

Probabilistic Discrepancies between Kalshi and Traditional Options

William Walz & Rushil Jaiswal

BK Capital Management

January 2026

Abstract

This paper presents a proof-of-concept for the Flow-Risk Sentiment Pricing Engine (FRiSPe) by documenting a lead-lag relationship between prediction markets (Kalshi) and options-implied volatility (VIX). Using 222 days of overlapping unemployment market data from January 2025 to January 2026, we find that changes in Kalshi-implied probabilities statistically precede movements in the VIX by approximately two days ($p = 0.024$). CPI prediction markets, used as a control, do not exhibit similar predictive power, likely due to limited data maturity. These findings suggest that regulated prediction markets may provide a useful early signal for risk management and volatility-based trading strategies.

1 Introduction

This project looks at the lead-lag relationship between event-based prediction markets (Kalshi) and standard financial derivatives. Treating Kalshi contracts as binary options makes it possible to extract an implied probability signal to compare it to a proxy for options-implied risk-neutral pricing derived from the volatility surface of S&P 500.

2 Literature Review

Our framework builds on the view that prediction markets are efficient information aggregators. Wolfers and Zitzewitz (2004, 2006) established that binary option prices in these markets generally correspond to the mean beliefs of traders, therefore providing useful estimates of event probabilities.

Recent empirical evidence from the Kalshi exchange by Burgi et al. (2025) confirms that these prices are informative and gain accuracy as markets approach resolution, though they exhibit a “favorite-longshot bias” where lower probability outcomes are often overpriced. This bias informs our use of the Discrepancy Coefficient (λ) to distinguish between fundamental belief shifts and retail-driven noise.

To connect prediction markets with traditional finance, we incorporate “model-free” methods for extracting event probabilities from options prices, as demonstrated by Langer and Lemoine (2020). However, Gemmill and Saflekos note that while option-implied distributions reflect market sentiment, they often function as reactive rather than predictive indicators. We therefore examine whether the “skin in the game” inherent in prediction markets allows these markets to incorporate information earlier than traditional risk-neutral densities.

3 Mathematical Methodology

3.1 Probability Extraction from Prediction Markets

Prediction markets like Kalshi trade binary contracts that pay \$1 if an event occurs, otherwise \$0. Given the market price $P_K \in [0, 1]$ and assuming risk neutrality under the measure \mathbb{Q} , the price of a contract for an event E at time t with expiration T is:

$$P_K(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[\mathbb{I}_E \mid \mathcal{F}_t] \quad (1)$$

where \mathbb{I}_E is the indicator function for the event. While prediction market prices do not necessarily reflect a risk-neutral measure in the typical sense, short maturities and limited leverage justify treating them as probability estimates. In practice, we treat Kalshi prices as approximately equal to probabilities for short-dated contracts, as the discount factor $e^{-r(T-t)}$ is negligible for near-term event maturities. Kalshi offers threshold contracts that imply a cumulative distribution. We recover the probability density function (PDF) by differencing adjacent contract thresholds, which serves as a discrete approximation of the derivative of the cumulative distribution function (CDF):

$$f_K(x) \approx \frac{P_K(X > x) - P_K(X > x + \Delta x)}{\Delta x} \quad (2)$$

3.2 Risk-Neutral Density from Options Markets

To derive a comparable signal from the traditional options market, we utilize the Black-Scholes framework. The probability that an option expires in-the-money (the risk-neutral probability) is given by the $N(d_2)$ component:

$$P_{opt} = N(d_2) = N\left(\frac{\ln(S/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right) \quad (3)$$

where σ is the implied volatility (IV) extracted from the volatility surface. In our empirical analysis, we use the VIX or VIX1D as a proxy for the short-term aggregate IV to measure institutional fear or expectation of a regime shift. Notably, as VIX is an aggregate volatility index rather than a strike-specific risk-neutral density, it serves as a coarse proxy for our comparisons.

3.3 The Discrepancy Coefficient (λ)

We define the Discrepancy Coefficient, λ (referred to internally as the *Arb Gap*), as the absolute difference between the Kalshi-implied probability and the options market signal (proxied by the VIX in our empirical analysis):

$$\lambda_t = |P_{Kalshi,t} - P_{Options,t}| \quad (4)$$

In this context, λ represents a sentiment and flow discrepancy between event-specific belief shifts and institutional pricing pressures, rather than a literal arbitrage gap. A high λ indicates a divergence between retail sentiment and options market positioning.

3.4 Lead-Lag Analysis and Granger Causality

To test whether Kalshi is a leading indicator of market volatility we use a Vector Autoregression (VAR) model and perform a Granger causality test. We assess whether or not past values of the Kalshi probability signal (X) help predict future values of the VIX (Y):

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=1}^p \gamma_j X_{t-j} + \epsilon_t \quad (5)$$

The null hypothesis $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_p = 0$ is tested to evaluate whether the prediction market provides unique information prior to the options market “correcting” its volatility pricing. This result reflects predictive content in the time series rather than a formal economic causal mechanism.

4 Empirical Results

4.1 Unemployment Markets as a Leading Indicator

Our primary analysis focused on Kalshi unemployment prediction markets compared against the VIX from January 2025 to January 2026 which totaled in 222 days of overlapping data. We found that Kalshi unemployment probabilities act as an influential indicator of market volatility.

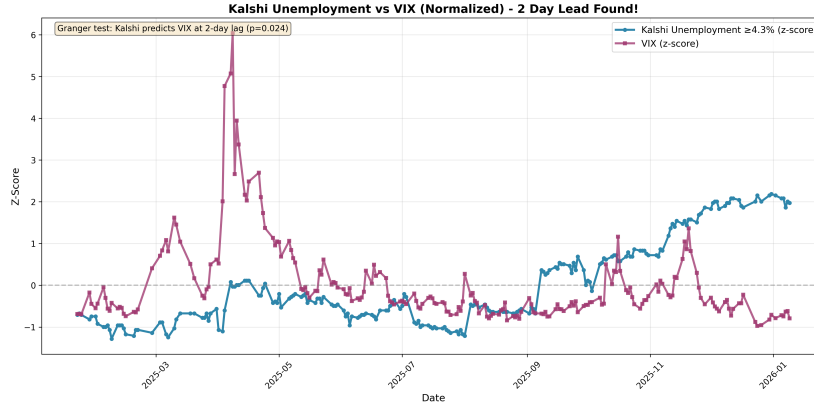


Figure 1: Normalized Time-Series Overlay: Kalshi Unemployment Sentiment vs. VIX. Note the leading movements in the Kalshi signal relative to VIX spikes.

- **Granger Causality:** The model yielded a p-value of 0.024 at a 2-day lag indicating that prediction market shifts statistically precede movements in the options-implied volatility surface.
- **Correlation Analysis:** We observed a correlation of -0.152 at a 5-day lag. This inverse relationship suggests that as the probability of a stable or improving unemployment regime increases, the aggregate market fear (VIX) decreases.

4.2 CPI Markets as a Statistical Control

To ensure the robustness of our methodology, we utilized CPI prediction markets as a control group. Due to limited historical depth (only 90 days of overlapping data), the CPI analysis did not yield significant Granger causality. This contrast is consistent with the interpretation that the signal in the unemployment markets is a function of increased data maturity and market liquidity, rather than a result of arbitrary “cherry-picking”.

5 Discussion and Interpretation

The 2-day predictive lead suggests that prediction markets respond faster to macroeconomic expectations than institutional options markets. This lends support to our hypothesis that the “skin in the game” inherent in Kalshi contracts enables more efficient information aggregation than the complex, frequently over-hedged traditional derivatives market.

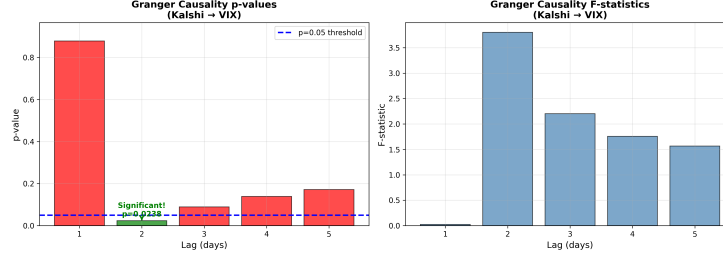


Figure 2: Granger Causality p-values by Lag. The significant p-value at Lag 2 confirms Kalshi as a leading indicator.

Notably, the negative correlation between unemployment probability and the VIX indicates that prediction markets are gauging economic stability. In this regime, stable employment forecasts act as a volatility suppressor, a signal that prediction markets reflect a predictive lead of approximately two days before it is manifested in the VIX.

6 Applications in Asset Pricing and Strategy

The empirical lead-lag relationship between Kalshi markets and options-implied volatility suggests several practical uses within the Flow-Risk Sentiment Pricing Engine (FRiSPe). The Discrepancy Coefficient (λ) measures timing differences between prediction markets and options markets and highlights periods of regime divergence.

6.1 Regime-Aware Volatility Analysis

One implication of the observed 2-day lead is its relevance for volatility-sensitive positioning decisions. Since Kalshi shifts precede VIX movements with a statistically significant p-value of 0.024, a shift in the unemployment probability signal can inform the directional bias of VIX-linked products (e.g., VXX, UVXY, SVXY).

- **Anticipatory Risk Management:** If λ widens due to a Kalshi-driven spike in unemployment, the signal can inform anticipatory hedging strategies before similar adjustments appear in options-implied volatility.
- **Signal Filtering and Sentiment Weighting:** A high λ may indicate periods where retail-driven sentiment on Kalshi is overextended relative to the volatility surface. In these instances, the discrepancy captures a short-term divergence between prediction-market sentiment and institutional pricing, helping to distinguish transient sentiment shifts from broader regime changes.

6.2 Flow-Driven Alpha and Regime Timing

Using Kalshi prices as a probability proxy provides a timing signal that is less dependent on narrative-driven media sentiment.

- **Regime-Based Timing Guidance:** Given the observed 2-day lead, these signals can be used as risk-aware inputs to identify potential regime shifts prior to their reflection in the broader volatility surface.
- **Expectation Arbitrage:** The model identifies specific windows where P_{Kalshi} and $P_{Options}$ diverge. This provides a framework to monitor the gap between retail expectations and institutional pricing pressures.

7 Conclusion and Future Work

This research successfully establishes a proof-of-concept for utilizing prediction markets as a regime-aware asset pricing tool within the FRiSPe framework.

7.1 Limitations

The main limitation of this study is the limited historical depth of regulated prediction markets. While the 222-day sample is sufficient for the current analysis, longer time series data will be required to assess performance across different economic cycles.

7.2 Next Steps

Future iterations of this research will focus on:

1. Expanding the analysis to additional macro indicators (e.g., Fed rate decisions and GDP growth).
2. Formally calculating the Arbitrage Coefficient (λ) to provide enhanced signal conditioning for volatility-based trades.
3. Conducting event-study analyses around high-impact federal releases.

References

- [1] Wolfers, J., & Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities*. Wharton, University of Pennsylvania.
- [2] Burgi, C., Deng, W., & Whelan, K. (2025). *Makers and Takers: The Economics of the Kalshi Prediction Market*. University College Dublin.
- [3] Langer, A., & Lemoine, D. (2020). *What Were the Odds? Estimating the Market's Probability of Uncertain Events*. University of Arizona Working Paper.

- [4] Gemmill, G., & Saffekos, A. *How Useful are Implied Distributions? Evidence from Stock-Index Options*. LIFFE Research.