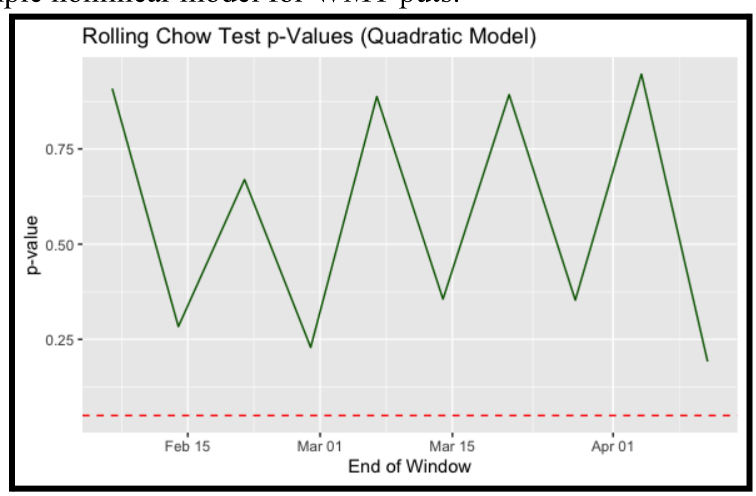
H_0 : data are normal distributed
H_1 : data are NOT normally distributed

This produces W = 0.86737 and a p-value = 0.01996, so we reject the null hypothesis, suggesting a statistically significant departure from normality.

- Stability in Regression - Though this is not a Gauss Markov assumption, we decided to test for any structural breaks in our significant model and determine whether coefficients in two linear regressions on different datasets are equal via the Chow test. We test our hypothesis:

H_0 : Coefficients are the same across both subsamples (no structural break)
H_1 : Coefficients are different (there is a structural break)

We also used a rolling regression of the WMT polynomial regression here and applied a rolling Chow test to different points in our data. The results below show that none of the rolling p-values follow below the 5% threshold. Hence, there is no evidence of significant breaks in the relationship between IV_GARCH and IV_BSM. This supports the use of a single, full-sample nonlinear model for WMT puts.



The quadratic model provides a statistically significant and better-fitting relationship between IV_BSM and IV_GARCH. It addresses some of the issues from the linear model, such as poor fit and heteroskedasticity. Minor normality concerns remain but are not severe. Overall, this model is more appropriate and stable for modeling the implied volatility of WMT puts.

Our conclusion makes sense since WMT is a stable, low-volatility stock so its implied volatility tends to respond nonlinearly to market signals. Also, put options often react strongly to downside risk and the nonlinear response in implied volatility can be better captured by quadratic terms. Last, we can see that at lower levels of GARCH-predicted volatility, IV_BSM may stay stable, but at higher levels, sensitivity increases or flattens depending on investor sentiment, causing curvature in the implied volatility relationship.

e. *APPL CALLS*

i. *Simple Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch, data = df_calls)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.51714 -0.26593 -0.19081 -0.09664  1.58463 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)  -0.3791     0.2847  -1.331  0.20291   
IV_Garch     49.3817    15.2087   3.247  0.00542 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5439 on 15 degrees of freedom
Multiple R-squared:  0.4127,	Adjusted R-squared:  0.3736 
F-statistic: 10.54 on 1 and 15 DF,  p-value: 0.005417
```

Here the coefficient is statistically significant, and the F-statistic ($p = 0.005417$) shows that the overall model has explanatory power. Furthermore, $R^2 = 0.4127$ and Adjusted $R^2 = 0.3736$, confirming this model could be a good fit.

ii. *Simple Linear Regression: Investigation of the Gauss Markov Assumptions*

- Linearity in terms of parameters – the relationship between IV_BSM and IV_GARCH is linear and the model is linear in the parameters.
- Multicollinearity – Upon conducting regression analysis, we get a valid output. Hence there is no perfect multicollinearity between variables.
- Independence & Randomness of Error terms –
  The performed Durbin-Watson test for positive autocorrelation within this model, we consider $\varepsilon_i = a\varepsilon_{i-1} + W_i$, where a is a deterministic factor and $\{W_i\}$ is a stochastic process, $i = 1, .., 17$. Within this model, we test our hypothesis:

$$H_0 : \alpha \ = \ 0 \text{ (no autocorrelation)}$$

$$H_1 : \alpha \ > \ 0 \text{ (positive autocorrelation)}$$

The test yielded a Durbin-Watson statistic of 2.0757 and a p-value of 0.4824. At a significance level of $\alpha = 0.05$, we fail to reject the null hypothesis that the residuals are uncorrelated. This implies that there is no statistically significant evidence of positive autocorrelation in the model residuals.

- Zero Conditional Mean – We use residuals to estimate errors. As observed from the sum of residuals being equal to 0, there is zero conditional mean.
- Homoscedasticity – For formal analysis, we will perform the Breusch-Pagan test within this model. Consider an auxiliary regression: $\varepsilon^2 = b_0 + b_1 x_1 + b_2 x_2$, we test our hypothesis:

$$H_0 : b_1 = b_2 = 0$$
$$H_1 : at\ least\ one\ b_i,\ where\ i = 1, 2,\ \neq 0 \quad \text{(heteroscedastic)}$$

The test yields a p-value of 0.8954, which is greater than 0.05, therefore there is enough evidence to NOT reject the null hypothesis of constant variance. This suggests that there is not enough evidence of heteroskedasticity.

- Normality of Error terms – We pursue the Shapiro-Wilk, so we test our hypothesis:

H_0 : data are normal distributed
H_1 : data are NOT normally distributed

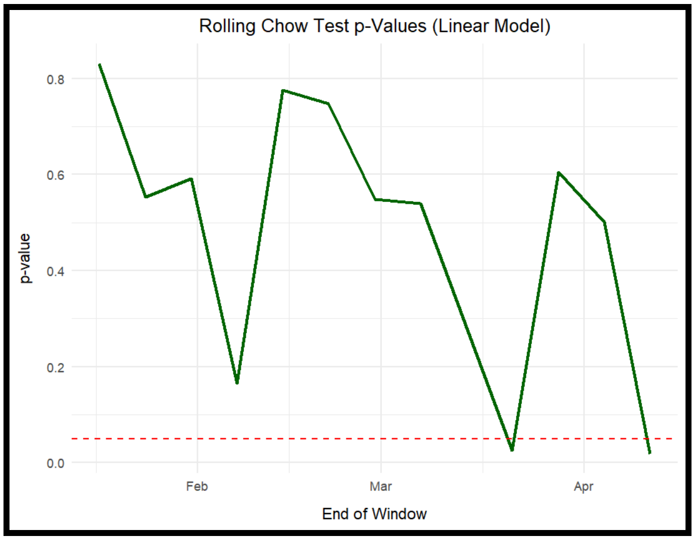
This produces W = 0.69567 and a p-value =0.0001033, so we reject the null hypothesis, suggesting a statistically significant departure from normality.

- Stability in Regression - Though this is not a Gauss Markov assumption, we decided to test for any structural breaks in our significant model and determine whether coefficients in two linear regressions on different datasets are equal via the Chow test. We test our hypothesis:

H_0 : Coefficients are the same across both subsamples (no structural break)
H_1 : Coefficients are different (there is a structural break)

The F value is 0.1143285 and the p-value was found to be greater than 0.05, so we do not reject the null hypothesis, which means that there is no structural break in the regression and shows stability.

We also used a rolling regression of the AAPL simple linear regression here and applied a rolling Chow test to different points in our data. The results show that one of the rolling p-values fall below the 5% threshold. However, after running the Cumulative Sum Test, it was found that the regression coefficients are stable gradually over time ($p>0.05$) and the Chow-type Structural Break Test finds that there is no structural break at any point ($p>0.05$). Hence, there is insufficient evidence of significant breaks in the relationship between IV_GARCH and IV_BSM. This supports the use of a single, full-sample linear model for AAPL calls.



The simple regression model provides a statistically significant and better-fitting relationship between IV_BSM and IV_GARCH. Minor normality concerns remain but are not severe. Overall, this model is more appropriate and stable for modeling the implied volatility of AAPL calls.

Our conclusion makes sense since AAPL is a high-beta growth stock. AAPL's options are actively traded and its implied volatility is more sensitive to market-wide volatility. Market participants respond to volatility in a more proportional manner due to high liquidity and institutional trading. Calls show a smoother implied volatility behavior especially since they are used more for speculation or leverage rather than downside protection, so they might not exhibit the same sharp shifts in implied volatility as puts. Last, for AAPL, increases in model-predicted volatility track linearly with option-implied volatility, without needing to adjust for curvature.

f. AAPL PUTS

i. Simple Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch, data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4009 -0.3293 -0.2700  0.1465  1.1045 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)    0.537      0.271   1.981   0.0662 .
IV_Garch      -5.276     14.477  -0.364   0.7206  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5178 on 15 degrees of freedom
Multiple R-squared:  0.008778,	Adjusted R-squared:  -0.0573 
F-statistic: 0.1328 on 1 and 15 DF,  p-value: 0.7206
```

Here the coefficient is statistically insignificant, and the F-statistic (p = 0.7206) shows that the overall model does not have explanatory power. Furthermore, $R^2 = 0.008778$ and Adjusted $R^2 = -0.0573$ confirms that this model is not a good fit.

ii. *Logarithmic Transformation Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ log(IV_Garch), data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.3938 -0.3273 -0.2729  0.1393  1.1083 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -0.06395    1.57931  -0.040    0.968
log(IV_Garch) -0.12303    0.37723  -0.326    0.749

Residual standard error: 0.5182 on 15 degrees of freedom
Multiple R-squared:  0.007041,	Adjusted R-squared:  -0.05916 
F-statistic: 0.1064 on 1 and 15 DF,  p-value: 0.7488
```

The coefficients are not statistically significant, indicating that the predictors do not have a meaningful impact on IV_BSM. Additionally, the F-statistic (p-value = 0.7488) suggests that the model, as a whole, is not statistically significant, meaning it doesn't explain the variability in IV_BSM effectively. The $R^2 = 0.007041$ and Adjusted $R^2 = -0.05916$ further confirm that the model has very poor explanatory power.

iii. *Quadratic Linear Regression*

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ poly(IV_Garch, 2, raw = TRUE), data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.3995 -0.3294 -0.2714  0.1441  1.1049 

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                      0.5566     1.1965   0.465    0.649
poly(IV_Garch, 2, raw = TRUE)1  -7.0796   107.9898  -0.066    0.949
poly(IV_Garch, 2, raw = TRUE)2  29.4002  1743.7159   0.017    0.987

Residual standard error: 0.5359 on 14 degrees of freedom
Multiple R-squared:  0.008799,	Adjusted R-squared:  -0.1328 
F-statistic: 0.06214 on 2 and 14 DF,  p-value: 0.94
```

The coefficients are statistically insignificant, and the F-statistic (p = 0.94) shows the overall model lacks explanatory power. Furthermore, $R^2 = 0.008799$ and Adjusted $R^2 = -0.1328$ confirming this model is also not a good fit.

iv. GLM (Generalized Linear Modeling) Regression

```
Call:
glm(formula = IV_BSM ~ IV_Garch, data = df)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.537      0.271   1.981   0.0662 .
IV_Garch      -5.276     14.477  -0.364   0.7206
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.2680745)

    Null deviance: 4.0567  on 16  degrees of freedom
Residual deviance: 4.0211  on 15  degrees of freedom
AIC: 29.736

Number of Fisher Scoring iterations: 2
```

The Generalized Linear Model (GLM) regression results indicate that the model provides minimal explanatory power for the Black-Scholes implied volatility of AAPL call options.

g. AAPL PUTS + VIX

i. Simple Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times VIX + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch + VIX, data = df)

Residuals:
    Min      1Q  Median      3Q     Max
-0.4325 -0.3383 -0.2370  0.1774  1.0837

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.76387    0.33651   2.270   0.0395 *
IV_Garch    11.24896   20.58434   0.546   0.5933
VIX         -0.02452    0.02189  -1.120   0.2815
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5134 on 14 degrees of freedom
Multiple R-squared:  0.09031,	Adjusted R-squared:  -0.03965
F-statistic: 0.6949 on 2 and 14 DF,  p-value: 0.5155
```

Here the coefficient is statistically insignificant, and the F-statistic ($p = 0.5155$) shows that the overall model does not have explanatory power. Furthermore, $R^2 = 0.09031$ and Adjusted $R^2 = -0.03965$, confirming this model is not a good fit.

ii. Logarithmic Transformation Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \beta_2 \times log(VIX) + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ log(IV_Garch) + log(VIX), data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4668 -0.3140 -0.1581  0.2208  1.0933 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     5.0052     4.0387   1.239    0.236
log(IV_Garch)   0.4897     0.5818   0.842    0.414
log(VIX)       -0.8486     0.6252  -1.357    0.196

Residual standard error: 0.5043 on 14 degrees of freedom
Multiple R-squared:  0.1225,	Adjusted R-squared:  -0.002854 
F-statistic: 0.9772 on 2 and 14 DF,  p-value: 0.4006
```

The coefficients are statistically insignificant, and the F-statistic ($p = 0.4006$) shows the overall model does not do a good job at explaining IV_BSM. Furthermore, $R^2 = 0.1225$ and Adjusted $R^2 = -0.002854$ confirms that this model is also not a good fit.

iii. Quadratic Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \beta_3 \times VIX + \beta_4 \times (VIX)^2 + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ poly(IV_Garch, 2) + poly(VIX, 2), data = df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.49677 -0.36602 -0.08058  0.31645  1.05999 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)   
(Intercept)          0.4495     0.1288   3.489  0.00447 **
poly(IV_Garch, 2)1   0.9159     0.9143   1.002  0.33623   
poly(IV_Garch, 2)2  -0.6337     0.7372  -0.860  0.40683   
poly(VIX, 2)1       -1.4908     1.0318  -1.445  0.17408   
poly(VIX, 2)2        0.2132     0.5612   0.380  0.71063   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5312 on 12 degrees of freedom
Multiple R-squared:  0.1652,	Adjusted R-squared:  -0.113 
F-statistic: 0.5938 on 4 and 12 DF,  p-value: 0.6738
```

The coefficients are statistically insignificant, and the F-statistic ($p = 0.6738$) shows the overall model lacks explanatory power. Furthermore, $R^2 = 0.1652$ and Adjusted $R^2 = -0.113$, confirms that this model is also not a good fit.

iv. GLM (Generalized Linear Modeling) Regression:

```
Call:
glm(formula = IV_BSM ~ IV_Garch + VIX, data = df)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.76387    0.33651   2.270   0.0395 *
IV_Garch    11.24896   20.58434   0.546   0.5933
VIX         -0.02452    0.02189  -1.120   0.2815
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.2635977)

    Null deviance: 4.0567  on 16  degrees of freedom
Residual deviance: 3.6904  on 14  degrees of freedom
AIC: 30.277

Number of Fisher Scoring iterations: 2
```

The Generalized Linear Model (GLM) regression results indicate that the model provides minimal explanatory power for the AAPL puts and VIX regression.

h. APPL PUTS + VIX+ CPI

i. Simple Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times VIX + \beta_3 \times CPI + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch + VIX + CPI, data = puts_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4012 -0.3477 -0.2729  0.2863  1.0280 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -24.32409   57.33661  -0.424    0.678
IV_Garch     10.04031   21.38493   0.470    0.646
VIX          -0.02601    0.02281  -1.141    0.275
CPI           0.07877    0.18002   0.438    0.669

Residual standard error: 0.5289 on 13 degrees of freedom
Multiple R-squared:  0.1035,	Adjusted R-squared:  -0.1034 
F-statistic: 0.5003 on 3 and 13 DF,  p-value: 0.6885
```

Here the coefficients are statistically insignificant, and the F-statistic ($p = 0.5003$) shows the overall model lacks explanatory power. Furthermore, $R^2 = 0.1035$ and Adjusted $R^2 = -0.1034$, confirming this model is also not a good fit.

ii. Logarithmic Transformation Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times log(IV_{GARCH}) + \beta_2 \times log(VIX) + \beta_3 \times log(CPI) + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ log_IV_Garch + log_VIX + log_CPI, data = puts_df)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4492 -0.3434 -0.1678  0.3106  1.0444 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -124.9588   342.4612  -0.365    0.721
log_IV_Garch   0.4243     0.6247   0.679    0.509
log_VIX       -0.8589     0.6458  -1.330    0.206
log_CPI       22.4991    59.2819   0.380    0.710

Residual standard error: 0.5204 on 13 degrees of freedom
Multiple R-squared:  0.1321,	Adjusted R-squared:  -0.06816 
F-statistic: 0.6597 on 3 and 13 DF,  p-value: 0.5913
```

The coefficients are statistically insignificant, and the F-statistic ($p = 0.5913$) shows the overall model does not do a good job at explaining IV_BSM. Furthermore, $R^2 = 0.1321$ and Adjusted $R^2 = -0.06816$, confirming this model is also not a good fit.

iii. Quadratic Linear Regression

$$IV_{BSM} = \beta_0 + \beta_1 \times IV_{GARCH} + \beta_2 \times (IV_{GARCH})^2 + \beta_3 \times VIX + \beta_4 \times (VIX)^2 + \beta_5 \times CPI + \beta_6 \times (CPI)^2 + \varepsilon$$

```
Call:
lm(formula = IV_BSM ~ IV_Garch + I(IV_Garch^2) + VIX + I(VIX^2) + 
    CPI + I(CPI^2), data = puts_df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.60680 -0.31982 -0.04813  0.19634  0.92742 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.217e+04  3.304e+04   0.974    0.353
IV_Garch      1.491e+02  3.139e+02   0.475    0.645
I(IV_Garch^2) -2.008e+03  4.657e+03  -0.431    0.675
VIX          -1.060e-01  1.655e-01  -0.641    0.536
I(VIX^2)      1.031e-03  3.055e-03   0.337    0.743
CPI          -2.019e+02  2.074e+02  -0.973    0.353
I(CPI^2)      3.168e-01  3.256e-01   0.973    0.353

Residual standard error: 0.5555 on 10 degrees of freedom
Multiple R-squared:  0.2393,	Adjusted R-squared:  -0.2171 
F-statistic: 0.5243 on 6 and 10 DF,  p-value: 0.7783
```

The coefficients are statistically insignificant, and the F-statistic ($p = 0.7783$) shows the overall model lacks explanatory power. Furthermore, $R^2 = 0.2393$ and Adjusted $R^2 = -0.2171$, confirming this model is also not a good fit.

iv. GLM (Generalized Linear Modeling) Regression

```
Call:
glm(formula = IV_BSM ~ IV_Garch + VIX + CPI, family = gaussian(link = "i
dentity"),
    data = puts_df)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -24.32409   57.33661  -0.424    0.678
IV_Garch     10.04031   21.38493   0.470    0.646
VIX          -0.02601    0.02281  -1.141    0.275
CPI           0.07877    0.18002   0.438    0.669

(Dispersion parameter for gaussian family taken to be 0.2797542)

    Null deviance: 4.0567  on 16  degrees of freedom
Residual deviance: 3.6368  on 13  degrees of freedom
AIC: 32.028

Number of Fisher Scoring iterations: 2
```

The coefficients are statistically insignificant which shows the overall model does not help explain the dependent variable.

IX. Modifications in the Trading Signals

To simulate a forward-looking trading environment, we forecast future values of IV_GARCH using a rolling GARCH(1,1)-SSTD model fitted separately for each target date. This step is crucial because it ensures that at each point in time, only information available up to that date is used, mimicking a true real-time trading situation. Using the GARCH forecasts, we then generate forecasted IV_BSM values by plugging the predicted IV_GARCH values into the previously trained regression model, polynomial regression for WMT:

$$\widehat{IV}_{BSM} = 1.1357 - 135.2223 \times IV_{GARCH} + 4291.6362 \times (IV_{GARCH})^2 ,$$

and a simple linear regression for AAPL:

$$\widehat{IV}_{BSM} = -0.3791 + 49.3817 \times IV_{GARCH}$$

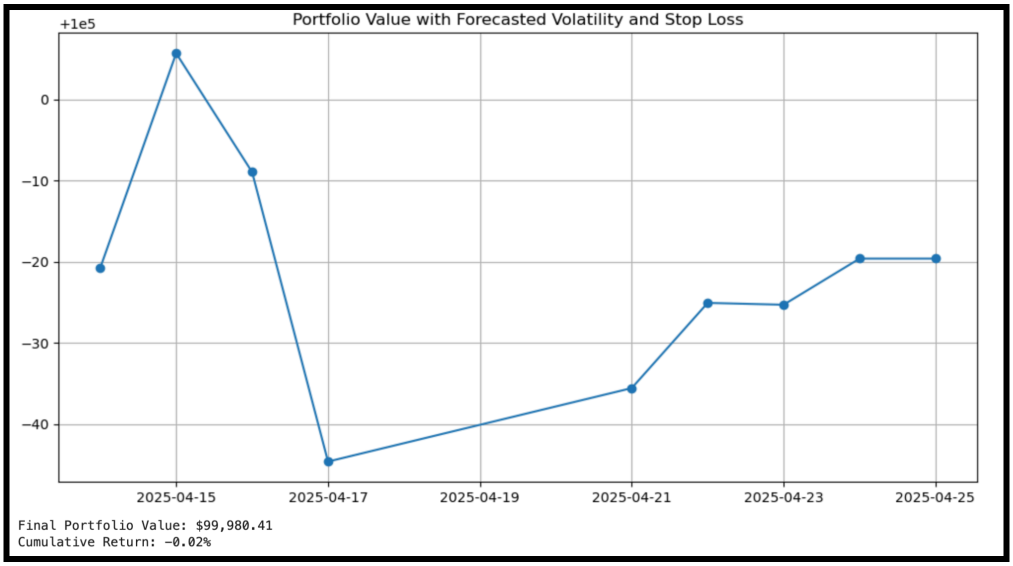
These forecasted IV_BSM values represent our model's estimate of what implied volatility "should" be, assuming the historical relationship holds.

In a parallel manner, to simulate realistic market behavior, we create synthetic realized IV_BSM paths. These are constructed by perturbing the forecasted IV_GARCH values with random skewed noise, intended to replicate real-world volatility shocks and fat-tailed behavior often observed in financial markets. Using the simulated IV_BSM and the Black-Scholes model, synthetic option prices are generated dynamically for each date, assuming fixed underlying parameters such as stock price, time to maturity, and risk-free rate.

Trading signals are then generated by comparing the simulated realized IV_BSM to the model-predicted $\widehat{IV}_{BSM}$. If the realized IV_BSM deviates upward beyond a predefined threshold from the predicted $\widehat{IV}_{BSM}$, a Sell signal is triggered; if it deviates downward beyond the threshold, a Buy signal is generated. If the deviation is within the threshold range, no trade (Hold) is executed. The threshold is dynamically scaled relative to the forecasted IV, allowing flexibility based on the current volatility regime.

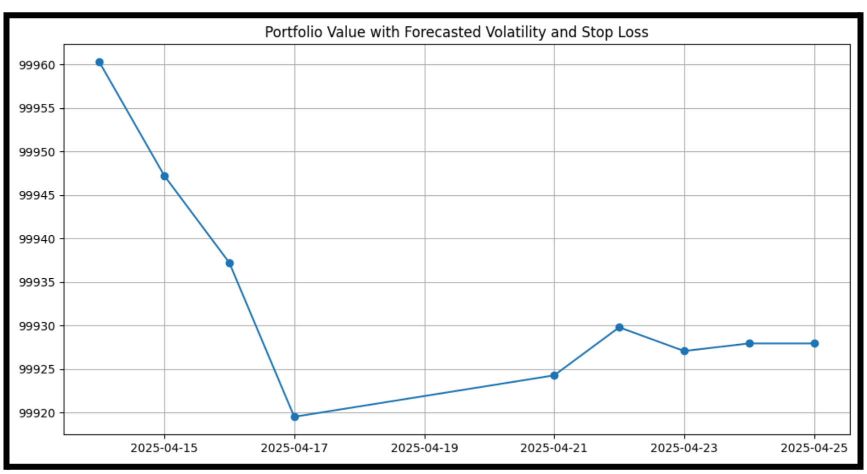
To manage risk exposure, a stop-loss mechanism is embedded into the trading logic. If a position's unrealized loss exceeds 15% of the initial option value, the position is closed immediately, capping downside risk. PnL (profit and loss) from each trade is computed based on the difference between entry and exit prices of the synthetic options, scaled appropriately by the contract multiplier.

a. Trading Strategies for WMT Put options



The plot shows the portfolio value evolution of WMT Put when trading based on the forecasted volatility and applying a stop-loss rule of 15%. Initially, the portfolio experiences a small gain, but shortly thereafter, a sharp drawdown occurs, triggering stop-loss exits. Despite partial recovery in subsequent periods, the portfolio finished slightly below the starting value, ending with a final portfolio value of $99,980.41 and a cumulative return of -0.02%. This suggests that while the strategy was able to contain losses through the stop-loss mechanism, the market conditions led to a series of unfavorable moves relative to the predicted volatility, resulting in minor underperformance.

b. Trading Strategies based off of AAPL call options



In the case of AAPL Call, the portfolio value declines steadily but less dramatically, indicating fewer extreme mispredictions between forecasted and actual implied volatilities. The final portfolio value settles at $99,927.94, corresponding to a cumulative return of -0.07%. The smoother downward trajectory suggests that trades were less volatile but more consistently unprofitable, likely due to persistent small forecast errors that did not trigger meaningful reversal opportunities.

These performances, though modest, demonstrate the potential for statistically informed volatility trading strategies to extract value from systematic inefficiencies in option markets. The reliance on a model-predicted benchmark allows the strategy to remain grounded in a consistent valuation framework, while the simplicity of the signal logic ensures interpretability and replicability. However, it is important to note that the backtest does not account for transaction costs, slippage, or changes in Vega exposure over time. Future iterations may benefit from incorporating additional macroeconomic variables, using out-of-sample testing, or transitioning to a delta-hedged setup to isolate pure volatility exposure more effectively.

X. Conclusion

This project set out to explore whether systematic mispricings in implied volatility could be identified and exploited using a statistically grounded framework. Initially, we modeled the relationship between backward-looking GARCH-implied volatility and forward-looking BSM-implied volatility using polynomial regression. We then extended the methodology to a dynamic trading environment by forecasting future GARCH volatilities, estimating corresponding *fair value* implied volatilities, and generating trading signals based on observed deviations.

The final trading strategies leveraged both synthetic IV paths and a real-time forecasting structure, incorporating stop-loss risk management to protect against large unfavorable moves. Despite some minor negative returns in the backtests, the results show that the strategy was able to maintain portfolio stability even in adverse market conditions. Losses were modest, and the behavior of the portfolio was consistent with a volatility-driven trading system focused on relative mispricings rather than directional market bets.

Importantly, the framework developed here lays a foundation for more advanced extensions, including dynamic threshold optimization, portfolio-wide risk budgeting, and combining volatility signals with macroeconomic information. In real-world deployment, improvements such as incorporating liquidity constraints, transaction costs, and more sophisticated volatility modeling techniques would be necessary. Nonetheless, this study demonstrates that statistically significant volatility signals, when properly forecasted and risk-managed, offer a viable path toward building an algorithmic options trading system grounded in volatility dynamics.