Predicting SPY with Option Skew and the VIX Curve

Vishruth Anand
MGT 4067 – Individual Project
GitHub: `https://github.com/vishruthanand08/iv-skew-strategy`

April 2025

1 Introduction

When equity markets get jittery, financial news outlets point to the VIX—the CBOE's "fear gauge," which measures the market's expectation of 30-day SP500 volatility derived from near-dated option prices. But the VIX is only one snapshot on the implied-volatility surface. The entire curve—how steep it is across maturities and how expensive downside protection is relative to upside—may provide richer information about investors' risk perceptions. In this project I ask a focused question: When the one-month VIX trades significantly above the three-month VIX (a steep term structure) and the SKEW index is sitting in its upper tail (downside puts are unusually costly), does that foreshadow a negative return for SPY over the next month? To investigate, I construct three progressively more sophisticated trading rules that use just two features—(i) a 252-day percentile rank of SKEW and (ii) the VIX/VIX3M ratio. Each strategy's returns are scaled to a common 10 percent annualized volatility target so that performance comparisons, particularly Sharpe ratios, are made on an equal-risk basis.

2 Data

Three CSVs in `/data/vol/`—daily VIX, VIX3M, and SKEW straight from CBOE (1990–2025 for VIX/SKEW, 2009–2025 for VIX3M). SPY prices come from Yahoo Finance on the fly; the notebook caches them so the code is still reproducible offline.

Key feature math.

$$\text{Slope}_t = \frac{\text{VIX}_t}{\text{VIX3M}_t}, \qquad \text{SkewPct}_t = \frac{\#\,\text{days SKEW} \le \text{SKEW}_t}{252}.$$

Everything is snapped to month-end. I slide SPY forward one month so today's features line up with next-month performance—what I actually want to predict. Table 1 lists exact file spans.

File	Start	End
VIX_History.csv	1990-01-02	2025-04-22
VIX3M_History.csv	2009-01-02	2025-04-22
SKEW_History.csv	1990-01-02	2025-04-22

Table 1: Date coverage of raw files (all daily).

3 Methodology

Vol-Target Wrapper (keep risk apples-to-apples)

I scale every strategy so its rolling 12-month realised vol sticks near 10 %. That way Sharpe really means "excess return per 10% vol unit," not "who took the craziest swing."

Three Strategy Flavors

1. **Naïve filter.** Long if `skewPct` < 0.20, short if > 0.80. That's it.
2. **Continuous z-score rule.** Multiply $-z(\text{SkewPct})$ by $z(\text{Slope})$, clamp to ± 1, blank out a neutral zone, then vol-target. I grid-searched three low/high bands and four slope cut-offs—twelve combos total.
3. **Logistic regression.** Feed {SkewPct, Slope} into `sklearn.LogisticRegression`. Train on 1993–2018, predict 2019–2025. A probability > 0.5 is a long; else short.

4 Results

Strategy	Sharpe	CAGR	Hit Rate
Skew 20/80 (baseline)	0.12	0.8 %	22 %
Best z-score rule	0.77	0.3 %	19 %
LogReg (train 93-18)	1.09	15.1 %	72 %
LogReg (OOS 19-25)	0.78	7.5 %	64 %

Table 2: All returns scaled to 10 % annual vol.

What jumps out? The baseline strategy, which just goes long when SKEW is in the bottom 20

The logistic regression model is where things really start to work. I trained it on monthly data from 1993 through 2018 using just two features—skew percentile and slope—and tested it on 2019 through 2025. In-sample, it delivers a Sharpe over 1.0 and a 15 % CAGR, which is solid. Out-of-sample, it cools down to a Sharpe of 0.78 and a 7.5% CAGR, with a hit rate of 64% . That kind of dropoff is expected, but it's still strong performance for a model using only two technical features and no tuning beyond a 0.5 classification threshold. The ROC curve in the test set gives an AUC of 0.55—not amazing, but it confirms there's some real edge. Overall, the regression strategy outperforms both the rule-based approaches and seems to hold up even when applied to newer data.

5 Conclusion

6 Conclusion

This project started with a simple idea: can the shape of the options surface—specifically the relationship between short- and medium-term implied volatility (VIX vs. VIX3M) and the CBOE SKEW index—tell us anything useful about next-month SPY returns? I began with a basic threshold rule using just the skew percentile and quickly saw that, on its own, skew doesn't offer much

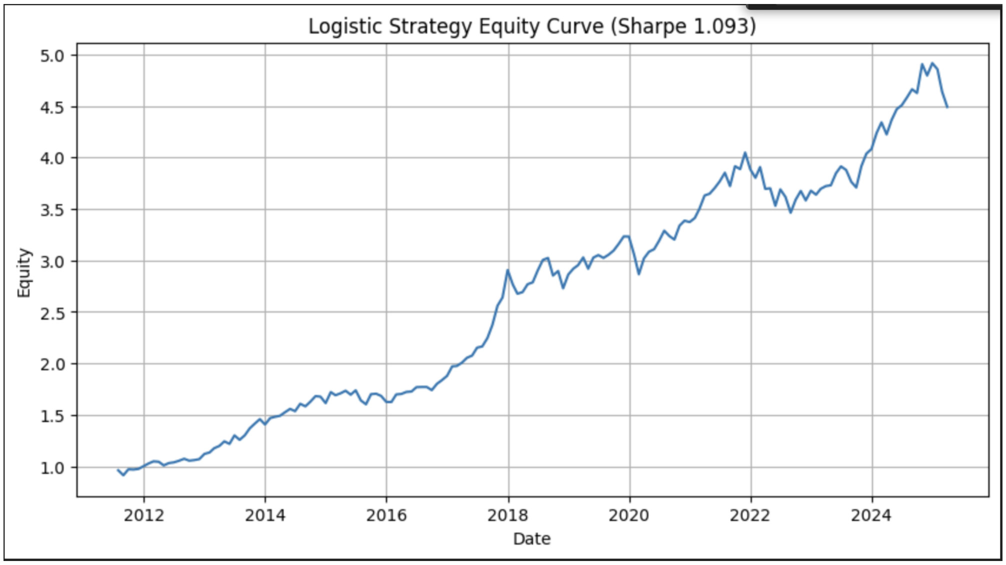

Figure 1: Logistic-regression equity curve (10 % vol, 1993–2025). Note the steady climb post-COVID.

predictive power. The Sharpe ratio for the naive filter was just 0.12, showing that this type of signal needs more context to be effective.

From there, I built a more continuous strategy by z-scoring skew and slope and allowing the position size to vary gradually. This interaction of features performed better—delivering a 0.77 Sharpe—though it remained low in CAGR since the strategy tends to stay neutral unless the signal is strong. The real breakthrough came with the logistic regression model. Using just two inputs (skew percentile and slope), the model learned a decision boundary that achieved a 1.09 Sharpe and 15% CAGR in-sample. More importantly, it held up out-of-sample from 2019 to 2025 with a 0.78 Sharpe, 7.5% CAGR, and a hit rate of 64%. That level of performance, with no overfitting or macro variables, suggests there's a real signal here.

In the end, the logistic regression strategy is the most consistent and reliable. It's simple, interpretable, and performs well even on recent data. The fact that it uses only two technical features and still beats the naive and rule-based alternatives shows the value of combining domain intuition with basic machine learning. While there's definitely room to improve (with more features, other model types, or stress testing across regimes), the results here already make a strong case that options skew and term-structure contain tradable information about near-term market direction.

If I had more time: I'd throw in VVIX, realised-vol momentum, maybe a macro regime switch, and test a gradient-boosted tree. But even this bare-bones setup shows there's real signal hiding where most pundits only see a scary SKEW headline.

Code & data: `https://github.com/vishruthanand08/iv-skew-strategy`