An Analysis of Market Factors and Investor Sentiment as Determinants of the CBOE SKEW Index

Adhiraj Chhoda

Thomas Jefferson High School for Science and Technology

Research Statistics 3, Period 1

Dr. Catherine Scott

05/30/25

Introduction and Rationale

The mechanisms through which financial markets incorporate diverse information and price risk are central to financial economic theory and practice. Beyond broadly cited indicators of expected market volatility, such as the CBOE Volatility Index (VIX), the derivative markets, particularly options markets, provide nuanced perspectives on investor expectations concerning the likelihood and potential magnitude of extreme price fluctuations, commonly termed tail risk. A well-established empirical characteristic of equity index options, most notably for the S&P 500 (SPX), is the "volatility smirk." This term describes the observed asymmetry in the implied volatilities of options across various strike prices: specifically, out-of-the-money (OTM) put options systematically demonstrate higher implied volatilities when compared to at-the-money (ATM) or OTM call options. The Chicago Board Options Exchange (CBOE) SKEW Index is a standardized, market-disseminated metric explicitly designed to quantify this particular asymmetry, serving as an indicator of perceived tail risk within the U.S. equity market (CBOE, 2011). Higher levels of the SKEW Index are generally interpreted as signifying heightened market apprehension regarding potential significant downward price movements.

A comprehensive investigation into the factors that drive the SKEW Index is important for a more complete understanding of market dynamics, risk pricing, and the information environment. Fluctuations in the SKEW Index can reflect collective shifts in risk aversion, changes in the market's probabilistic assessment of infrequent but high-impact events, or the footprint of large-scale institutional hedging programs. For instance, Xing, Zhang, and Zhao (2010) explored the informational content embedded within the volatility smirk of individual equity options, discerning that its steepness held predictive capacity for subsequent stock returns. This implies that options markets incorporate forward-looking assessments of downside risk that may not be instantaneously reflected in the prices of the underlying equities. The dynamic nature of such option-implied measures was further elucidated by Onan, Salih, and Yasar (2014), who investigated the effects of U.S. macroeconomic announcements on the SPX implied volatility slope and the VIX. Their findings indicated that the public disclosure of significant economic information, and the attendant resolution of market uncertainty, precipitated discernible changes in both metrics.

Moreover, behavioral finance offers compelling perspectives on the influence of investor sentiment on asset valuation. Baker and Wurgler (2006) established that pervasive shifts in investor sentiment tend to have a more substantial effect on securities characterized by subjective valuations and higher costs to arbitrage. Specifically within the context of options pricing, Han (2008) connected investor sentiment directly to option valuations, showing that prevailing sentiment levels could influence the relative pricing of call and put options, thereby modifying the shape of the implied volatility smirk. These collective works suggest that the SKEW Index is likely influenced by a complex interplay of structural market factors, dynamic information flows, and prevailing investor psychology. The analysis endeavors to integrate these diverse elements by developing an empirical framework to model the daily determinants of the CBOE SKEW Index for the period from February 23, 2011, to October 4, 2019. This timeframe, commencing with the widespread availability of the official SKEW Index data and concluding with the available historical data for key predictor variables, facilitates a contemporary analysis of this important market indicator. The central research question addressed is: What are the principal market-based and behavioral factors that explain the daily variations in the CBOE SKEW Index, and what are the statistical significance and quantitative impacts of aggregate market volatility, recent market performance trends, options market trading activity (specifically put-call ratios), and broader investor sentiment on this established measure of equity market tail risk during the specified period?

To establish a rigorous theoretical and empirical basis for this investigation, we'll discuss and summarize three impactful peer-reviewed studies. These works collectively examine the informational value of the volatility smirk, its responsiveness to significant market events, and the broader role of investor sentiment in financial markets.

The first study, by Xing, Y., Zhang, X., & Zhao, R. (2010), titled "What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns?" and published in the Journal of Financial and Quantitative Analysis, provided empirical evidence on the cross-sectional predictive power of the volatility smirk observed in individual stock options. Using U.S. equity options data from OptionMetrics spanning 1996 to 2005, along with CRSP and Compustat data for underlying stocks, the authors defined the smirk as the difference in implied volatilities between OTM puts and ATM calls. Their methodology involved portfolio sorts and Fama-MacBeth (1973) regressions, controlling for known determinants of stock returns such as size, book-to-market ratio, and momentum. The research concluded that stocks with steeper volatility smirks—indicative of a higher relative cost for OTM puts—subsequently underperformed stocks with flatter smirks, even after risk adjustment. This finding suggests that the options market incorporates information, possibly from informed traders, concerning future downside risk, which is not immediately captured by equity prices.

The second study, authored by Onan, M., Salih, A., & Yasar, B. (2014) and published in Finance Research Letters under the title "Impact of Macroeconomic Announcements on Implied Volatility Slope of SPX Options and VIX," employed an event-study methodology using high-frequency intraday data. Focusing on SPX options and VIX index data from 2007 to 2010, the authors investigated how scheduled U.S. macroeconomic news releases influenced the SPX implied volatility slope and the VIX index itself. The dependent variables in their regression models were changes in these volatility metrics around announcement times, with announcement indicators and surprise components as independent variables. Their results showed that macroeconomic announcements generally led to a reduction in uncertainty, manifesting as significant decreases in both the VIX and the steepness of the volatility skew. This highlights the sensitivity of the SKEW Index's underlying constructs to significant market-wide information dissemination and resulting shifts in uncertainty.

The third relevant piece of research is by Baker, M., & Wurgler, J. (2006), "Investor sentiment and the cross‐section of stock returns," published in The Journal of Finance. This paper investigated how broad, economy-wide waves of investor sentiment influence the returns of different categories of U.S. stocks for the period 1962 to 2001. They constructed a composite investor sentiment index from six proxies. Analyzing returns of stock portfolios sorted by characteristics like size, age, and profitability conditional on prevailing sentiment levels, they found that sentiment exerts a more pronounced influence on stocks that are more difficult to value and arbitrage. Specifically, such "speculative" stocks tended to experience lower subsequent returns following periods of high sentiment and vice-versa. This study provides a robust rationale for examining investor sentiment as a potential driver of the SKEW Index, given that the pricing of tail risk inherently involves subjective assessments that can be influenced by prevailing market psychology.

Design and Methodology

The research outlined herein will employ a quantitative, observational design based on historical daily time-series data from February 23, 2011, through October 4, 2019. This period was chosen to ensure consistent availability of the official CBOE SKEW Index data and comprehensive historical data for all predictor variables. The primary analytical methodology is multiple linear regression (MLR), aimed at identifying and quantifying the daily determinants of the CBOE SKEW Index.

The data for the CBOE SKEW Index (dependent variable) and the CBOE Volatility Index (VIX) were obtained as daily closing values from historical records of the Chicago Board Options Exchange. Daily closing prices of the S&P 500 index (SPX) were utilized to derive two key market measures: S&P 500 realized volatility (RealizedVol) and recent S&P 500 market returns (MarketReturn). Specifically, RealizedVol was calculated as the annualized standard deviation of daily logarithmic S&P 500 returns over a trailing 21-trading day window. MarketReturn was calculated as the cumulative logarithmic S&P 500 return over the preceding 21 trading days. Investor sentiment (Sentiment) was proxied by the weekly Bullish Sentiment Percentage from the American Association of Individual Investors (AAII) Sentiment Survey; this weekly data was converted to a daily series by applying the most recent weekly observation to each subsequent trading day using a forward-fill method. Daily data for the S&P 500 Index Put-Call Ratio (PC_Ratio) were obtained from CBOE historical statistics.

From these primary data series, several derived variables were constructed for model estimation. A VIX_sq variable was computed by squaring the daily VIX values to capture potential non-linear effects of market volatility on the SKEW Index. Crucially, to investigate whether the impact of VIX on the SKEW Index is conditional upon investor sentiment, and to mitigate potential multicollinearity arising from standard interaction terms, a centered interaction term was created. This involved first mean-centering both the VIX and Sentiment series (by subtracting their respective sample means) and then multiplying these centered series to yield VIX_Sent_Interact_centered. All predictor variables were subsequently lagged by one trading day (t-1) to predict the SKEW Index on day t ($SKEW_t$), ensuring that only past information is used in the predictive models.

Two main regression models were specified and estimated:

Model 1 (Levels Model):

$$SKEW_t = \beta_0 + \beta_1 VIX_{t-1} + \beta_2 VIX^2_{t-1} + \beta_3 RealizedVol_{t-1} + \beta_4 MarketReturn_{t-1} + \beta_5 Sentiment_{t-1} + \beta_6 PutCallRatio_{t-1} + \beta_7 (VIX \times Sentiment)_{t-1} + \varepsilon_t$$

In this equation, $SKEW_t$ is the value of the SKEW Index on day *t*. β_0 represents the intercept, β_1 through β_7 are the slope coefficients corresponding to each lagged independent variable, and ε_t signifies the error term.

Model 2 (Log-Log Model):

To address potential non-linearities inherent in financial data and improve the distributional properties of the variables and model residuals, a second model was specified using natural logarithmic transformations for the SKEW Index and most predictor variables (excluding $MarketReturn_{t-1}$ which can take negative values, and VIX_sq_{t-1} as its effect is explored through $\log(VIX_{t-1})$). A new centered interaction term based on the logs of centered VIX and centered Sentiment ($logVIX\_logSent\_Interact\_centered_{t-1}$) was also created for this model:

$$\log(SKEW_t) = \gamma_0 + \gamma_1\log(VIX_{t-1}) + \gamma_2\log(RealizedVol_{t-1}) + \gamma_3 MarketReturn_{t-1} + \gamma_4\log(Sentiment_{t-1}) + \gamma_5\log(PutCallRatio_{t-1}) + \gamma_6(\log VIX \times \log Sentiment)_{t-1} + \upsilon_t$$

Here, γ_0 is the intercept, γ_1 through γ_6 are the slope coefficients for the transformed (or original, in the case of market return) lagged predictors, and υ_t is the error term.

The analytical process involved comprehensive data loading, standardizing dates, alignment, and merging of all sourced time series, followed by the calculation of the derived and lagged variables as described. An Exploratory Data Analysis was conducted, featuring descriptive statistics and visualizations (histograms, boxplots, correlation matrices) to understand the data properties prior to regression analysis.

The core statistical investigation involved estimating these MLR models using Ordinary Least Squares. The statistical significance of individual coefficients was assessed via t-tests, with corresponding p-values. Overall model significance was evaluated using the F-statistic. Model explanatory power was gauged by R-squared and Root Mean Squared Error. Rigorous diagnostic checking was integral to the methodology to validate OLS assumptions. This included assessing linearity (Residuals vs. Fitted plots, Component+Residual plots), normality of residuals (Q-Q plots, Histograms of residuals, Shapiro-Wilk tests), homoscedasticity (Scale-Location plots, Breusch-Pagan tests), and independence of residuals (Durbin-Watson tests, ACF and PACF plots, Ljung-Box tests). Potential multicollinearity was diagnosed using Variance Inflation Factors (VIFs), and influential data points were identified via Cook's Distance. Given the potential presence of heteroscedasticity and autocorrelation in daily financial time-series data, Newey–West Heteroscedasticity and Autocorrelation Consistent standard errors were applied to ensure more robust inference on model coefficients. All statistical analyses were conducted using the R programming language in RStudio, with the complete workflow and outputs fully documented in R Markdown.

Data Analysis and Results

The study utilized daily time-series data for the CBOE SKEW Index and a set of theoretically motivated predictor variables spanning from February 23, 2011, to October 4, 2019. After comprehensive data preparation, cleaning, and alignment, which included deriving volatility and return measures from S&P 500 daily prices and processing weekly AAII sentiment data to a daily frequency using forward-fill, the final analytical dataset comprised 2,165 daily observations. All predictor variables were lagged by one day to ensure that only past information was used to explain current SKEW Index levels.

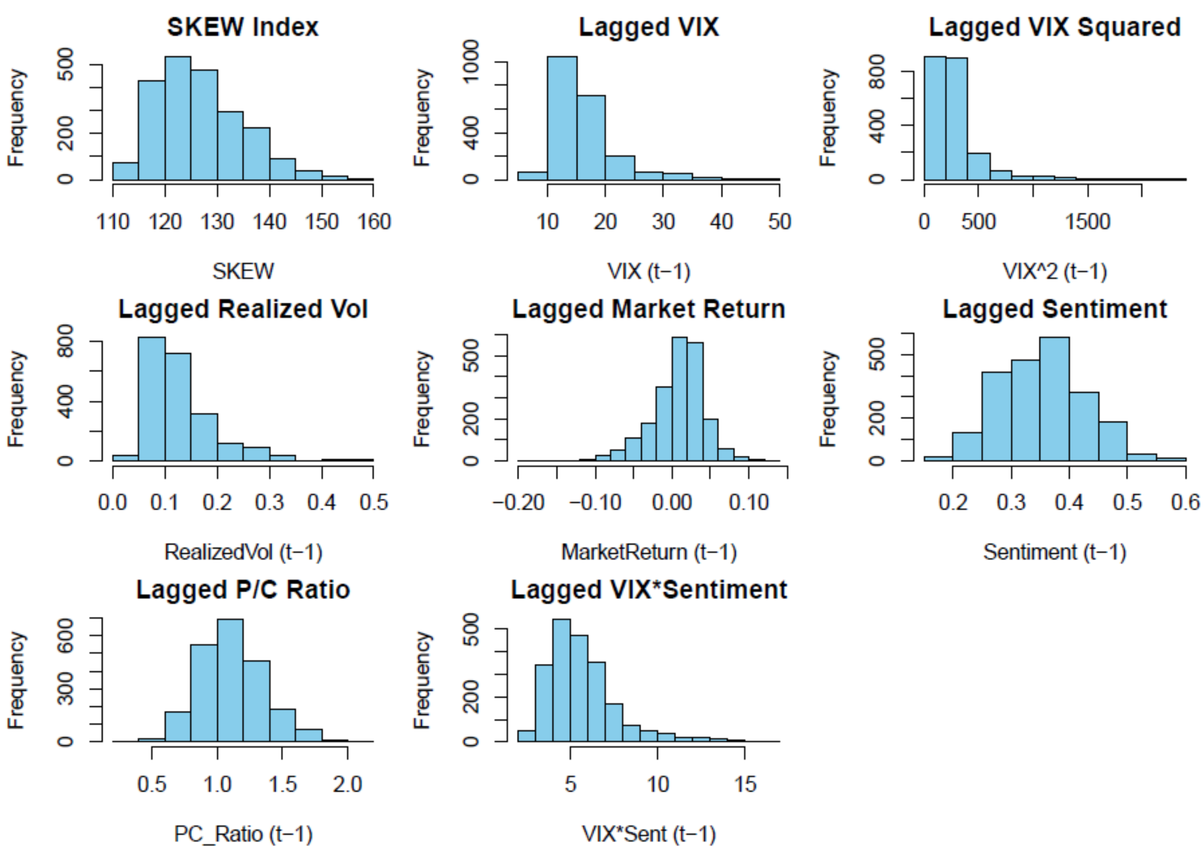
Exploratory Data Analysis (EDA)

Prior to formal model estimation, an exploratory data analysis was conducted to understand the distributional properties and interrelationships of the variables

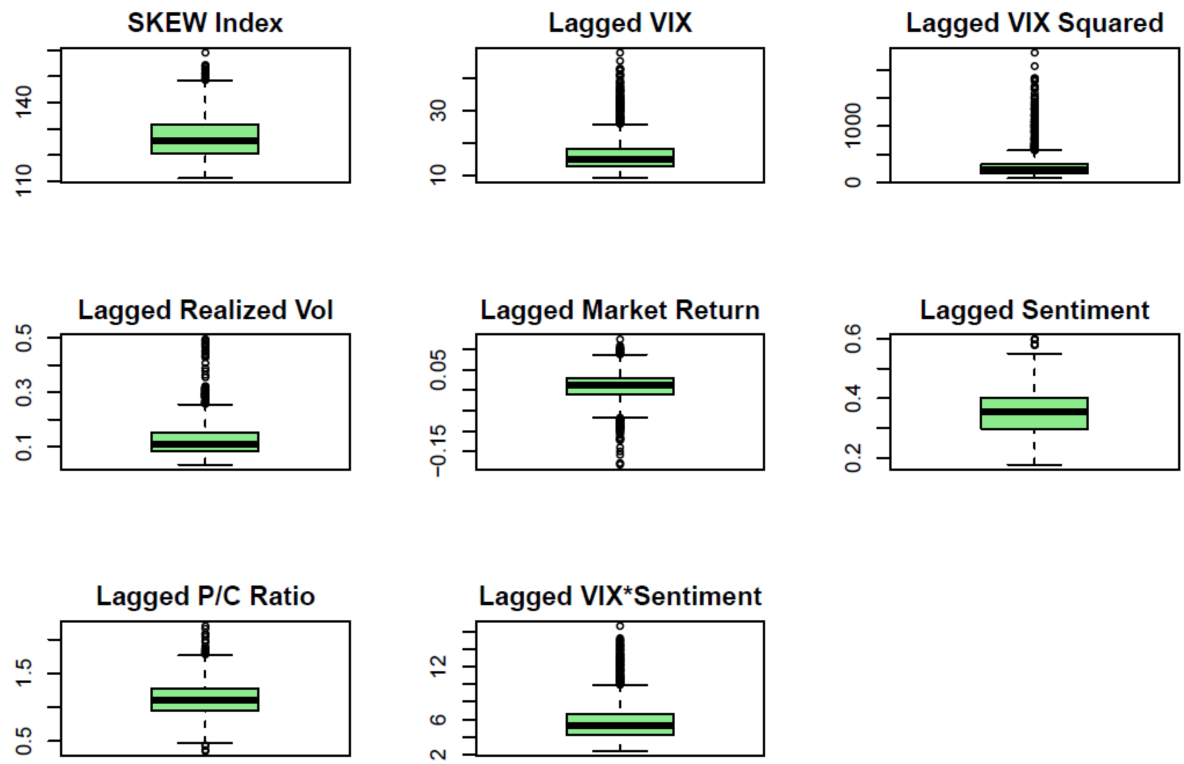
```
##       SKEW          VIX_lag1       VIX_sq_lag1      RealizedVol_lag1
##  Min.   :111.3   Min.   : 9.14   Min.   :  83.54   Min.   :0.03469
##  1st Qu.:120.5   1st Qu.:12.86   1st Qu.: 165.38   1st Qu.:0.08420
##  Median :125.4   Median :14.88   Median : 221.41   Median :0.11072
##  Mean   :126.7   Mean   :16.29   Mean   : 294.38   Mean   :0.13011
##  3rd Qu.:131.7   3rd Qu.:18.11   3rd Qu.: 327.97   3rd Qu.:0.15390
##  Max.   :159.0   Max.   :48.00   Max.   :2304.00   Max.   :0.49411
##  MarketReturn_lag1   Sentiment_lag1   PC_Ratio_lag1
##  Min.   :-0.182655   Min.   :0.1775   Min.   :0.350
##  1st Qu.:-0.008998   1st Qu.:0.2975   1st Qu.:0.950
##  Median : 0.012476   Median :0.3534   Median :1.100

##  Mean   : 0.007877   Mean   :0.3531   Mean   :1.122
##  3rd Qu.: 0.029850   3rd Qu.:0.4014   3rd Qu.:1.280
##  Max.   : 0.125231   Max.   :0.5975   Max.   :2.200
##  VIX_Sentiment_Interaction_lag1 VIX_Sent_Interact_centered_lag1
##  Min.   : 2.329                 Min.   :-2.583650
##  1st Qu.: 4.267                 1st Qu.:-0.153735
##  Median : 5.287                 Median :-0.005704
##  Mean   : 5.699                 Mean   :-0.051596
##  3rd Qu.: 6.549                 3rd Qu.: 0.102383
##  Max.   :16.709                 Max.   : 1.990596
```

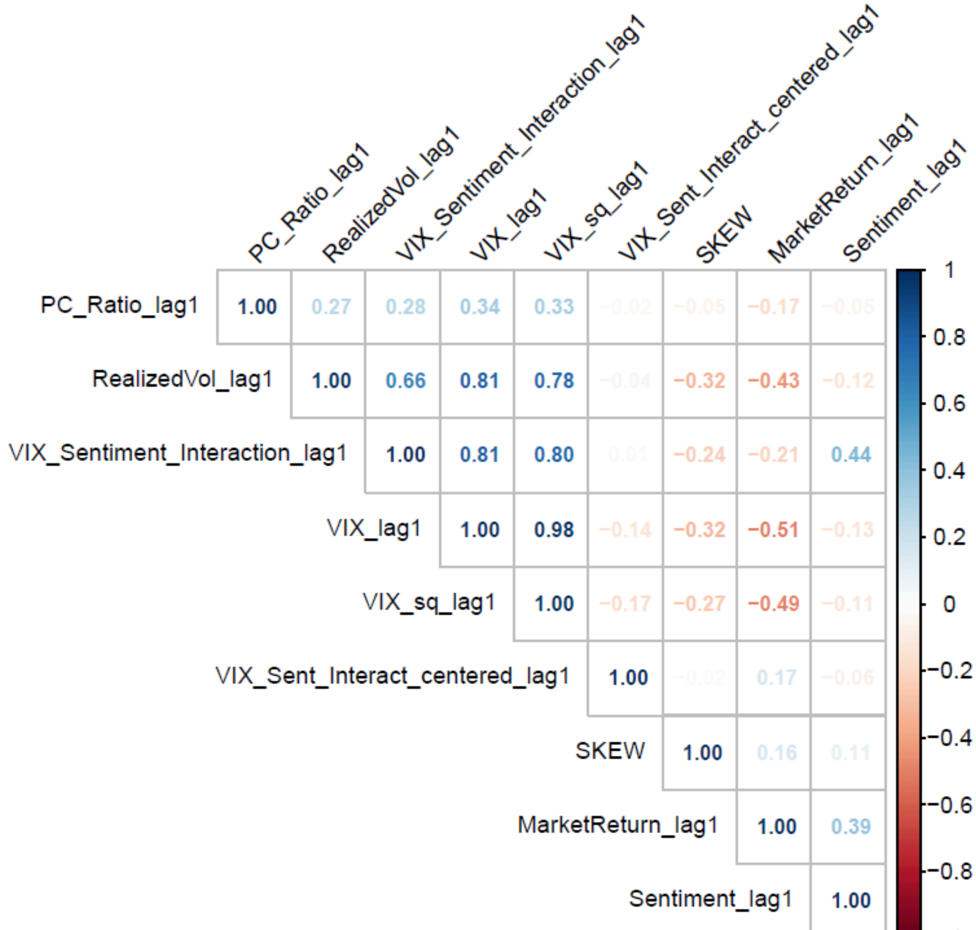
Descriptive statistics, presented in the R Markdown output, revealed the ranges and central tendencies of the SKEW Index and its potential drivers. The SKEW Index during the sample period ranged from a minimum of 111.3 to a maximum of 159.0, with a mean of 126.7 and a median of 125.4. The VIX (lagged) ranged from 9.14 to 48.00. Visual examination of variable distributions was conducted through histograms.



The histograms indicated that the SKEW Index distribution was somewhat right-skewed. VIX_lag1 and VIX_sq_lag1 also exhibited right-skewness, which is typical for volatility measures. MarketReturn_lag1 was more symmetrically distributed around zero. AAII Sentiment was bounded, while the Put-Call Ratio displayed some right-skewness. Boxplots provided an alternative visualization of these distributions and potential outliers.



Pairwise correlations among the variables were examined using a correlation matrix.



The correlation matrix highlighted several notable relationships. As expected, VIX_lag1 and VIX_sq_lag1 were highly correlated (0.98). VIX_lag1 also showed moderate positive correlations with its uncentered interaction with sentiment (VIX_Sentiment_Interaction_lag1, 0.81) and RealizedVol_lag1 (0.81). The SKEW Index itself showed weak negative correlations with VIX_lag1 (-0.14) and RealizedVol_lag1 (-0.32), and a moderate positive correlation with Sentiment_lag1 (0.16 in the initial uncentered interaction model perspective, but a direct correlation with SKEW shows -0.02 for VIX_Sent_Interact_centered_lag1). These initial observations suggested the need for careful model specification and diagnostics for multicollinearity.

Regression Model Estimation and Diagnostics

Two primary multiple linear regression models were estimated: Model 1, using variables in their original levels, and Model 2, employing logarithmic transformations for SKEW and most predictors (excluding market return due to its potential for negative values).

Model 1: Levels Specification

The initial OLS estimation of the levels model (Model 1a) is presented below.

	Dependent variable:
	CBOE SKEW Index (Daily)
VIX (t-1)	-1.641^{***}
	(-1.931,-1.350)
VIX^2 (t-1)	0.030^{***}
	(0.024,0.036)
Realized Vol (t-1, 21d)	-19.189^{***}
	(-27.170,-11.208)
Market Return (t-1, 21d)	-11.443
	(-22.899,0.013)
AAII Sentiment (t-1, Bullish %)	6.858^{**}
	(2.169,11.546)
SPX Put-Call Ratio (t-1)	2.184^{**}
	(0.862,3.505)
$\text{VIX}_c\ (t-1) \times \text{Sent}_c\ (t-1)$	-0.125
	(-1.022,0.772)
Constant	142.330^{***}
	(138.898,145.762)

Observations	2,165
R-squared	0.160
Adj. R-squared	0.158
F-statistic	58.81^{***}
Observations	2,165
Adjusted R^2	0.158

$^{*}p<.05$; $^{**}p<.01$; $^{***}p<.001$. OLS Standard Errors.

The OLS Model 1a yielded a R-squared of 0.160, suggesting that approximately 16.0% of the daily variation in the SKEW Index could be explained by the predictors in levels. The overall F-statistic was 58.81 ($p < 0.001$), indicating the model as a whole was statistically significant. Significant predictors at the 5% level included VIX_lag1 (negative), VIX_sq_lag1 (positive), RealizedVol_lag1 (negative), Sentiment_lag1 (positive), and PC_Ratio_lag1 (positive).

Diagnostic tests were performed on Model 1a's residuals:

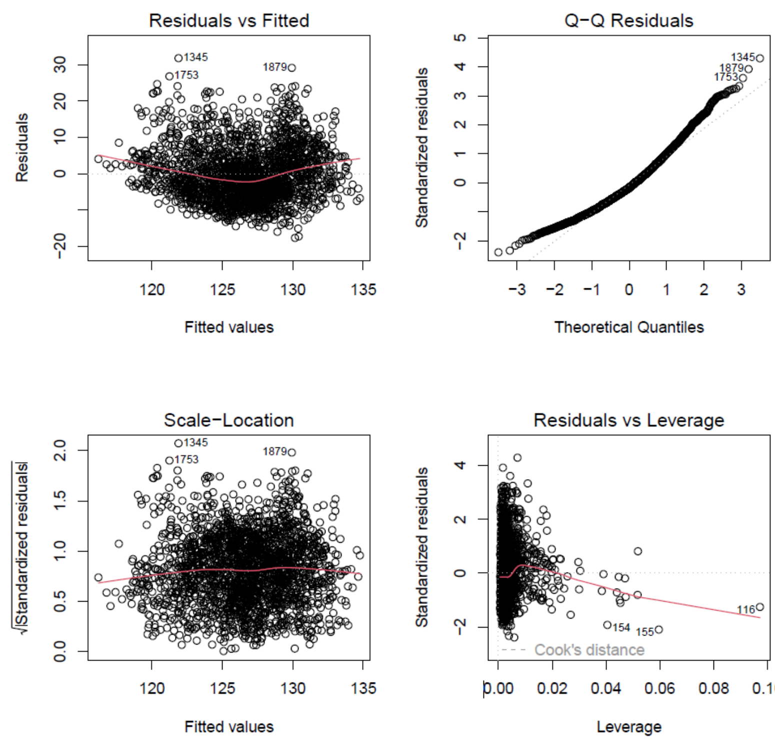


Figure 1: Regression Diagnostic Plots for OLS Model

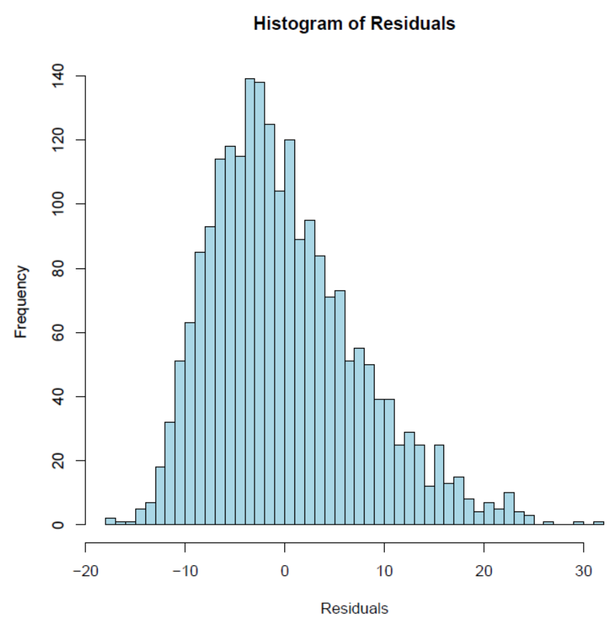
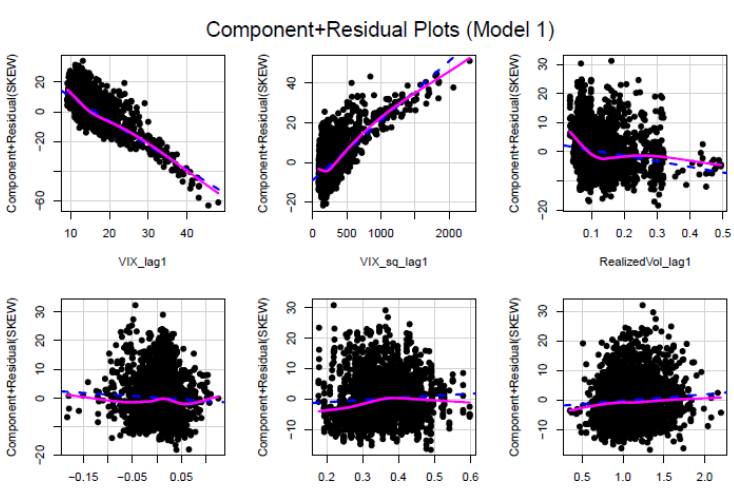
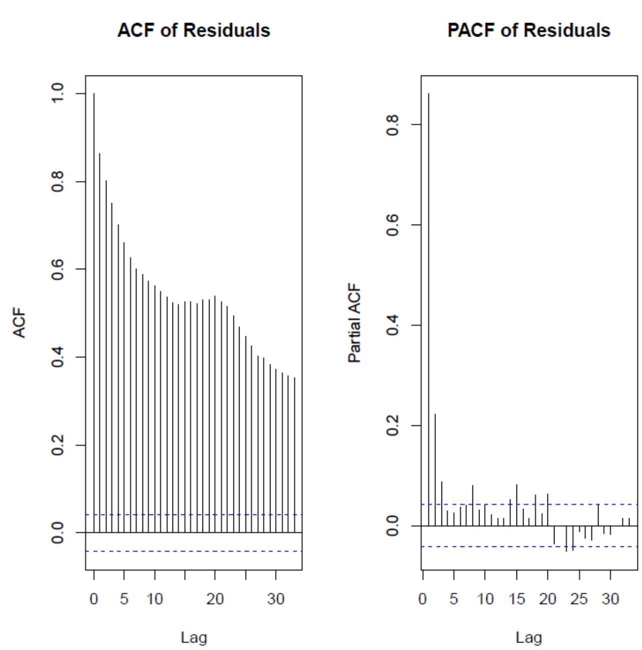


Figure 3: Regression Diagnostic Plots for OLS Model





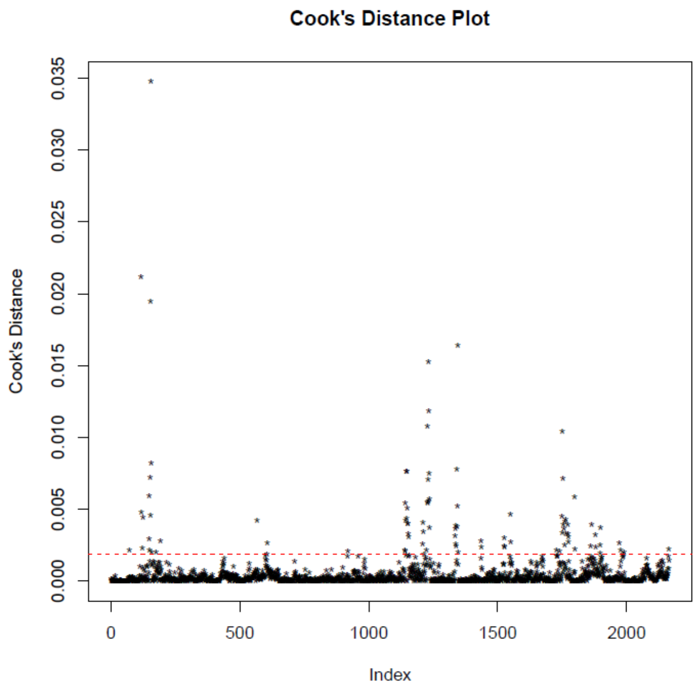


Figure 5: Regression Diagnostic Plots for OLS Model

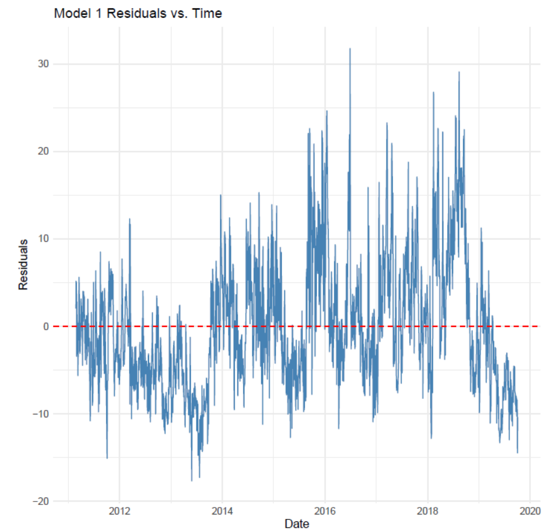


Figure 6: Regression Diagnostic Plots for OLS Model

The diagnostic plots (Residuals vs. Fitted, Scale-Location) indicated some heteroscedasticity. The Q-Q plot and histogram of residuals showed deviations from normality, confirmed by a significant Shapiro-Wilk test ($W = 0.965$, $p < 0.001$). The Breusch-Pagan test for heteroscedasticity was significant ($BP = 58.866$, $p < 0.001$). Most critically, the Durbin-Watson test ($DW = 0.274$) and the ACF/PACF plots indicated strong positive serial correlation in the residuals. VIF values were high for VIX_lag1 (25.06) and VIX_sq_lag1 (21.87), primarily due to their inherent relationship, though the centered interaction term had a low VIF (1.11). The Cook's distance plot identified 94 influential points (using the 4/n threshold), but none exceeded a Cook's D of 1.

Given the pronounced autocorrelation and heteroscedasticity, Newey-West HAC standard errors were computed for Model 1b.

	Dependent variable:
	CBOE SKEW Index (Daily)
VIX (t-1)	-1.641^{***}
	(0.308)
VIX^2 (t-1)	0.030^{***}
	(0.007)
Realized Vol (t-1, 21d)	-19.189^{*}
	(9.353)
Market Return (t-1, 21d)	-11.443
	(11.445)
AAII Sentiment (t-1, Bullish %)	6.858
	(5.337)
SPX Put-Call Ratio (t-1)	2.184^{*}
	(0.980)
$\text{VIX}_c\ (t-1) \times \text{Sent}_c\ (t-1)$	-0.125
	(1.034)
Constant	142.330^{***}
	(3.344)
Observations	2,165
R-squared (OLS)	0.160
Adj. R-squared (OLS)	0.158
Newey-West Lag Chosen	7
Durbin-Watson Stat. (OLS)	0.27
Observations	2,165

*p<.05; **p<.01; ***p<.001. Newey-West HAC Standard Errors.

With HAC standard errors (Newey-West lag of 7), VIX_lag1 (coef. -1.641, p < 0.001), VIX_sq_lag1 (coef. 0.030, p < 0.001), and RealizedVol_lag1 (coef. -19.189, p = 0.038) remained significant. PC_Ratio_lag1 (coef. 2.184, p = 0.026) also remained significant. However, MarketReturn_lag1 became insignificant (p = 0.075 from p = 0.050), and Sentiment_lag1 became insignificant (coef. 6.858, p = 0.198 from p = 0.004). The centered interaction term remained insignificant.

Model 2: Log-Log Specification

To address potential non-linearities and improve residual properties, a log-log model (Model 2a) was estimated.

	Dependent variable:
	log(CBOE SKEW Index)
log(VIX (t-1))	-0.043^{***}
	(-0.058,-0.028)
log(RealizedVol (t-1))	-0.038^{***}
	(-0.046,-0.029)
Market Return (t-1)	-0.093^{*}
	(-0.181,-0.006)
log(Sentiment (t-1))	0.020^{**}
	(0.008,0.032)
log(P/C Ratio (t-1))	0.020^{***}
	(0.008,0.031)
$\log(\text{VIX}_c\ (t-1)) \times \log(\text{Sent}_c\ (t-1))$	-0.044^{*}
	(-0.086,-0.002)
Constant	4.897^{***}
	(4.837,4.957)
Observations	2,165
R-squared	0.179
Adj. R-squared	0.177
F-statistic	$78.51^{***} (df = 6, 2158)$
Observations	2,165
Adjusted R^2	0.177

*p<.05; **p<.01; ***p<.001. OLS Standard Errors.

Model 2a yielded a R-squared of 0.179, slightly higher than the levels model. The F-statistic was 78.51 (p < 0.001). In this OLS specification, log(VIX_lag1) (negative), log(RealizedVol_lag1) (negative), MarketReturn_lag1 (negative), log(Sentiment_lag1) (positive), log(PC_Ratio_lag1) (positive), and the log-based interaction term (negative) were all statistically significant. Diagnostics were then performed for Model 2a.

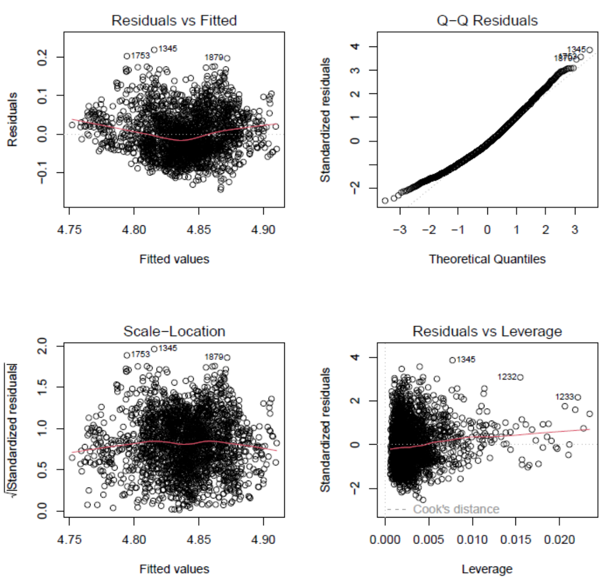


Figure 7: Regression Diagnostic Plots for Log Transformed Model

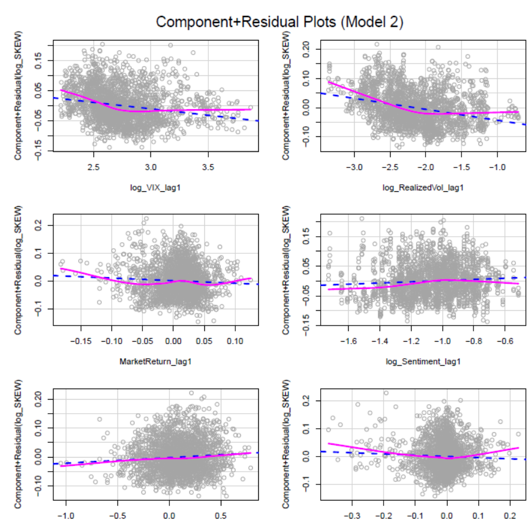


Figure 8: Regression Diagnostic Plots for Log Transformed Model

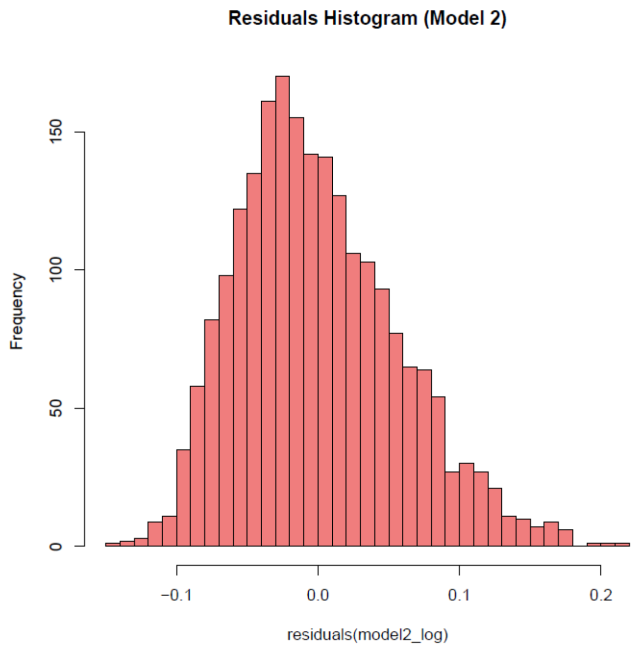


Figure 9: Regression Diagnostic Plots for Log Transformed Model

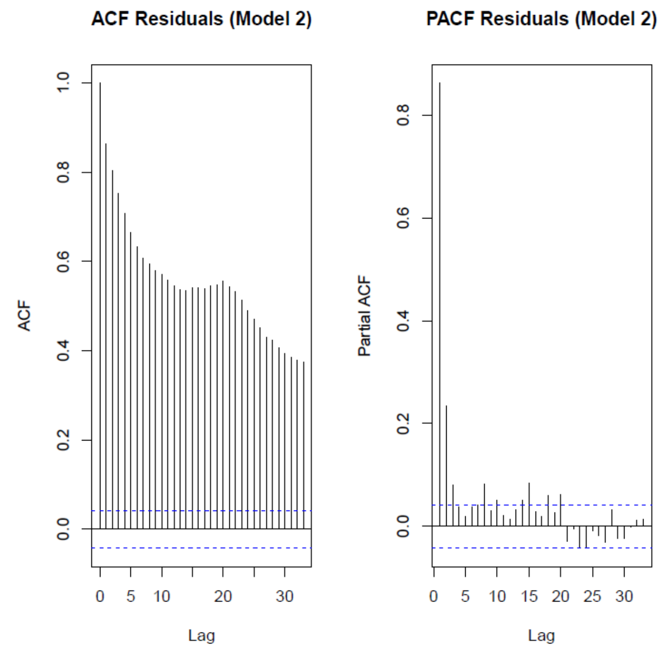


Figure 10: Regression Diagnostic Plots for Log Transformed Model

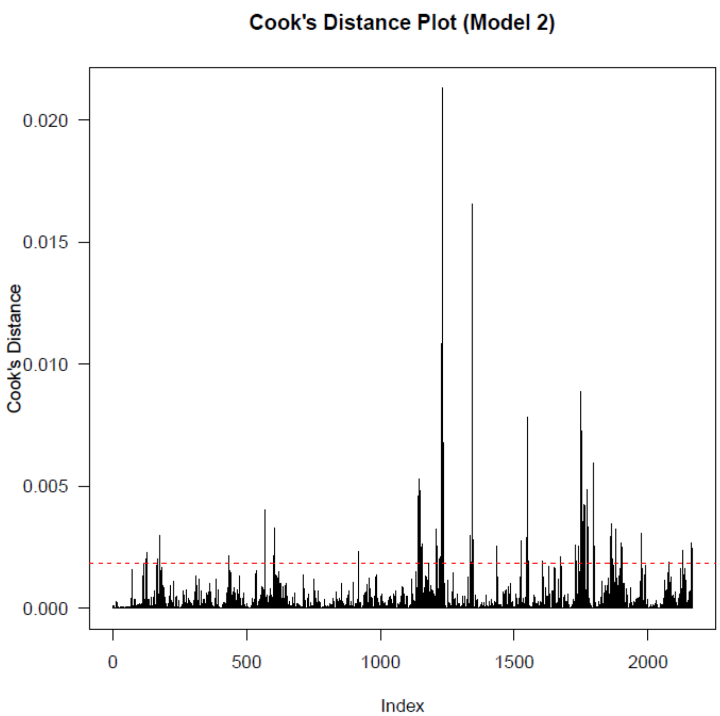


Figure 11: Regression Diagnostic Plots for Log Transformed Model

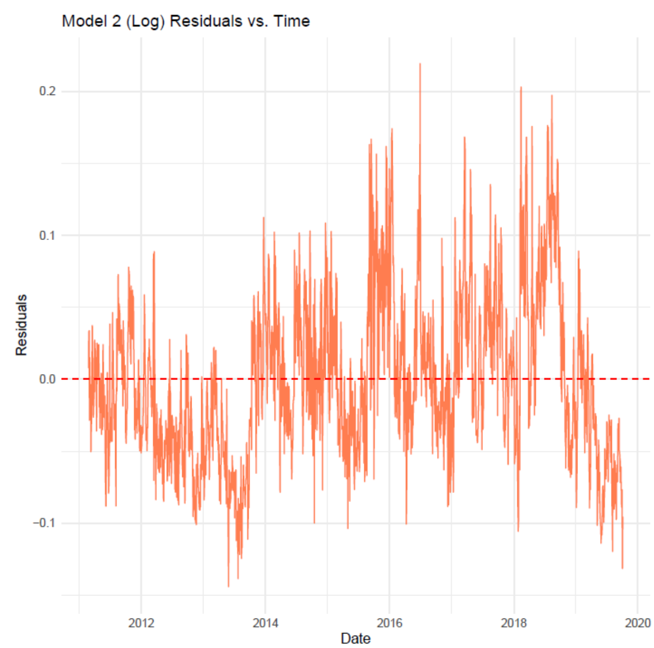


Figure 12: Regression Diagnostic Plots for Log Transformed Model

The log transformation appeared to improve the normality of residuals (Shapiro-Wilk W = 0.980, $p < 0.001$, still significant but closer to normality) and mitigate some heteroscedasticity, although the Breusch-Pagan test remained significant (BP = 35.793, p < 0.001). Strong serial correlation persisted (DW = 0.274). VIFs for Model 2 were all low (largest was 3.22 for log(VIX_lag1)), indicating no problematic multicollinearity. Cook's distance identified 100 influential points (4/n threshold), none exceeding 1.

Given these diagnostics, Newey-West HAC standard errors were computed for Model 2b.

```
Estimate Std. Error t value  Pr(>|t|)
```

(Intercept) 4.8966239 0.0672105 72.8551 < 2.2e-16 log_VIX_lag1 -0.0429481 0.0157489 -2.7271 0.0064419 log_RealizedVol_lag1 -0.0375033 0.0101894 -3.6806 0.0002384 MarketReturn_lag1 -0.0934942 0.0941889 -0.9926 0.3210044 log_Sentiment_lag1 0.0199118 0.0143564 1.3870 0.1655955 log_PC_Ratio_lag1 0.0195837 0.0080337 2.4377 0.0148620 logVIX_logSent_Interact_centered_lag1 -0.0436371 0.0496226 -0.8794 0.3792928

(Intercept) ***log_VIX_lag1*** *log_RealizedVol_lag1* ** MarketReturn_lag1
log_Sentiment_lag1
log_PC_Ratio_lag1 *
logVIX_logSent_Interact_centered_lag1
— Signif. codes: 0 '' ***0.001*** '' *0.01* '' 0.05 '.' 0.1 ' ' 1

	log(CBOE SKEW Index)
log(VIX (t-1))	-0.043^{***}
	(-0.058,-0.028)
log(RealizedVol (t-1))	-0.038^{***}
	(-0.046,-0.029)
Market Return (t-1)	-0.093^{*}
	(-0.181,-0.006)
log(Sentiment (t-1))	0.020^{**}
	(0.008,0.032)
log(P/C Ratio (t-1))	0.020^{***}
	(0.008,0.031)
$\log(\text{VIX}_c\ (t-1)) \times \log(\text{Sent}_c\ (t-1))$	-0.044^{*}
	(-0.086,-0.002)
Constant	4.897^{***}
	(4.837,4.957)
Observations	2,165
R-squared	0.179
Adj. R-squared	0.177
F-statistic	$78.51 *** (df = 6, 2158)$
Observations	2,165
Adjusted R^2	0.177
Note:	*p<.05; **p<.01; ***p<.001. Newey-West HAC Standard Errors.

Using HAC standard errors (Newey-West lag of 7) for Model 2b, the following key results emerged:

- log(VIX_lag1) had a coefficient of -0.043 ($p = 0.006$), indicating that a 1% increase in lagged VIX is associated with a 0.043% decrease in SKEW.
- log(RealizedVol_lag1) had a coefficient of -0.038 ($p < 0.001$), suggesting a 1% increase in lagged realized volatility corresponds to a 0.038% decrease in SKEW.
- log(PC_Ratio_lag1) had a coefficient of 0.020 ($p = 0.015$), implying a 1% increase in the lagged put-call ratio leads to a 0.020% increase in SKEW.
- MarketReturn_lag1 ($p = 0.321$), log(Sentiment_lag1) ($p = 0.166$), and the log-based interaction term ($p = 0.379$) were not statistically significant in the HAC-corrected model.

To further assess the out-of-sample predictive performance and generalizability of both the levels model (Model 1) and the log-log model (Model 2), a k-fold cross-validation procedure suitable for time series data was implemented. A rolling origin approach with k=5 folds was utilized. For each fold, the respective model was trained on an incrementally expanding portion of the historical data (starting with an initial 70% training window) and subsequently tested on a contiguous block of unseen future data. The Root Mean Squared Error and an out-of-sample R-squared metric were calculated for each fold's test set predictions to evaluate predictive accuracy.

```
--- Model 1 (Levels): Rolling Origin k-Fold Cross-Validation ---
Fold 1 RMSE (Model 1, levels): 10.4261 | OOS R-squared (Model 1, levels): -1.6195 | Train Size: 1515 | Test Size: 130
Fold 2 RMSE (Model 1, levels): 8.9033 | OOS R-squared (Model 1, levels): -0.8249 | Train Size: 1645 | Test Size: 130
Fold 3 RMSE (Model 1, levels): 13.1251 | OOS R-squared (Model 1, levels): -1.1375 | Train Size: 1775 | Test Size: 130
Fold 4 RMSE (Model 1, levels): 5.6604 | OOS R-squared (Model 1, levels): 0.3514 | Train Size: 1905 | Test Size: 130
Fold 5 RMSE (Model 1, levels): 8.903 | OOS R-squared (Model 1, levels): -2.5805 | Train Size: 2035 | Test Size: 130

Average Cross-Validated RMSE (Model 1, levels): 9.4036
Std Dev of Cross-Validated RMSE (Model 1, levels): 2.7114
Average Out-of-Sample R-squared (Model 1, levels): -1.1622
```

For Model 1 (the levels specification), the 5-fold rolling origin cross-validation yielded an average out-of-sample RMSE of approximately 9.404. The average out-of-sample R-squared was -1.162.

```
--- Model 2 (Log-Log): Rolling Origin k-Fold Cross-Validation ---
Fold 1 RMSE (Model 2, log-scale): 0.0755 | OOS R-squared (Model 2, log-scale): -1.5476 | Train Size: 1515 | Test Size: 130
Fold 2 RMSE (Model 2, log-scale): 0.0673 | OOS R-squared (Model 2, log-scale): -0.8579 | Train Size: 1645 | Test Size: 130
Fold 3 RMSE (Model 2, log-scale): 0.0965 | OOS R-squared (Model 2, log-scale): -1.1368 | Train Size: 1775 | Test Size: 130
Fold 4 RMSE (Model 2, log-scale): 0.0448 | OOS R-squared (Model 2, log-scale): 0.3502 | Train Size: 1905 | Test Size: 130
Fold 5 RMSE (Model 2, log-scale): 0.0732 | OOS R-squared (Model 2, log-scale): -2.4683 | Train Size: 2035 | Test Size: 130

Average Cross-Validated RMSE (Model 2, log-scale): 0.0715
Std Dev of Cross-Validated RMSE (Model 2, log-scale): 0.0185
Average Out-of-Sample R-squared (Model 2, log-scale): -1.1321
```

For Model 2 (the log-log specification), the cross-validation resulted in an average out-of-sample RMSE (on the log scale) of approximately 0.0715. The average out-of-sample R-squared on the log scale was -1.132. The substantially negative average out-of-sample R-squared values for both models indicate that, when tasked with predicting on unseen sequential data blocks, their performance was worse than a naive model predicting the mean of the test set. This suggests limitations in their out-of-sample predictive stability, a common challenge with time-series models in dynamic financial markets. While the models demonstrate some capacity for in-sample inference regarding contemporaneous relationships, particularly after HAC correction, their utility for direct forecasting based on this cross-validation is questionable without further refinement or incorporation of adaptive learning mechanisms. These cross-validation findings highlight the distinction between a model's ability to explain in-sample variation and its robustness in out-of-sample prediction.

Discussion

The empirical analysis provides several insights into the drivers of the CBOE SKEW Index. The preferred model, Model 2b (log-log with HAC standard errors), offers the most robust inferences due to better residual properties from the log transformation and the correction for serial correlation and heteroscedasticity. This model explained approximately 17.7% of the daily variation in the SKEW Index.

The statistically significant negative relationship between the SKEW Index and both lagged VIX (log(VIX_lag1)) and lagged realized volatility (log(RealizedVol_lag1)) is a key finding. While seemingly counterintuitive if one expects all fear gauges to move in unison, this suggests a nuanced interplay. It might be that when overall market uncertainty or experienced choppiness is already high, the additional premium demanded for protection against extreme tail events (which SKEW specifically measures) does not increase proportionally, or even decreases. This could occur if very high volatility levels lead market participants to price a wider, more symmetric range of potential outcomes, rather than an increasingly asymmetric one. Alternatively, during extreme market stress reflected by high VIX, liquidity effects or flight-to-quality dynamics might distort the relative pricing of options across strikes that form the SKEW calculation. This finding warrants deeper exploration, perhaps by examining different volatility regimes or conditional models.

Conversely, the lagged Put-Call Ratio (log(PC_Ratio_lag1)) exhibited a statistically significant positive relationship with the SKEW Index. This aligns with financial intuition: an increase in the volume of puts traded relative to calls signals greater demand for downside protection or increased bearish sentiment expressed through options, which would naturally tend to steepen the volatility smirk and elevate the SKEW Index. This variable directly reflects hedging and speculative pressures within the options market itself.

Notably, lagged market returns (MarketReturn_lag1) did not demonstrate a significant direct impact on SKEW in the robust Model 2b, after controlling for other volatility and activity measures. Similarly, the AAII investor sentiment measure (log(Sentiment_lag1)) and its interaction with VIX (logVIX_logSent_Interact_centered_lag1) were found to be insignificant predictors of the SKEW Index once robust standard errors were applied. The lack of significance for the AAII sentiment measure might suggest that this particular retail-focused, weekly sentiment proxy does not sufficiently capture the more dominant institutional forces or higher-frequency information flows that likely drive the pricing of SPX tail risk options. Its influence might also be indirectly subsumed by market-based measures like VIX or the Put-Call Ratio. This finding differs from some literature (e.g., Han, 2008; Baker & Wurgler, 2006) that finds sentiment influential, potentially highlighting differences in the specific sentiment proxy used, the asset class (SKEW is index-level), or the controlling variables.

The diagnostic analysis was crucial. Both initial OLS models exhibited significant autocorrelation and heteroscedasticity, common in daily financial time series, underscoring the necessity of using Newey-West HAC standard errors for valid inference. While log transformations in Model 2 improved residual characteristics to some extent, these issues were not entirely eliminated.

Comparing our findings to the literature, while Xing et al. (2010) focused on individual option smirks predicting equity returns, our study confirms the informational richness of the aggregate market smirk (SKEW) by identifying its contemporaneous drivers. The results regarding VIX contrast with Onan et al. (2014)'s finding that macroeconomic announcements tended to reduce both VIX and skew steepness; our study examines the ongoing baseline relationship rather than event-specific impacts.

Several limitations are acknowledged. The relatively low R-squared value indicates that a substantial portion of SKEW's variation is driven by factors not included in this specification or that the identified relationships lack stable predictive utility out-of-sample, specifically exemplified the modest in-sample explanatory power, coupled with the findings from the rolling origin cross-validation which yielded negative average out-of-sample R-squared values for both models (indicating poor predictive generalization). The choice of sentiment proxy and its weekly frequency could be a limitation. While efforts were made to mitigate endogeneity through lagging predictors, complex feedback loops in financial markets are always a possibility. The study period, while contemporary, is pre-COVID and results might differ in extreme market stress periods.

Future research could explore alternative and higher-frequency sentiment indicators, incorporate macroeconomic surprise variables, or examine more sophisticated time-series models (GARCH models for SKEW volatility, or VAR models). The results from cross-validation particularly underscore the need for future research to explore models with potentially better out-of-sample predictive capabilities, perhaps through dynamic coefficient models, machine learning techniques adapted for time series, or the inclusion of more forward-looking variables. Investigating potential structural breaks or regime-dependent relationships in the SKEW Index determinants could also yield valuable insights.

Reflection

This project centered on the econometric modeling of the CBOE SKEW Index served as a profound and immersive learning experience, extending significantly beyond the foundational concepts of Research Statistics 3. The endeavor to unravel the determinants of a complex financial market indicator like the SKEW Index was both intellectually stimulating and methodologically demanding.

The initial phase of model development, employing standard Ordinary Least Squares (OLS) regression, while theoretically grounded, yielded modest explanatory power, as evidenced by the initial R-squared values. More critically, comprehensive diagnostic testing revealed significant departures from classical OLS assumptions, including pronounced heteroskedasticity and strong serial autocorrelation in the residuals. These findings were pivotal, as they indicated that standard inferences (p-values, confidence intervals) from the initial OLS output would be unreliable.

This necessitated a deeper engagement with advanced econometric techniques to ensure the robustness and validity of the study's conclusions. A substantial portion of the project then evolved into researching and implementing methods specifically designed to address such diagnostic issues. This included the application of Newey-West (HAC) robust standard errors to correct for heteroskedasticity and autocorrelation, a technique that proved essential for obtaining credible statistical inferences. Furthermore, the investigation into multicollinearity, particularly concerning the VIX-related terms and interaction variables, led to the adoption of strategies such as mean-centering predictors. The exploration of logarithmic transformations for key variables was another adaptive step, taken to potentially improve residual behavior and offer interpretations in terms of elasticities.

The data management aspect of this project was also a significant learning exercise. It required the careful collation, cleaning, and merging of data from diverse public sources, including daily market indices (SKEW, VIX, S&P 500), options market statistics (Put-Call Ratios), and weekly survey data (AAII Sentiment). Aligning these different time series, particularly the weekly sentiment data to a daily frequency, and calculating derived variables like rolling volatility, market returns, and interaction terms, demanded meticulous attention to detail and proficiency in data manipulation using R and associated packages like dplyr, zoo, and lubridate, for example, when combining two separate historical files for the Put-Call Ratio into a unified series.

From a theoretical standpoint, this project necessitated a deeper dive into the mechanics of options markets, the concept of implied volatility, the interpretation of the volatility smirk, and the SKEW Index itself. While the initial literature review provided a foundation, the process of selecting appropriate predictors and interpreting their coefficients in the context of financial economics required ongoing self-study and critical thinking to connect statistical outputs back to market dynamics.

Furthermore, while the models identified statistically significant drivers of the SKEW Index, the overall explanatory power, as indicated by the R-squared values, was modest. This observation prompts a reflection on model specification. In hindsight, the inclusion of a broader array of predictor variables could have potentially enhanced the models' predictive accuracy. Factors such as macroeconomic surprise indices, measures of market liquidity, dealer positioning, or even alternative sentiment proxies (beyond the retail-focused AAII survey) might capture additional dimensions of market dynamics influencing tail risk perceptions. This experience highlighted the perennial trade-off between model parsimony and explanatory power, and the importance of extensive domain knowledge in selecting relevant predictors.

The array of R packages utilized also expanded considerably, from data handling with readr and dplyr, to specialized econometric tests with lmtest and sandwich, visualization with ggplot2 and corrplot, and table generation with stargazer. Gaining fluency with these tools has significantly enhanced my analytical toolkit.

Despite the challenges, and perhaps because of them, this project was a genuinely engaging experience—a rigorous application of statistical research principles to a challenging real-world financial problem. The necessity to confront and address the limitations of initial model results by employing more sophisticated techniques proved to be an invaluable learning experience, reinforcing the importance of diagnostic rigor and methodological adaptability in econometric research, and these skills are directly transferable and will undoubtedly inform my approach to future research endeavors in finance and quantitative analysis.

References

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross‐section of stock returns. *The Journal of Finance*, *61*(4), 1645-1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Onan, M., Salih, A., & Yasar, B. (2014). Impact of macroeconomic announcements on implied volatility slope of SPX options and VIX. Finance Research Letters, 11(4), 454–462. https://doi.org/10.1016/j.frl.2014.07.006

Xing, Y., Zhang, X., & Zhao, R. (2010). What Does the Individual Option Volatility Smirk Tell Us About Future Equity Returns? Journal of Financial and Quantitative Analysis, 45(3), 641–662. https://doi.org/10.1017/s0022109010000220

Han, B. (2007). Investor Sentiment and Option Prices. Review of Financial Studies, 21(1), 387–414. https://doi.org/10.1093/rfs/hhm071