

## News-Based Uncertainty and Macro-Market Fusion for VIX Direction Forecasting: Evidence from 2015-2024 FRED Panel

Hailin Zhou\*<sup>1</sup>, Kai Zhang<sup>2</sup>

Email: [hailin.zhou1668@yahoo.com](mailto:hailin.zhou1668@yahoo.com)

<sup>1</sup>Applied Analytics, Columbia University, NY, USA

<sup>2</sup>Financial Engineering, Baruch College, NY, USA

\*Corresponding Author

### Abstract

*This paper evaluates news-based uncertainty and macro-market fusion for one-trading-day VIX direction forecasting using a 2015-2024 daily FRED panel. The dependent variable equals one when the next trading-day VIX close exceeds the current close. The final panel contains 2,512 processed observations from 2015-02-04 to 2024-12-30, with 1,997 training observations from 2015-2022, 257 validation observations in 2023, and 258 holdout observations in 2024. The feature set combines VIX market-state variables, the daily newspaper-based Economic Policy Uncertainty index, the effective federal funds rate, 10-year and 2-year Treasury yields, lagged CPI inflation, lagged unemployment, and interaction terms. Expanding cross-validation on the 2015-2022 training sample gives the highest average ROC AUC to Fusion Random Forest (0.5620). The 2023 validation window selects a weighted fusion ensemble with ROC AUC 0.6005; its weights are fixed before the 2024 test. In the 2024 holdout window, VIX-only Logistic achieves the highest one-day ROC AUC (0.5869) and F1 (0.5448), while the weighted fusion ensemble reaches ROC AUC 0.5638 and F1 0.4786. Event-window diagnostics show that EPU shocks following calm VIX states have a next-day VIX-up rate of 0.5672, compared with 0.4445 on other trading days. The findings support a cautious interpretation: news-based uncertainty contains conditional information, but one-day practical forecasting reliability remains modest and VIX state variables remain the strongest 2024 holdout benchmark.*

**Keywords:** *VIX direction forecasting; Economic Policy Uncertainty; FRED; Treasury yield curve; machine learning.*

### I. INTRODUCTION

Volatility forecasting is central to risk management, hedging, portfolio allocation, and stress monitoring. The Cboe Volatility Index (VIX) condenses option-implied expectations of near-term S&P 500 volatility and is widely used as a market stress measure. Classical volatility research shows that volatility is persistent, clustered, and asymmetric, while implied-volatility research shows that option-based volatility measures contain information about future risk states. These findings justify using lagged VIX levels, changes, returns, and rolling volatility as the first block of predictors.

The forecasting question is not limited to volatility persistence. Public information can change the interpretation of risk even before it appears fully in price dynamics. News-based uncertainty measures are therefore useful because they translate public newspaper coverage into a daily macro-financial signal. In this study, the news signal is the FRED Economic Policy Uncertainty index for the United States. It is combined with interest-rate, inflation, unemployment, and VIX-state variables to test whether a transparent FRED panel can improve next-day VIX direction classification. The empirical design expands the sample to a ten-year panel rather than a single-

year case. After lag and rolling-window construction, the panel contains 2,512 observations and 71 full-fusion features. This larger sample directly reduces the feature-to-observation imbalance that would arise in a one-year design. The 2024 holdout window is kept separate from model development, while 2023 is used only for validation and ensemble-weight selection.

The paper evaluates three hypotheses. H1 is that news-based uncertainty variables contain measurable information about next-day VIX direction. H2 is that macro-market fusion improves probability ranking in at least the validation or robustness setting relative to single-block models. H3 is that high EPU shocks following calm VIX states are followed by a higher next-day VIX-up rate than other trading days. These hypotheses are evaluated with time-ordered cross-validation, a validation window, a 2024 holdout test, ablation analysis, event-window diagnostics, feature importance, and horizon robustness. The results are intentionally framed as measured evidence rather than as proof of a deployable trading rule. The best one-day 2024 ROC AUC is 0.5869, and the weighted fusion ensemble ROC AUC is 0.5638. These values indicate modest signal. The main contribution is therefore diagnostic: a transparent FRED-based news and macro panel can reveal conditional uncertainty effects, but operational use requires threshold monitoring, additional years, and model governance.

## **II. LITERATURE REVIEW**

Volatility modeling began with conditional variance models. Engle (1982) introduced ARCH, and Bollerslev (1986) extended it to GARCH. Glosten et al. (1993) modeled asymmetric volatility responses, and Andersen and Bollerslev (1998) emphasized disciplined forecast evaluation. These studies motivate using lagged volatility, recent changes, and rolling volatility features rather than relying only on a contemporaneous VIX level.

Implied-volatility research connects option markets to forward-looking risk expectations. Whaley (1993, 2000) established the VIX as an interpretable volatility benchmark, and Fleming et al. (1995) documented forecasting content in implied-volatility indexes. Giot (2005) showed that implied volatility indexes are closely connected to equity-market risk states. The present target follows this literature by forecasting whether the next trading-day VIX close is higher than the current close.

Text and uncertainty research explains why news-based variables may matter for volatility. Tetlock (2007) and Tetlock et al. (2008) show that media language is related to market outcomes. Loughran and McDonald (2011) demonstrate the importance of domain-specific textual measurement in finance. Baker et al. (2016) construct economic policy uncertainty indexes from newspaper coverage and show that uncertainty is economically meaningful. Manela and Moreira

(2017) also link news-implied volatility and disaster concerns to risk premia. The daily EPU index used here provides a transparent, directly observed news-based uncertainty proxy.

Machine-learning models can capture nonlinear interactions among VIX state, uncertainty, and macro conditions. Random forests reduce variance through tree aggregation (Breiman, 2001), while gradient boosting builds additive nonlinear decision rules (Friedman, 2001). Logistic regression remains important because it provides a stable and interpretable probability baseline. This paper uses linear, sparse, tree-based, and ensemble models, but it avoids overstating model complexity because the 2024 holdout evidence remains modest.

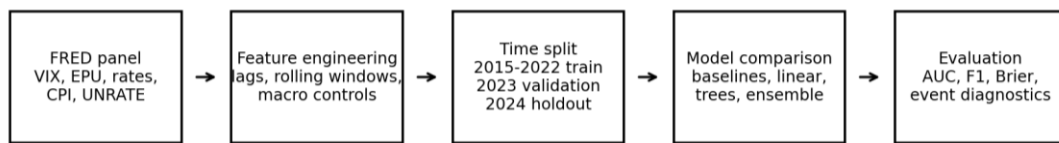


Figure 1. Experimental pipeline for the FRED-based news and macro-market fusion design

### III. RESEARCH METHOD

The empirical pipeline is summarized in Figure 1. The design begins with FRED series, constructs leakage-controlled daily features, uses a 2015-2022 training period, selects ensemble weights in 2023, and reserves 2024 for final holdout evaluation.

Table 1. Data sources and roles in the 2015-2024 FRED panel

Source	Input	Scope	Derived Features	Role
FRED / CBOE VIXCLS	Daily VIX close	2015-2024 trading days	VIX level, changes, returns, moving averages, volatility	Target construction and market volatility state
FRED / USEPUIN DXD	Daily news-based Economic Policy Uncertainty Index	2015-2024 calendar days aligned to trading days	EPU level, changes, rolling means, volatility, z-scores, lags	News-based uncertainty signal
FRED / DFF	Daily effective federal funds rate	2015-2024 daily panel	Policy-rate level, changes, lagged changes, real-rate proxy	Monetary-policy control
FRED / DGS10 and DGS2	Daily Treasury yields	2015-2024 trading-day panel	10-year yield, 2-year yield, term spread, rate changes, rate volatility	Yield-curve and macro-financial conditions
FRED / CPIAUCSL and UNRATE	Monthly CPI and unemployment rate	Monthly observations shifted one month and aligned to trading days	Lagged CPI inflation, CPI monthly change, unemployment level and change	Inflation and labor-market controls

Table 1 lists the data sources. VIXCLS provides the target and market-volatility state. USEPUINDXD provides the news-based uncertainty signal. DFF, DGS10, DGS2, CPIAUCSL, and UNRATE provide monetary-policy, yield-curve, inflation, and labor-market controls. Daily series are aligned to VIX trading dates using the latest available observation. Monthly CPI and

unemployment variables are shifted forward by one month before daily forward-filling, which avoids using same-month macro values as if they were known on every day of that month.

Table 2 reports the feature inventory. The full fusion block contains 71 variables: 20 VIX market-state features, 20 news-based uncertainty features, 27 macro-rate and inflation-labor controls, and 4 interaction terms. All lagged and rolling features are computed before target alignment, so the predictor vector at date  $t$  is used only to forecast the direction from  $t$  to  $t+1$ .

Table 2. Feature inventory and modeling purpose

Block	Feature Count	Variables	Modeling Purpose
VIX market state	20	VIX level, daily change, return, rolling means, rolling volatility, lagged changes and returns	Captures volatility persistence, stress level, and short-run reversal
News-based uncertainty	20	Daily EPU level, log EPU, changes, 5-day and 22-day means and volatility, z-scores, lags	Measures newspaper-based policy uncertainty and attention shocks
Macro rates and inflation-labor controls	27	Federal funds rate, 10-year and 2-year Treasury yields, term spread, real-rate proxies, lagged CPI and unemployment	Controls for policy stance, yield-curve state, inflation pressure, and labor conditions
Fusion interactions	4	EPU-VIX, EPU-curve, VIX-real-rate, and EPU-inflation interactions	Allows uncertainty signals to depend on volatility and macro states
Full fusion	71	Union of all VIX, uncertainty, macro, and interaction features	Tests the news-macro-market fusion hypothesis

The target is defined as  $y_t = 1[VIX_{t+1} > VIX_t]$ . After removing rows lost to lagged features, rolling features, and target construction, the final one-day panel contains 2,512 observations. Table 3 shows the training, validation, and holdout windows. The holdout positive rate is 0.4457, so accuracy alone is not sufficient; precision, recall, F1, ROC AUC, PR AUC, and Brier score are reported together.

Table 3. Sample construction and target distribution

Sample	N	Start	End	Positive Target Rate
Processed observations	2512	2015-02-04	2024-12-30	0.4510
Training observations	1997	2015-02-04	2022-12-30	0.4537
Validation observations	257	2023-01-03	2023-12-29	0.4358
Holdout observations	258	2024-01-02	2024-12-30	0.4457

Table 4 provides descriptive statistics for the main variables. The VIX average is 18.2589 over the processed sample, with a maximum of 82.6900 during the 2020 stress period. The EPU series is highly dispersed, with a mean of 123.5088 and a maximum of 1026.3800. These distributions justify using rolling transformations, z-scores, and lagged features instead of raw levels alone.

Table 4. Descriptive statistics for main variables

Variable	count	mean	std	min	25%	50%	75%	max
Vix	2512.0000	18.2589	7.3037	9.1400	13.3300	16.2700	21.3425	82.6900
vix_change	2512.0000	0.0000	1.9116	-17.6400	-0.7600	-0.1000	0.5700	24.8600
Epu	2512.0000	123.5088	94.2352	3.3200	67.1200	96.7650	144.8325	1026.3800
epu_change	2512.0000	0.0290	76.1273	-952.7800	-31.4150	-1.1800	31.0825	872.8200
epu_z22	2512.0000	0.0024	1.0246	-2.5253	-0.7093	-0.2068	0.5390	4.3136
Dff	2512.0000	1.8131	1.8792	0.0400	0.1300	1.1600	2.4000	5.3300
dgs10	2512.0000	2.4952	1.0591	0.5200	1.7300	2.3300	3.1600	4.9800
dgs2	2512.0000	2.0283	1.5791	0.0900	0.7000	1.5800	3.1550	5.1900
term_spread_10y2y	2512.0000	0.4669	0.6638	-1.0800	0.0700	0.4700	1.0100	1.7700
cpi_yoy_lag1m	2512.0000	2.8980	2.2900	-0.2299	1.4876	2.2915	3.3608	8.9794
unrate_lag1m	2512.0000	4.6820	1.7308	3.4000	3.7000	4.1000	5.0000	14.8000
real_10y_rate	2512.0000	-0.4027	2.1622	-6.3094	-0.6236	0.3292	0.8787	2.4650

Table 5 reports selected correlations. The raw correlations with the next-day target are small, which is expected for a one-day direction problem. VIX level and VIX z-score have the largest absolute correlations with the target, while EPU variables are weaker unconditionally. This pattern motivates both ablation tests and conditional event-window analysis rather than a claim that news uncertainty dominates the VIX state by itself.

Table 5. Selected correlations with next-day VIX direction and current VIX change

Variable	Corr. with VIX-up Target	Corr. with Current VIX Change
Vix	-0.1095	0.1309
vix_z22	-0.1057	0.3592
epu_ma22	-0.0551	-0.0455
vix_change	-0.0531	1.0000
Epu	-0.0382	-0.0534
epu_change	0.0267	-0.0015
unrate_lag1m	-0.0218	-0.0123
cpi_yoy_lag1m	-0.0164	0.0041
term_spread_10y2y	0.0156	-0.0067
real_10y_rate	0.0154	-0.0060
dgs2	-0.0092	0.0004
Dff	-0.0079	0.0025
dgs10	-0.0039	-0.0035
epu_z22	0.0020	-0.0049

The model set is reported in Table 6. Logistic models use standardized inputs and balanced class weights. The sparse fusion model applies ANOVA feature selection with  $k = 24$  before logistic regression. Random forest and gradient boosting provide nonlinear alternatives. The weighted fusion ensemble combines four probability series: News Uncertainty Logistic, Fusion SelectK

Logistic, VIX-only Logistic, and Fusion Gradient Boosting. Its weights are selected by a convex grid search on the 2023 validation window and then fixed for the 2024 holdout test.

Table 6. Model specifications and hyperparameters

Model	Inputs	Main Hyperparameters	Purpose
Majority Class	none	most frequent training class; probability equals training positive rate	naive class-balance benchmark
VIX Lag Rule	vix_change_lag1	predicts VIX-up when lagged VIX change is positive	simple volatility-persistence benchmark
VIX-only Logistic	VIX market state block	standardized logistic regression, C=0.7, balanced class weights	parametric VIX baseline
News Uncertainty Logistic	news-based uncertainty block	standardized logistic regression, C=0.7, balanced class weights	news-only baseline
Macro-Market Logistic	VIX + macro-rate controls	standardized logistic regression, C=0.7, balanced class weights	macro-market baseline
Fusion Logistic	all features	standardized logistic regression, C=0.7, balanced class weights	interpretable full fusion model
Fusion SelectK Logistic	all features	ANOVA feature selection k=24 plus standardized logistic regression	sparse fusion model for correlated features
Fusion Random Forest	all features	120 trees, max depth 5, minimum leaf 12	nonlinear bagging model
Fusion Gradient Boosting	all features	120 trees, depth 2, learning rate 0.045, minimum leaf 10	additive nonlinear model
Fusion Weighted Ensemble	news, sparse fusion, VIX, and boosting probabilities	convex weights selected on the 2023 validation window	main probability-ranking forecast

#### IV. FINDINGS AND DUSCUSSION

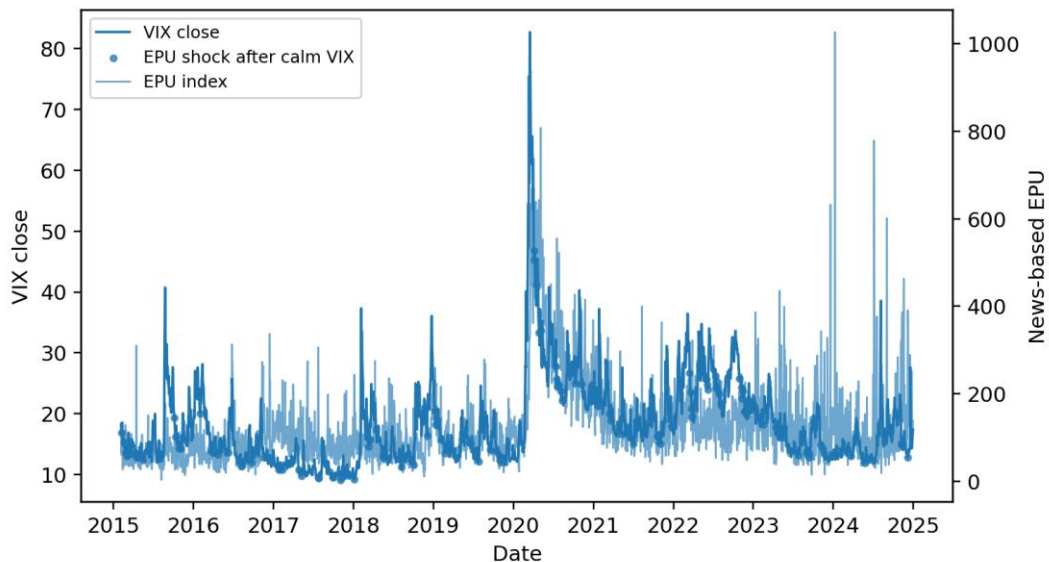


Figure 2. VIX path, EPU path, and conditional EPU shock windows

Figure 2 plots the VIX and EPU paths over the processed sample. The figure shows that the sample includes the 2020 volatility shock, the inflation and rate-hike period, and the 2024 holdout year. The marked event windows are not simply high VIX days; they are days when EPU is

elevated while VIX is below its recent trend, which is the conditional setting used in the event-window diagnostic. Figure 3 shows the yearly next-day VIX-up rate. The positive class is close to balanced but varies by year, which supports the use of balanced accuracy and AUC in addition to raw accuracy. Figure 4 shows that VIX, EPU, and macro-rate variables are correlated in economically sensible ways, especially during stress and rate-cycle periods.

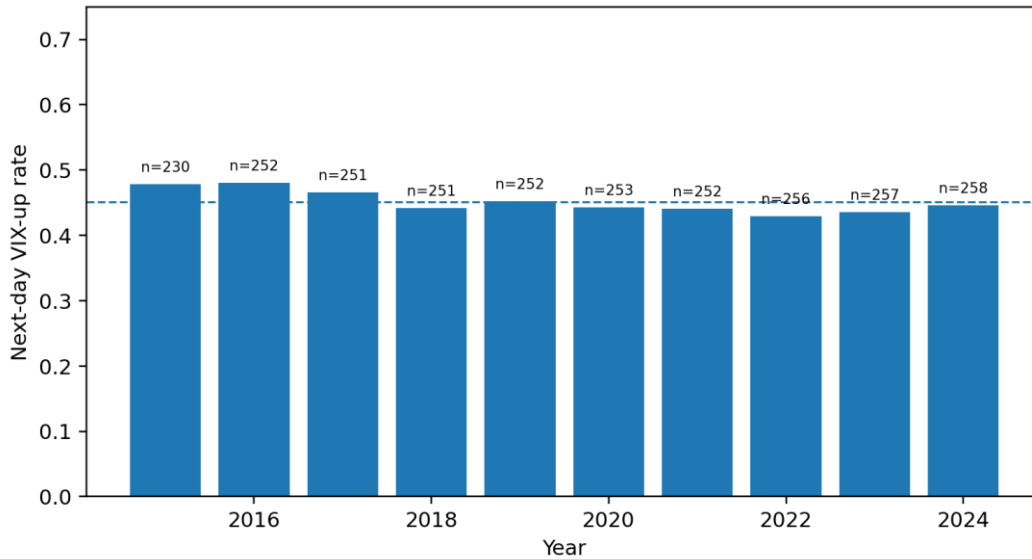


Figure 3. Yearly next-day VIX-up rates in the processed panel

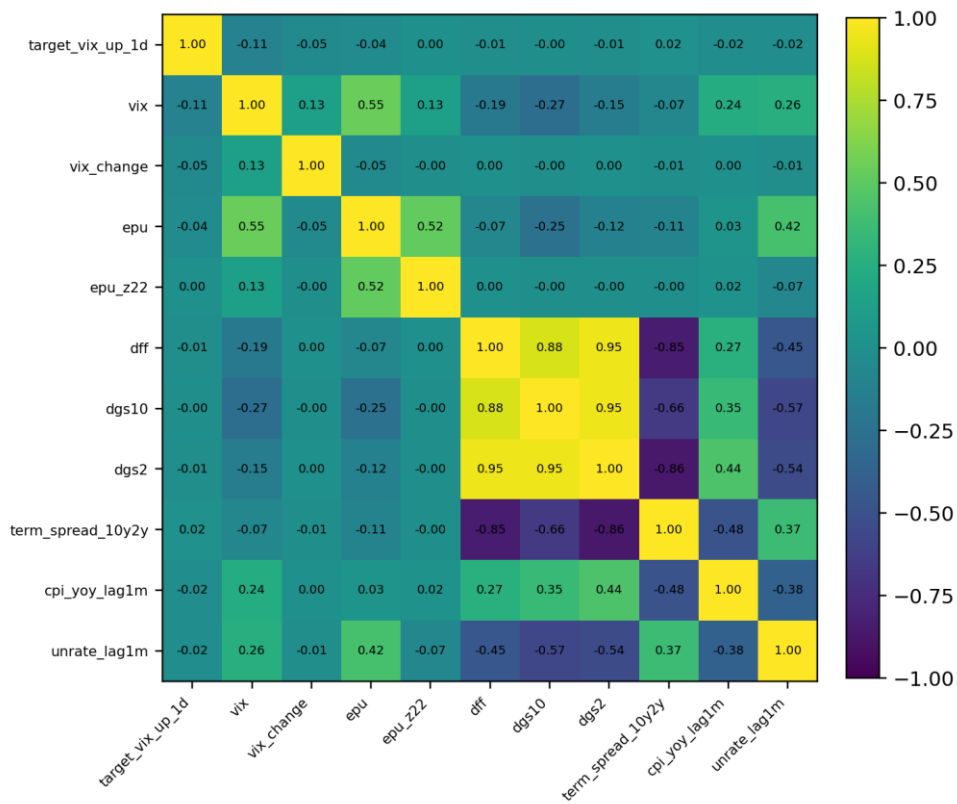


Figure 4. Correlation structure of selected target, uncertainty, VIX, and macro variables

Table 7 reports expanding time-series cross-validation on the 2015-2022 training sample. Fusion Random Forest has the highest average ROC AUC (0.5620), followed closely by VIX-only Logistic (0.5597) and News Uncertainty Logistic (0.5550). The differences are small, so the cross-validation evidence supports only a modest predictive signal.

Table 7. Expanding time-series cross-validation metrics

Model	Family	Feature Count	Accuracy	Balanced Accuracy	F1	ROC AUC	PR AUC	Brier
Fusion Random Forest	Fusion	71	0.5602	0.5387	0.4114	0.5620	0.5147	0.2453
VIX-only Logistic	VIX	20	0.5407	0.5316	0.4766	0.5597	0.5065	0.2497
News Uncertainty Logistic	News	20	0.5526	0.5226	0.3369	0.5550	0.5009	0.2647
Macro-Market Logistic	Macro+Market	47	0.5420	0.5235	0.4274	0.5480	0.4966	0.3147
Fusion Logistic	Fusion	71	0.5382	0.5233	0.4394	0.5428	0.4932	0.3191
Fusion Gradient Boosting	Fusion	71	0.5370	0.5169	0.3989	0.5377	0.5018	0.2612
Fusion SelectK Logistic	Fusion	71	0.5307	0.5173	0.4339	0.5369	0.4948	0.2787
Majority Class	Baseline	0	0.5489	0.5000	0.0000	0.5000	0.4511	0.2479
VIX Lag Rule	Baseline	1	0.4818	0.4759	0.4243	0.4759	0.4406	0.5182

Table 8 reports the 2023 validation results used for model development. The weighted fusion ensemble has the highest validation ROC AUC of 0.6005. Table 9 shows the selected weights: 0.1000 for News Uncertainty Logistic, 0.1000 for Fusion SelectK Logistic, 0.6000 for VIX-only Logistic, and 0.2000 for Fusion Gradient Boosting. The high VIX-only weight is economically informative because validation performance still depends strongly on volatility-state variables.

Table 8. Validation results for ensemble construction

Model	Family	Feature Count	Accuracy	Balanced Accuracy	F1	ROC AUC	PR AUC	Brier
Fusion Weighted Ensemble	Fusion	71	0.5486	0.5553	0.5397	0.6005	0.5451	0.2446
VIX-only Logistic	VIX	20	0.5097	0.5340	0.5625	0.5890	0.5582	0.2490
Fusion Gradient Boosting	Fusion	71	0.5914	0.5678	0.4503	0.5800	0.5159	0.2432
Fusion SelectK Logistic	Fusion	71	0.5992	0.5757	0.4607	0.5499	0.4971	0.2452
News Uncertainty Logistic	News	20	0.5720	0.5577	0.4762	0.5402	0.4916	0.2460

Table 10 gives the main 2024 holdout comparison. VIX-only Logistic has the highest one-day ROC AUC (0.5869) and F1 (0.5448). Macro-Market Logistic has the highest accuracy (0.5736) and the lowest Brier score among the reported models (0.2428). The weighted fusion ensemble

reaches ROC AUC 0.5638 and F1 0.4786. These results do not show a strong one-day forecasting model; they show a small but structured signal that should be interpreted cautiously.

Table 9. Validation-selected ensemble weights

Component	Validation Weight
News Uncertainty Logistic	0.1000
Fusion SelectK Logistic	0.1000
VIX-only Logistic	0.6000
Fusion Gradient Boosting	0.2000

The ROC curves in Figure 5 and the AUC comparison in Figure 6 show the same pattern visually. VIX-only and macro-market specifications rank holdout observations more effectively than the pure news model, while the fusion ensemble remains above the majority and lag-rule benchmarks but below the best VIX-only baseline. This ranking is consistent with the moderate conclusion that news-based uncertainty helps in selected states rather than replacing volatility-state information.

Table 10. 2024 holdout experimental comparison on next-day VIX direction

Model	Family	FC	Acc	BA	Prec	Rec	F1	ROC AUC	PR AUC	Brier
VIX-only Logistic	VIX	20	0.5271	0.5377	0.4771	0.6348	0.5448	0.5869	0.5834	0.2448
Macro-Market Logistic	Macro+Market	47	0.5736	0.5532	0.5316	0.3652	0.4330	0.5706	0.5288	0.2428
Fusion Weighted Ensemble	Fusion	71	0.5271	0.5232	0.4706	0.4870	0.4786	0.5638	0.5213	0.2450
Fusion Gradient Boosting	Fusion	71	0.5659	0.5394	0.5231	0.2957	0.3778	0.5577	0.5222	0.2455
Fusion Random Forest	Fusion	71	0.5465	0.5398	0.4911	0.4783	0.4846	0.5563	0.5409	0.2467
Fusion SelectK Logistic	Fusion	71	0.5581	0.5461	0.5051	0.4348	0.4673	0.5420	0.4830	0.2522
Fusion Logistic	Fusion	71	0.5465	0.5245	0.4868	0.3217	0.3874	0.5231	0.4668	0.2604
News Uncertainty Logistic	News	20	0.5155	0.5136	0.4597	0.4957	0.4770	0.5147	0.4724	0.2637
Majority Class	Baseline	0	0.5543	0.5000	0.0000	0.0000	0.0000	0.5000	0.4457	0.2471
VIX Lag Rule	Baseline	1	0.4961	0.4901	0.4348	0.4348	0.4348	0.4901	0.4410	0.5039

FC = Feature Count; Acc = Accuracy; BA = Balanced Accuracy; Prec = Precision; Rec = Recall

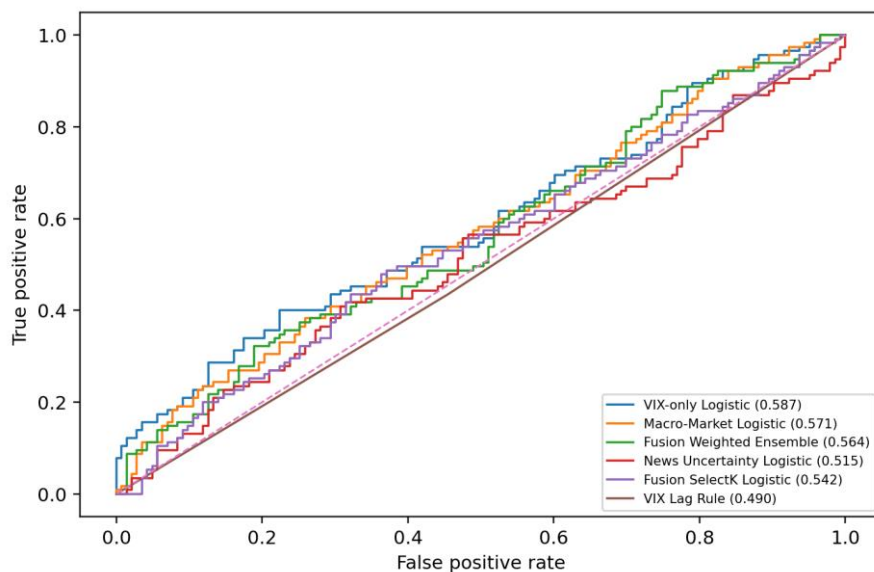


Figure 5. 2024 holdout ROC curves for the main comparison models

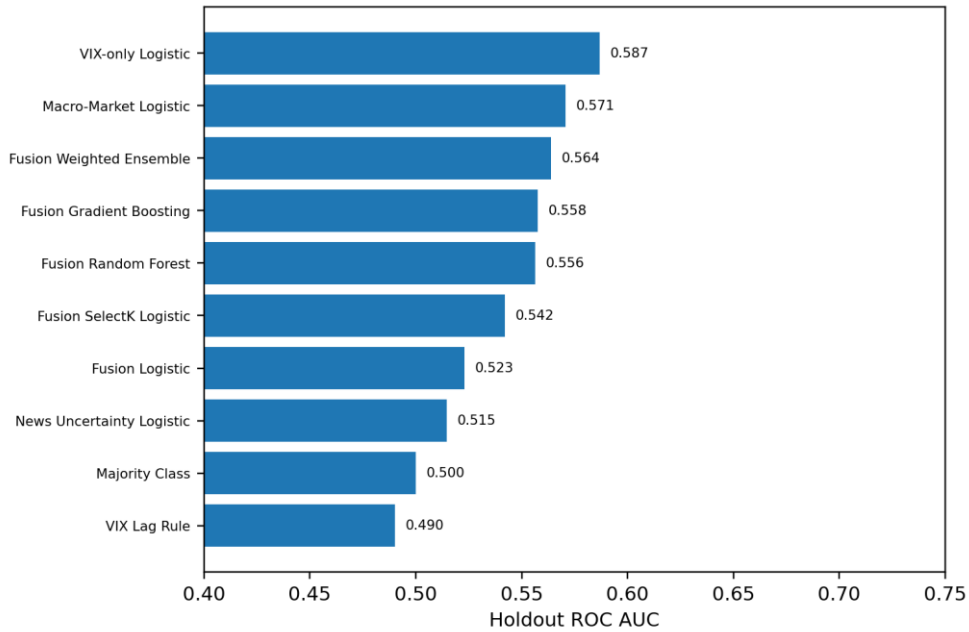


Figure 6. 2024 holdout model comparison by ROC AUC

Table 11 and Figure 7 report the confusion matrix for the weighted fusion ensemble at the standard 0.50 threshold. The model correctly classifies 80 down/flat days and 56 VIX-up days, while producing 63 false positives and 59 false negatives. The resulting positive-class recall is 0.4870. This is not high enough for a standalone early-warning system, but it provides a transparent baseline for threshold adjustment in applications that value recall more than false-alarm control.

Table 11. Confusion matrix for the weighted fusion ensemble

<b>Confusion Matrix</b>	<b>Predicted Down/Flat</b>	<b>Predicted Up</b>
Actual Down/Flat	80	63
Actual Up	59	56

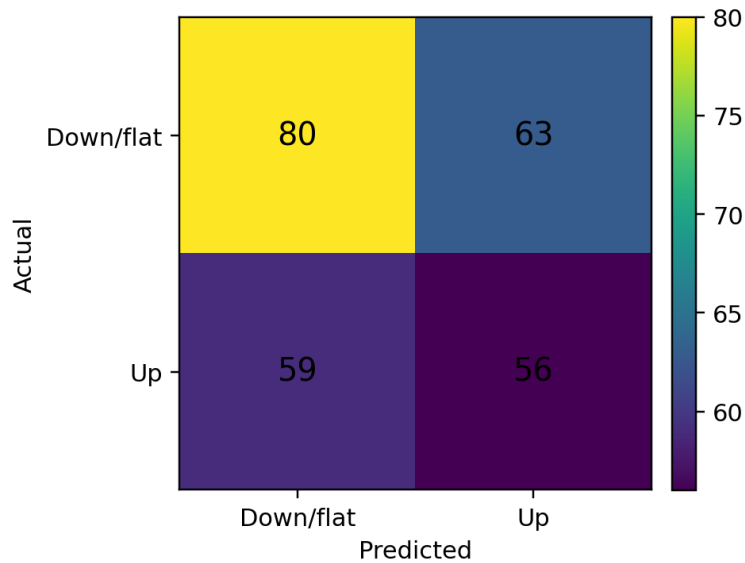


Figure 7. Confusion matrix for the weighted fusion ensemble

Table 12 reports ablation results relative to the weighted fusion ensemble. The VIX-only model exceeds the ensemble by 0.0230 ROC AUC and 0.0661 F1. Macro-Market Logistic exceeds the ensemble by 0.0068 ROC AUC but has lower F1. News Uncertainty Logistic has similar F1 to the ensemble but a lower ROC AUC. The ablation results show that one-day direction remains dominated by VIX-state variables, while EPU and macro variables are more useful as conditional diagnostics and robustness features.

Table 12. Ablation study against the weighted fusion ensemble

Model	Family	Feature Count	F1	ROC AUC	Brier	AUC Delta vs. Ensemble	F1 Delta vs. Ensemble
VIX-only Logistic	VIX	20	0.5448	0.5869	0.2448	0.0230	0.0661
Macro-Market Logistic	Macro+Market	47	0.4330	0.5706	0.2428	0.0068	-0.0456
Fusion Weighted Ensemble	Fusion	71	0.4786	0.5638	0.2450	0.0000	0.0000
Fusion Gradient Boosting	Fusion	71	0.3778	0.5577	0.2455	-0.0061	-0.1009
Fusion Random Forest	Fusion	71	0.4846	0.5563	0.2467	-0.0075	0.0059
Fusion SelectK Logistic	Fusion	71	0.4673	0.5420	0.2522	-0.0218	-0.0113
Fusion Logistic	Fusion	71	0.3874	0.5231	0.2604	-0.0407	-0.0912
News Uncertainty Logistic	News	20	0.4770	0.5147	0.2637	-0.0491	-0.0016
Majority Class	Baseline	0	0.0000	0.5000	0.2471	-0.0638	-0.4786
VIX Lag Rule	Baseline	1	0.4348	0.4901	0.5039	-0.0737	-0.0438

Figure 8 shows permutation importance for the sparse fusion model. VIX log level, VIX z-score, VIX return, and EPU log level appear among the leading features. The ranking is economically coherent because it places volatility state first while retaining EPU level as an incremental signal. The small magnitude of most importance values also reinforces the conclusion that the forecasting signal is diffuse and modest.

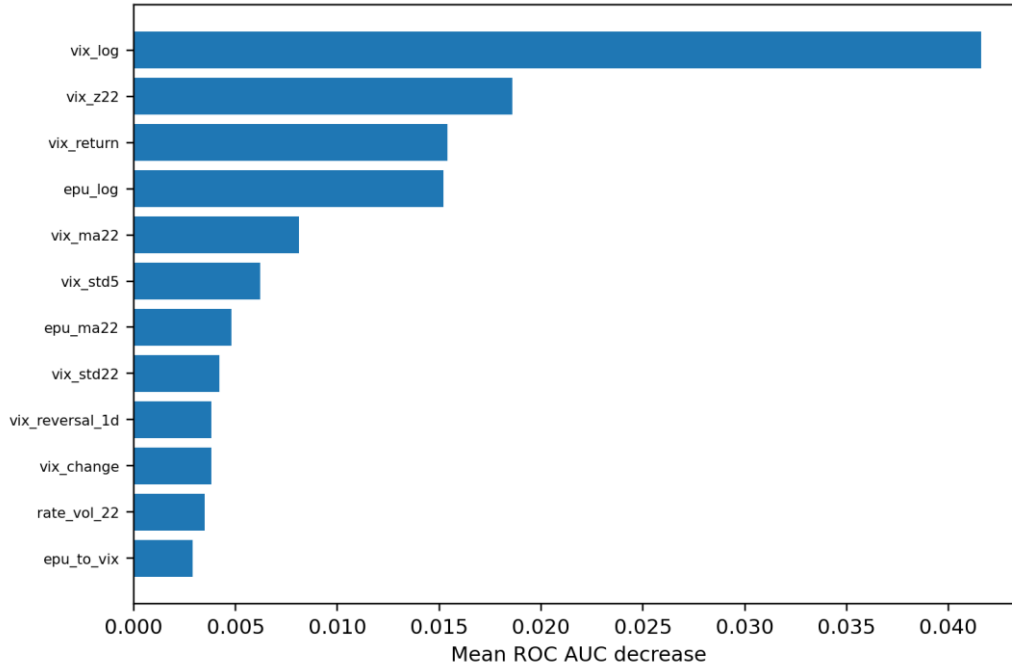


Figure 8. Permutation feature importance for the sparse fusion model

Table 13 evaluates conditional event windows. EPU shocks after calm VIX states have 134 observations and a next-day VIX-up rate of 0.5672. Other trading days have a next-day VIX-up rate of 0.4445. The difference of 0.1227 supports H3: news-based uncertainty is most informative when it arrives before volatility has already moved above its recent trend.

Table 13. Event-window diagnostics for EPU shocks after calm VIX states

Subsample	N	Positive Target Rate	Mean VIX Change	Mean EPU	Mean EPU z-score	Mean VIX z-score
EPU shock after calm VIX state	134	0.5672	-0.6628	212.1574	1.4576	-1.3722
Other trading days	2378	0.4445	0.0374	118.5135	-0.0796	0.0251

Figure 9 plots calibration curves for selected models. The curves show that predicted probabilities require monitoring and may not be perfectly calibrated in the 2024 holdout year. This matters for practice because a ranking model and a capital-allocation model are not the same object: even when AUC is above 0.50, probability levels should be treated cautiously.

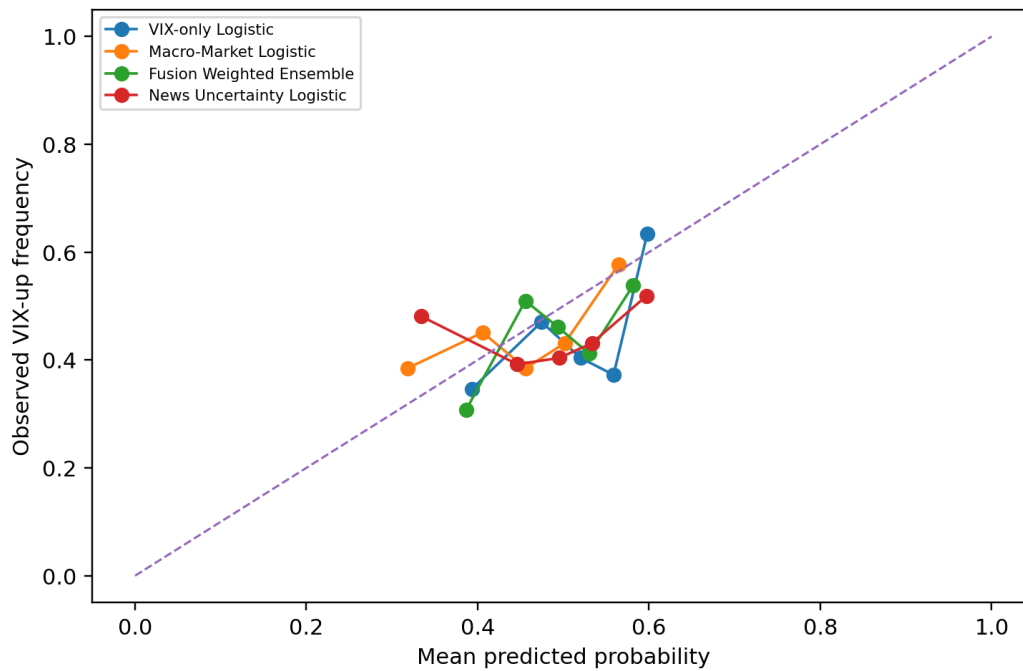


Figure 9. Holdout calibration curves for selected models

Table 14 reports horizon robustness. At the one-day horizon, VIX-only Logistic remains the best AUC model. At the two-day horizon, Fusion Gradient Boosting has the highest ROC AUC of 0.6319. At the five-day horizon, Macro-Market Logistic has the highest ROC AUC of 0.6842. This pattern suggests that one-day direction is dominated by VIX state, while longer horizons allow macro-market conditions and nonlinear fusion to become more informative.

Table 14. Horizon robustness for one-day, two-day, and five-day VIX direction

Forecast Horizon	Best Model	Accuracy	Balanced Accuracy	F1	ROC AUC	PR AUC	Brier	Test Positive Rate
1-day	VIX-only Logistic	0.5271	0.5377	0.5448	0.5869	0.5834	0.2448	0.4457
2-day	Fusion Gradient Boosting	0.5953	0.5786	0.4286	0.6319	0.6396	0.2338	0.4669
5-day	Macro-Market Logistic	0.6378	0.6378	0.5354	0.6842	0.7009	0.2330	0.5000

## V. CONCLUSION AND RECOMMENDATION

This paper revises the forecasting design around a 2015-2024 FRED panel and evaluates news-based uncertainty and macro-market fusion for next-day VIX direction. The expanded sample contains 2,512 processed observations and removes the need for document-level narrative scoring. The study uses a time-ordered training, validation, and holdout design, multiple baselines, linear and nonlinear models, validation-selected ensemble weights, ablation analysis, feature importance, event-window diagnostics, and horizon robustness. The evidence is qualified. The 2023 validation results favor the weighted fusion ensemble, and the conditional event analysis shows that EPU shocks following calm VIX states have a higher next-day VIX-up rate.

However, the 2024 one-day holdout results favor VIX-only Logistic in ROC AUC and F1. The practical conclusion is therefore not that fusion uniformly dominates, but that news-based uncertainty provides a useful supplementary signal when interpreted with VIX state and macro controls.

For applied risk monitoring, the recommended workflow is a two-layer dashboard. The first layer should maintain a VIX-state model as the primary one-day benchmark. The second layer should monitor EPU shocks, macro-rate conditions, and fusion probabilities as diagnostic overlays. Institutions that value early warning should tune thresholds on validation data to increase recall, while institutions that value fewer false alarms should require stronger probability confirmation. The current results should be viewed as modest predictive evidence, not as a standalone trading strategy. Future work should expand the same FRED-based framework with additional real-time vintages, alternative volatility targets, and rolling out-of-sample windows. The current ten-year evidence shows that transparent news-based uncertainty can be studied without opaque text extraction, and it clarifies where the signal is strongest: conditional states in which policy uncertainty rises before VIX has already moved above its recent trend.

## REFERENCES

- Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885-905. <https://doi.org/10.2307/2527343>
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636. <https://doi.org/10.1093/qje/qjw024>
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Chicago Board Options Exchange. CBOE Volatility Index: VIX [VIXCLS]. Retrieved from FRED, Federal Reserve Bank of St. Louis. <https://fred.stlouisfed.org/series/VIXCLS>
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007. <https://doi.org/10.2307/1912773>
- Federal Reserve Bank of St. Louis. Federal Reserve Economic Data. <https://fred.stlouisfed.org>
- Fleming, J., Ostdiek, B., & Whaley, R. E. (1995). Predicting stock market volatility: A new measure. *Journal of Futures Markets*, 15(3), 265-302. <https://doi.org/10.1002/fut.3990150303>

- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Giot, P. (2005). Relationships between implied volatility indexes and stock index returns. *The Journal of Portfolio Management*, 31(3), 92-100. <https://doi.org/10.3905/jpm.2005.500363>
- Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5), 1779-1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
- Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), 873-889. <https://doi.org/10.1002/jae.800>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Manela, A., & Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1), 137-162. <https://doi.org/10.1016/j.jfineco.2016.01.032>
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168. <https://doi.org/10.1111/j.1540-6261.2007.01232.x>
- Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3), 1437-1467. <https://doi.org/10.1111/j.1540-6261.2008.01362.x>
- Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. *The Journal of Derivatives*, 1(1), 71-84. <https://doi.org/10.3905/jod.1993.407868>
- Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management*, 26(3), 12-17. <https://doi.org/10.3905/jpm.2000.319728>