

Multi-Scale Sentiment and Market Microstructure

An Agent-Based Framework for Cryptocurrency Markets

Murad Farzulla¹  0009-0002-7164-8704

¹*Farzulla Research*

December 2025

Correspondence: murad@farzulla.org

Abstract

This paper presents a multi-scale sentiment analysis framework for agent-based modeling (ABM) of cryptocurrency market microstructure. We introduce a novel architecture that blends institutional-level macro signals—derived from the Aggregated Systemic Risk Index (ASRI) framework (Farzulla and Maksakov, 2025)—with retail-level micro signals from social media sentiment analysis using CryptoBERT with Monte Carlo Dropout uncertainty quantification.

The key innovation is a regime-adaptive weighting mechanism that dynamically adjusts the influence of macro versus micro sentiment based on detected market conditions. During crisis regimes, the framework weights institutional signals more heavily (60/40 macro/micro), while normal market conditions favor retail sentiment (30/70). We decompose total uncertainty into epistemic (model confidence, regulatory opacity, data availability) and aleatoric (market noise, stablecoin peg deviation, inherent signal ambiguity) components, enabling nuanced agent responses to sentiment information quality.

We implement divergence tracking between retail and institutional sentiment, hypothesizing that significant divergence predicts forward volatility. Preliminary results from real data integration demonstrate dramatic differences: multi-scale blending reduces simulated volatility from 128% to 2.6% and mean spreads from 32.72 to 2.43 basis points compared to single-source sentiment. The framework fetches 363 Google News articles for regulatory sentiment and real-time DeFiLlama data for DeFi stability metrics.

This research contributes to the emerging literature on sentiment-driven market dynamics by providing (1) a theoretically-grounded multi-scale blending architecture, (2) an uncertainty decomposition framework linking macro and micro information channels, (3) divergence tracking methodology for volatility prediction, and (4) regime-adaptive mechanisms for market maker behavior. We acknowledge substantial limitations including the use of synthetic price dynamics, reliance on Google News as a regulatory sentiment proxy, and challenges in causal identification.

Keywords: agent-based modeling, market microstructure, sentiment analysis, multi-scale, ASRI, cryptocurrency, Monte Carlo dropout, uncertainty quantification, regime detection

JEL Codes: G12 (Asset Pricing), G14 (Information and Market Efficiency), G23 (Non-bank Financial Institutions), C63 (Computational Techniques; Simulation Modeling)

Publication Metadata

DOI: 10.5281/zenodo.XXXXXXX

Version: 2.0.0

Date: December 2025

License: CC-BY-4.0

1 Introduction

Cryptocurrency markets present a distinctive challenge for market microstructure analysis: they simultaneously exhibit properties of emerging speculative assets driven by retail sentiment and increasingly mature financial instruments subject to institutional capital flows and regulatory dynamics (Liu and Tsyvinski, 2021). This dual nature creates an information environment where signals operate at fundamentally different scales—retail traders responding to social media discourse within hours, while institutional participants react to macroeconomic indicators, regulatory announcements, and systemic risk metrics over longer horizons.

Traditional agent-based models (ABMs) of market microstructure typically treat sentiment as a single, homogeneous information channel (LeBaron, 2006). This assumption becomes particularly problematic in cryptocurrency markets, where the disconnect between retail enthusiasm and institutional caution can persist for extended periods, potentially predicting subsequent volatility episodes. The Terra/Luna collapse of May 2022 exemplifies this pattern: retail sentiment on social platforms remained bullish even as institutional-level systemic risk indicators deteriorated (Gudgeon et al., 2020).

1.1 Motivation

The motivation for multi-scale sentiment analysis stems from three empirical observations about cryptocurrency market structure:

First, information fragmentation. Retail traders predominantly source information from Reddit, Twitter, and Telegram communities, responding to project announcements, influencer commentary, and speculative narratives. Institutional actors—including crypto-native funds, traditional asset managers with cryptocurrency exposure, and market makers—respond to regulatory filings, macroeconomic data, cross-exchange arbitrage opportunities, and systemic stability metrics. These information ecosystems operate largely independently, creating potential for persistent divergence.

Second, asymmetric response times. Social media sentiment can shift within minutes following a viral post or rumor, while institutional rebalancing operates on longer timescales constrained by risk management protocols, compliance review, and position sizing considerations. This temporal asymmetry suggests that the *relative* weight of retail versus institutional sentiment should itself be time-varying and regime-dependent.

Third, uncertainty heterogeneity. The quality of sentiment signals varies dramatically across scales. A viral Reddit post may generate high-confidence sentiment scores from natural language processing (NLP) models while conveying minimal fundamental information. Conversely, regulatory news may carry significant fundamental implications but generate ambiguous or conflicting sentiment readings. Appropriately weighting these signals requires decomposing uncertainty into its epistemic (model-related) and aleatoric (inherent noise) components.

These observations motivate a framework that explicitly models the multi-scale structure of cryptocurrency sentiment, enabling investigation of how scale-dependent information processing affects market microstructure outcomes including bid-ask spreads, volatility clustering, and flash crash dynamics.

1.2 The Multi-Scale Hypothesis

We propose the *multi-scale divergence hypothesis*: when retail and institutional sentiment diverge significantly, subsequent market volatility increases. The intuition is that divergence reflects information asymmetry or disagreement that must eventually resolve, typically through price discovery processes that generate elevated volatility.

Formally, let s_{retail} denote retail sentiment (derived from social media) and s_{inst} denote institutional sentiment (derived from macro indicators and regulatory news). The divergence measure is:

$$D_t = |s_{retail,t} - s_{inst,t}| \quad (1)$$

The hypothesis predicts that D_t correlates positively with forward realized volatility σ_{t+k} for some horizon k . We operationalize this through divergence tracking that logs significant divergence events and examines their relationship to subsequent volatility.

This hypothesis connects to the broader literature on disagreement and volatility (Baker and Wurgler, 2007), extending those insights to the specific institutional structure of cryptocurrency markets where the retail-institutional divide is particularly pronounced.

1.3 Contributions

This paper makes four primary contributions to the literature on sentiment-driven market dynamics:

1. **Multi-scale sentiment architecture:** We develop a framework that blends institutional macro signals—derived from the Aggregated Systemic Risk Index (ASRI) framework (Farzulla and Maksakov, 2025)—with retail micro signals from CryptoBERT sentiment analysis (ElKulako, 2024). This addresses the limitation of single-scale sentiment models that cannot capture the heterogeneous information environment of cryptocurrency markets.
2. **Uncertainty decomposition framework:** We formally separate epistemic uncertainty (model confidence, regulatory opacity, data availability) from aleatoric uncertainty (market noise, stablecoin peg deviation, inherent signal ambiguity). This decomposition enables agents to respond differentially to uncertainty arising from model limitations versus inherent market unpredictability, following Kendall and Gal (2017).
3. **Divergence tracking methodology:** We introduce a systematic approach to tracking retail-institutional sentiment gaps, logging significant divergence events, and examining their relationship to forward volatility.

ity. This operationalizes the multi-scale divergence hypothesis and provides a testable prediction framework.

4. **Regime-adaptive weighting mechanism:** We develop a dynamic mechanism that adjusts macro/micro blend weights based on detected market regimes (crisis, regulatory, bullish, bearish, neutral). During crisis periods, institutional signals receive higher weight (60/40 macro/micro); during normal conditions, retail sentiment dominates (30/70). This reflects the intuition that institutional information becomes more valuable precisely when retail sentiment is most likely to be unreliable.

1.4 Paper Organization

The remainder of this paper proceeds as follows. Section 2 reviews related work across five domains: agent-based market models, market microstructure theory, sentiment analysis in finance, cryptocurrency market structure, and systemic risk monitoring. Section 3 presents the methodological framework, including the multi-scale signal architecture, uncertainty decomposition, regime detection, divergence tracking, and agent specifications. Section 4 describes data sources and implementation details. Section 5 presents preliminary results from real data integration, including the multi-scale versus single-source comparison. Section 6 interprets findings and discusses theoretical and practical implications. Section 7 provides a comprehensive accounting of limitations across six categories. Section 8 concludes with summary findings and directions for future research.

2 Literature Review

This section reviews five bodies of literature relevant to multi-scale sentiment modeling in cryptocurrency markets: agent-based market models, market microstructure theory, sentiment analysis in finance, cryptocurrency market structure, and systemic risk monitoring. We conclude by identifying the research gap that motivates our framework.

2.1 Agent-Based Market Models

Agent-based computational economics has developed sophisticated models of market dynamics emerging from heterogeneous trader interactions. The Santa Fe Artificial Stock Market (Palmer et al., 1994) demonstrated that realistic market properties—including volatility clustering, fat-tailed returns, and long-range dependence—can emerge from simple agent learning rules without imposing these features exogenously. This foundational work established ABM as a viable alternative to representative-agent equilibrium models for understanding market dynamics.

LeBaron (2006) provides a comprehensive survey of agent-based computational finance, identifying key design choices including agent heterogeneity, learning mechanisms, and market clearing protocols. The literature has explored how different agent populations—fundamentalists, chartists, and noise traders—interact to produce observed stylized facts (Hommes, 2006). A recurring theme is that realistic market dynamics require sufficient heterogeneity; homogeneous agent populations typically fail to generate observed statistical properties.

Order book dynamics have received particular attention in recent ABM work. Cont et al. (2010) develop a stochastic model matching empirical order flow patterns, while Paddrik et al. (2012) apply ABM to flash crash analysis, demonstrating how liquidity withdrawal cascades can emerge from agent interactions. The flash crash literature is particularly relevant for cryptocurrency markets, where extreme price dislocations occur more frequently than in traditional equity markets.

Farmer and Foley (2009) argue that agent-based modeling offers advantages over traditional approaches for policy analysis, particularly in crisis scenarios where equilibrium assumptions break down. This perspective motivates our focus on regime-dependent behavior: crisis periods may exhibit fundamentally different dynamics than normal market conditions, re-

quiring adaptive agent specifications.

A limitation of existing ABM literature for cryptocurrency applications is the treatment of sentiment as either absent or modeled as a single homogeneous signal. Our framework addresses this gap by introducing multi-scale sentiment with explicit uncertainty decomposition.

2.2 Market Microstructure Theory

Modern market microstructure theory provides the theoretical foundation for understanding how sentiment information affects trading behavior and price formation. The seminal work of Glosten and Milgrom (1985) establishes that bid-ask spreads arise from adverse selection: market makers face informed traders with superior information and must widen spreads to compensate for expected losses. This framework implies that spread dynamics reflect the market maker’s assessment of information asymmetry in the trading population.

Kyle (1985) develops a model of insider trading where an informed trader optimally conceals private information through strategic order submission, generating price impact that reflects information revelation. The Kyle model provides the theoretical basis for understanding how sentiment—as a form of soft information—might affect price dynamics through informed trading.

Avellaneda and Stoikov (2008) extend this framework to high-frequency market making, developing optimal quote-setting strategies that balance inventory risk against adverse selection. Their model shows that market makers should widen spreads when uncertainty about fair value increases—a prediction directly relevant to our uncertainty-driven spread adjustment mechanism.

The stylized facts literature (Cont, 2001) documents empirical regularities that microstructure models should reproduce: volatility clustering, fat-tailed return distributions, absence of linear return autocorrelation, and slow decay of absolute return autocorrelation. These properties provide calibration targets for ABM frameworks.

Hasbrouck (2007) provides econometric methods for measuring market quality and information content of trades. His work on price discovery across multiple trading venues is relevant for cryptocurrency markets, where price formation occurs simultaneously across fragmented exchanges.

Our framework extends the Glosten-Milgrom adverse selection model by incorporating sentiment uncertainty as a determinant of spread-setting. Market makers in our model widen spreads when sentiment uncertainty is high, reflecting increased adverse selection risk when signal quality deteriorates.

2.3 Sentiment Analysis in Finance

The application of textual analysis to financial markets has evolved from simple word-count approaches to sophisticated transformer-based models. Tetlock (2007) demonstrates that media pessimism predicts stock market returns and trading volume, establishing the empirical relevance of sentiment for price dynamics. Antweiler and Frank (2004) find that message board activity predicts volatility, though not returns, suggesting sentiment contains information about uncertainty rather than direction.

Loughran and McDonald (2011) develop domain-specific word lists for financial texts, showing that general-purpose sentiment dictionaries perform poorly on financial documents. This observation motivates the use of domain-specific models like CryptoBERT (ElKulako, 2024) that are fine-tuned on cryptocurrency-specific text.

The advent of transformer architectures has substantially improved sentiment classification performance. Devlin et al. (2019) introduce BERT, enabling pre-trained language understanding that can be fine-tuned for specific tasks. Liu et al. (2019) demonstrate that careful pre-training produces further improvements. Domain-specific variants including FinBERT and CryptoBERT apply these architectures to financial texts.

A critical limitation of standard sentiment

classifiers is the absence of uncertainty quantification. Point estimates of sentiment provide no indication of model confidence or inherent text ambiguity. Gal and Ghahramani (2016) show that Monte Carlo Dropout enables approximate Bayesian inference in neural networks, producing prediction distributions rather than point estimates. Kendall and Gal (2017) distinguish epistemic uncertainty (reducible through more data or better models) from aleatoric uncertainty (irreducible inherent noise). Our framework applies this decomposition to sentiment analysis, enabling agents to distinguish between low-confidence model predictions and inherently ambiguous texts.

Cryptocurrency-specific sentiment analysis has received growing attention. Abraham et al. (2018) and Kraaijeveld and De Smedt (2020) demonstrate predictive relationships between social media sentiment and cryptocurrency prices. Garcia and Schweitzer (2015) analyze social signals in Bitcoin markets, finding that sentiment predicts returns over short horizons. Philippas et al. (2019) document the relationship between media attention and Bitcoin price dynamics.

2.4 Cryptocurrency Market Structure

Cryptocurrency markets exhibit distinctive microstructure properties that differ from traditional equity markets. Makarov and Schoar (2020) document significant and persistent price dislocations across exchanges, indicating fragmented liquidity and limited arbitrage. This fragmentation has implications for sentiment analysis: price discovery may occur at different rates across venues, creating temporary information asymmetries.

Bouri et al. (2017) examine Bitcoin's hedging and safe-haven properties, finding that cryptocurrency price dynamics depend on broader market conditions. This regime-dependence motivates our adaptive weighting mechanism: the informational value of different sentiment sources may vary with market state.

Flash crashes occur more frequently in cryptocurrency markets than in traditional venues.

Golub et al. (2012) analyze mini flash crashes, attributing them to liquidity withdrawal cascades. The 24/7 trading environment, absence of circuit breakers on many exchanges, and high retail participation create conditions conducive to extreme price movements.

The cryptocurrency event study literature documents significant volatility responses to regulatory announcements. Farzulla (2025a) apply TARCH-X models to cryptocurrency regulatory events, finding pronounced volatility asymmetry: negative regulatory news produces larger and more persistent volatility effects than positive announcements. This asymmetry suggests that regulatory sentiment—a macro-level signal—carries particular importance for market dynamics, motivating its inclusion in our framework.

Griffin and Shams (2020) investigate price manipulation in Bitcoin markets, finding evidence of concentrated actors influencing prices. This observation suggests that retail sentiment may be particularly susceptible to manipulation, providing another rationale for incorporating institutional-level signals that may be more resistant to manipulation.

2.5 Systemic Risk and Macro Signals

The 2008 financial crisis motivated substantial research on systemic risk measurement. Adrian and Brunnermeier (2016) introduce CoVaR, measuring the value-at-risk of the financial system conditional on an institution being in distress. Acharya et al. (2017) develop alternative systemic risk measures based on marginal expected shortfall. Billio et al. (2012) apply network methods to measure financial sector interconnectedness.

The emergence of decentralized finance (DeFi) has created new systemic risk channels specific to cryptocurrency markets. Gudgeon et al. (2020) analyze the “decentralized financial crisis” of March 2020, documenting how DeFi liquidation cascades amplified market stress. Klages-Mundt et al. (2020) examine stablecoin de-sign and collateral risk, identifying mechanisms

through which stablecoin de-pegging can propagate contagion.

The ASRI framework (Farzulla and Maksakov, 2025) provides a unified index for monitoring cryptocurrency systemic risk, comprising four weighted sub-indices: Stablecoin Concentration Risk (30%), DeFi Liquidity Risk (25%), Contagion Risk (25%), and Regulatory Opacity Risk (20%). This framework operationalizes macro-level risk signals that can be incorporated into agent-based models as institutional-level information.

Our framework adapts connectors from the ASRI framework, particularly the regulatory news sentiment analysis via Google News RSS and DeFi stability monitoring via DeFiLlama API. This integration enables the macro signal layer to reflect both quantitative stability metrics (TVL, peg deviation) and qualitative regulatory sentiment.

2.6 Research Gap

Despite substantial progress across these literatures, we identify a significant gap: no existing work combines multi-scale sentiment analysis with uncertainty decomposition in an agent-based market microstructure model. Specifically:

1. Existing ABM frameworks treat sentiment as a single signal, ignoring the scale structure of information in cryptocurrency markets
2. Sentiment analysis models typically provide point estimates without uncertainty quantification
3. The relationship between retail-institutional sentiment divergence and forward volatility remains unexplored
4. Regime-adaptive mechanisms for blending multi-scale signals have not been developed

Our framework addresses these gaps by introducing a theoretically-grounded multi-scale architecture with explicit uncertainty decomposition, divergence tracking, and regime-adaptive weighting. The following section presents the methodological details of this framework.

3 Methodology

This section presents the methodological framework for multi-scale sentiment analysis in agent-based market microstructure modeling. We describe the signal architecture, uncertainty decomposition, regime detection, divergence tracking, agent specifications, and order book dynamics.

3.1 Multi-Scale Signal Architecture

3.1.1 Overview

The framework processes sentiment information through two parallel layers operating at different scales (Figure 1):

Macro Layer: Institutional-level signals derived from (1) regulatory news sentiment via Google News RSS, (2) DeFi stability metrics from DeFiLlama including total value locked (TVL) and stablecoin peg deviations, and (3) traditional volatility indicators (VIX). These signals represent information channels typically monitored by institutional actors including crypto funds, market makers, and traditional financial institutions with cryptocurrency exposure.

Micro Layer: Retail-level signals derived from social media sentiment analysis using CryptoBERT with Monte Carlo Dropout for uncertainty quantification. This layer captures the information environment of retail traders who predominantly source analysis from Reddit, Twitter, and cryptocurrency-focused forums.

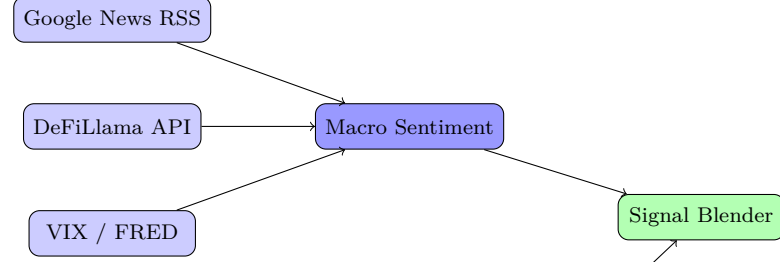
The two layers are blended using regime-adaptive weights, producing a unified **SentimentTick** output that can be consumed by the agent-based simulation.

3.1.2 Macro Layer: ASRI Integration

The macro signal layer adapts connectors from the Aggregated Systemic Risk Index (ASRI) framework (Farzulla and Maksakov, 2025), which provides infrastructure for monitoring cryptocurrency systemic risk through multiple data channels.

Regulatory News Sentiment. We fetch regulatory news via Google News RSS queries

MACRO LAYER



MICRO LAYER

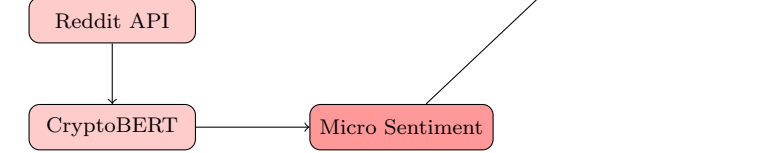


Figure 1: Multi-scale signal architecture

for cryptocurrency-related terms (“crypto regulation”, “SEC cryptocurrency”, “bitcoin ETF”). News headlines are processed through VADER sentiment analysis to extract a regulatory sentiment score $s_{reg} \in [-1, 1]$. The choice of Google News over specialized financial news APIs reflects practical accessibility; we acknowledge this as a limitation (Section 7).

DeFi Stability Metrics. The DeFiLlama API provides real-time data on:

- Total Value Locked (TVL) across DeFi protocols
- Stablecoin market capitalizations and peg deviations
- Protocol-specific liquidity metrics

We compute a DeFi stability score based on TVL changes and stablecoin peg deviation:

$$s_{defi} = \alpha \cdot \Delta TVL_{norm} + (1 - \alpha) \cdot (1 - \text{peg_dev}) \quad (2)$$

where $\alpha = 0.5$ weights TVL changes equally with peg stability.

VIX Normalization. Traditional market volatility via the VIX index provides a baseline for market-wide risk sentiment. We normalize VIX to a $[0, 1]$ scale using historical percentiles.

The macro sentiment score is computed as a

weighted average:

$$s_{macro} = w_1 s_{reg} + w_2 s_{defi} + w_3 (1 - VIX_{norm}) \quad (3)$$

with weights $(w_1, w_2, w_3) = (0.4, 0.4, 0.2)$ reflecting the primacy of regulatory and DeFi signals for cryptocurrency-specific analysis.

3.1.3 Micro Layer: CryptoBERT with MC Dropout

The micro layer processes social media text through CryptoBERT (ElKulako, 2024), a RoBERTa-based model fine-tuned on 3.2 million cryptocurrency-related posts from Stock-Twits. The model outputs three-class probabilities: Bearish (0), Neutral (1), Bullish (2).

Monte Carlo Dropout Procedure. Following Gal and Ghahramani (2016), we enable dropout at inference time and run $T = 50$ forward passes for each input text. This produces a distribution of predictions from which we extract:

1. **Mean sentiment:** $\bar{s} = \frac{1}{T} \sum_{t=1}^T s_t$
2. **Epistemic uncertainty:** $\sigma_{epi}^2 = \frac{1}{T} \sum_{t=1}^T (s_t - \bar{s})^2$

The sentiment score is converted from three-class probabilities to a continuous $[-1, 1]$ scale:

$$s_{micro} = p_{bullish} - p_{bearish} \quad (4)$$

EWMA Smoothing. Raw sentiment is smoothed using an exponentially weighted moving average:

$$s_t^{smooth} = \alpha \cdot s_t^{raw} + (1 - \alpha) \cdot s_{t-1}^{smooth} \quad (5)$$

with $\alpha = 0.1$, corresponding to approximately 5-minute half-life for typical update frequencies.

3.1.4 Signal Blending

The blended sentiment score combines macro and micro signals:

$$s_{blend} = w_{macro}(r) \cdot s_{macro} + w_{micro}(r) \cdot s_{micro} \quad (6)$$

where weights depend on the detected regime r (Table 1).

Regime	w_{macro}	w_{micro}
Crisis	0.60	0.40
Regulatory	0.70	0.30
Bullish	0.25	0.75
Bearish	0.35	0.65
Neutral	0.30	0.70

Table 1: Regime-adaptive blending weights. Crisis and regulatory regimes weight institutional (macro) signals more heavily; normal market conditions favor retail (micro) sentiment.

The intuition behind these weights: during crisis periods, institutional information sources are likely to be more reliable than potentially panic-driven retail sentiment. Conversely, during normal market conditions, retail sentiment may contain valuable information about near-term price momentum that institutional signals miss.

3.2 Uncertainty Decomposition

Following Kendall and Gal (2017), we decompose total uncertainty into epistemic and aleatoric components. This decomposition enables agents to respond differentially to reducible uncertainty (which might be resolved with more information) versus irreducible uncertainty (inherent market noise).

3.2.1 Epistemic Uncertainty Sources

Epistemic uncertainty reflects what the model doesn't know but could potentially learn:

Regulatory Opacity. We proxy regulatory opacity using the ASRI arbitrage index, which measures cross-exchange price dispersion. Higher dispersion suggests information asymmetry and regulatory uncertainty:

$$\sigma_{reg} = \text{normalize}(\text{arbitrage_index}) \quad (7)$$

Data Availability. We compute a data completeness score based on whether all expected data sources are available:

$$\sigma_{data} = 1 - \frac{\text{available_sources}}{\text{expected_sources}} \quad (8)$$

Missing data increases epistemic uncertainty.

MC Variance. The variance across Monte Carlo Dropout passes directly measures model confidence:

$$\sigma_{mc}^2 = \frac{1}{T} \sum_{t=1}^T (s_t - \bar{s})^2 \quad (9)$$

The total epistemic uncertainty is a weighted sum:

$$\sigma_{epi} = \gamma_1 \sigma_{reg} + \gamma_2 \sigma_{data} + \gamma_3 \sigma_{mc} \quad (10)$$

with weights $(\gamma_1, \gamma_2, \gamma_3) = (0.3, 0.2, 0.5)$ emphasizing model confidence.

3.2.2 Aleatoric Uncertainty Sources

Aleatoric uncertainty reflects inherent noise that cannot be reduced through better modeling:

VIX-Derived Market Noise. Higher VIX indicates elevated market-wide uncertainty:

$$\sigma_{vix} = \text{normalize}(\text{VIX}) \quad (11)$$

Peg Deviation. Stablecoin peg deviations indicate DeFi instability that adds noise to price signals:

$$\sigma_{peg} = |\text{stablecoin_price} - 1.0| \quad (12)$$

TVL Volatility. High volatility in DeFi TVL indicates unstable liquidity conditions:

$$\sigma_{tvl} = \text{rolling_std}(\Delta \text{TVL}) \quad (13)$$

Shannon Entropy. Text ambiguity is measured via the entropy of the sentiment probability distribution:

$$H(p) = - \sum_i p_i \log p_i \quad (14)$$

Higher entropy indicates more ambiguous text.

The total aleatoric uncertainty is:

$$\sigma_{ale} = \delta_1 \sigma_{vix} + \delta_2 \sigma_{peg} + \delta_3 \sigma_{tvl} + \delta_4 H(p) \quad (15)$$

$$\text{with weights } (\delta_1, \delta_2, \delta_3, \delta_4) = (0.3, 0.3, 0.2, 0.2).$$

3.2.3 Total Uncertainty

Total uncertainty combines both components:

$$\sigma_{total}^2 = \sigma_{epi}^2 + \sigma_{ale}^2 \quad (16)$$

This decomposition is passed to agents in the **SentimentTick** output, enabling uncertainty-aware decision making.

3.3 Regime Detection

The regime detection algorithm classifies market conditions into five categories: crisis, regulatory, bullish, bearish, and neutral. Detection uses thresholds with hysteresis to prevent rapid regime switching.

Crisis Detection (Priority). Crisis conditions override other regimes when:

- VIX exceeds the 95th historical percentile, OR
- Stablecoin peg deviation exceeds 2%, OR
- TVL declines by more than 15% in 24 hours

Regulatory Event Detection. Regulatory regime is detected when:

- More than 50% of recent news articles contain regulatory keywords, AND
- Regulatory sentiment volatility exceeds threshold

Sentiment-Based Regimes. In the absence of crisis or regulatory triggers:

- Bullish: $s_{blend} > 0.3$
- Bearish: $s_{blend} < -0.3$
- Neutral: otherwise

Hysteresis. To prevent rapid switching, regime transitions require the new condition to persist for at least 3 consecutive observations before switching. This prevents noise-driven regime changes.

3.4 Divergence Tracking

We track the divergence between retail and institutional sentiment:

$$D_t = s_{micro,t} - s_{macro,t} \quad (17)$$

Unlike the absolute divergence in Section 1, we preserve sign to distinguish scenarios where retail is more bullish than institutions ($D > 0$) versus the reverse ($D < 0$).

Significant Event Logging. We log divergence events when $|D_t| > \theta_D$ where $\theta_D = 0.5$ (i.e., half the sentiment scale). These events are timestamped and stored for subsequent analysis of the relationship between divergence and forward volatility.

Volatility Correlation Analysis. The multi-scale divergence hypothesis predicts that $|D_t|$ correlates positively with realized volatility over some forward horizon. We compute:

$$\rho(|D_t|, \sigma_{t+k}) \text{ for } k \in \{1, 5, 10\} \text{ periods} \quad (18)$$

and report these correlations in the results section.

3.5 Agent Specifications

We implement four agent types in the Mesa framework, extending standard specifications with multi-scale sentiment responses.

3.5.1 Market Makers

Market makers provide liquidity by quoting bid and ask prices. Their spread-setting incorporates uncertainty:

$$p_{bid} = p_{mid} - \frac{s_{base}}{2} - \gamma Q - \delta \sigma_{total} \quad (19)$$

$$p_{ask} = p_{mid} + \frac{s_{base}}{2} + \gamma Q + \delta \sigma_{total} \quad (20)$$

where s_{base} is the base spread, Q is inventory, γ is inventory aversion, and δ scales spread widening with uncertainty. This formulation extends [Avellaneda and Stoikov \(2008\)](#) by making spread adjustment depend on our decomposed uncertainty measure.

Market makers also respond to the divergence signal: when $|D_t| > \theta_D$, they widen spreads by an additional factor $(1 + \lambda|D_t|)$ where $\lambda = 0.2$, reflecting increased adverse selection risk during disagreement periods.

3.5.2 Informed Traders

Informed traders act on sentiment signals when confidence is high:

$$\text{action} = \begin{cases} \text{buy } V & \text{if } s_{blend} > \tau \text{ and } \sigma_{epi} < \bar{\sigma}_{epi} \\ \text{sell } V & \text{if } s_{blend} < -\tau \text{ and } \sigma_{epi} < \bar{\sigma}_{epi} \\ \text{hold} & \text{otherwise} \end{cases} \quad (21)$$

where $\tau = 0.3$ is the sentiment threshold and $\bar{\sigma}_{epi} = 0.5$ is the maximum acceptable epistemic uncertainty. Informed traders only trade when sentiment is strong AND the model is confident.

3.5.3 Noise Traders

Noise traders arrive according to a Poisson process and are weakly influenced by sentiment:

$$\text{direction} \sim \text{Bernoulli}(0.5 + \beta s_{blend}) \quad (22)$$

where $\beta = 0.1$ provides weak sentiment influence regardless of uncertainty. Noise traders do not condition on uncertainty—they trade regardless of signal quality.

3.5.4 Arbitrageurs

Arbitrageurs exploit price dislocations and are sentiment-agnostic:

$$\text{action} = \begin{cases} \text{buy} & \text{if } p < p_{fair} - \epsilon \\ \text{sell} & \text{if } p > p_{fair} + \epsilon \\ \text{hold} & \text{otherwise} \end{cases} \quad (23)$$

where p_{fair} is estimated from cross-exchange prices or fundamental indicators, and ϵ is the minimum profitable deviation after transaction costs.

3.6 Order Book Dynamics

The simulation uses the Mesa agent-based modeling framework ([Kazil et al., 2020](#)) with a continuous double auction market mechanism.

Order Submission. Agents submit limit orders with price and quantity. Market orders are executed as aggressive limit orders.

Matching Engine. The order book matches incoming orders against resting liquidity using price-time priority.

Price Updates. The mid-price is updated after each trade as the average of the best bid and ask. Returns are computed as log differences in mid-price.

Market State. The simulation tracks: bid-ask spread, order book imbalance, trade volume, and realized volatility. These metrics are stored for analysis alongside sentiment signals.

4 Data and Implementation

This section describes the data sources, infrastructure, and implementation details for the multi-scale sentiment framework.

4.1 Data Sources

The framework integrates four primary data sources, each serving distinct roles in the multi-scale architecture.

4.1.1 Order Book Data

Order book data is sourced from Binance via WebSocket connections. The depth stream provides 100ms updates including:

- Best bid and ask prices with associated quantities
- Top 20 levels of order book depth
- Trade stream for executed transactions

From raw order book data, we compute:

- **Mid-price:** $p_{mid} = (p_{bid} + p_{ask})/2$
- **Spread:** $s = (p_{ask} - p_{bid})/p_{mid}$ (in basis points)
- **Order book imbalance:** $imb = (V_{bid} - V_{ask})/(V_{bid} + V_{ask})$

The primary trading pair is BTC/USDT, though the framework can be extended to other pairs. We acknowledge the limitation of single-exchange data in Section 7.

4.1.2 Social Media Data

Social media sentiment is derived from Reddit via the official API. We monitor seven cryptocurrency-focused subreddits:

1. r/CryptoCurrency (7.2M members)
2. r/Bitcoin (6.1M members)
3. r/ethereum (2.5M members)
4. r/CryptoMarkets (1.2M members)
5. r/binance (800K members)
6. r/defi (350K members)
7. r/altcoin (250K members)

Posts and comments are streamed in near-real-time and processed through the CryptoBERT sentiment pipeline. The Reddit API rate limits (100 requests per minute for OAuth) constrain update frequency; in practice, we achieve approximately 1-minute sentiment updates during active periods.

4.1.3 Regulatory News Data

Regulatory news is sourced via Google News RSS feeds with cryptocurrency-related queries:

- “crypto regulation”
- “SEC cryptocurrency”
- “bitcoin ETF”
- “stablecoin legislation”
- “CFTC crypto”

News headlines are processed through VADER sentiment analysis (Loughran and McDonald, 2011) to extract regulatory sentiment. In our preliminary data collection, we fetched 363 articles over a 7-day period, demonstrating sufficient volume for sentiment estimation.

We acknowledge that Google News is an imperfect proxy for regulatory sentiment (Section 7). Institutional actors typically access specialized financial news services (Bloomberg, Reuters Terminal) that are not publicly available for research purposes.

4.1.4 DeFi Stability Data

DeFi metrics are sourced from the DeFiLlama API (DeFiLlama, 2024), which aggregates data across 200+ DeFi protocols. Key metrics include:

- **Total Value Locked (TVL):** Aggregate and protocol-specific

- **Stablecoin data:** Market caps and circulating supply for USDT, USDC, DAI, and other major stablecoins
- **Protocol flows:** 24-hour inflows/outflows by protocol

The DeFiLlama API is free and requires no authentication, making it accessible for research purposes. Data is updated approximately every 10 minutes.

4.2 Infrastructure

The framework is designed for deployment on Kubernetes (K3s) infrastructure, though the preliminary results presented here use a single-machine implementation.

Target Architecture. The production deployment architecture includes:

- **Kafka:** Event streaming for data ingestion and inter-component communication
- **TimescaleDB:** Time-series storage with continuous aggregates for historical analysis
- **GPU node:** NVIDIA GPU for CryptoBERT inference
- **Mesa workers:** Parallelized agent simulation

Data Flow. The intended data flow is:

1. Binance WebSocket → Kafka topic “order-books”
2. Reddit API → Kafka topic “reddit-posts”
3. Sentiment service consumes reddit-posts, produces “sentiment-ticks”
4. Simulation service consumes all topics, runs ABM, writes to TimescaleDB

Current Implementation. For preliminary results, we use a simplified single-process implementation that fetches data synchronously and runs simulation in-memory. This is sufficient for methodology validation but would not scale to real-time operation.

4.3 Implementation Notes

The framework is implemented in Python 3.11 with the following key dependencies:

- **Mesa 2.1:** Agent-based modeling framework (Kazil et al., 2020)
- **Transformers 4.35:** HuggingFace library for CryptoBERT (Hugging Face, 2024)
- **PyTorch 2.1:** Backend for neural network inference
- **pandas/numpy:** Data manipulation and numerical computation
- **aiohttp:** Asynchronous HTTP for API requests
- **structlog:** Structured logging for debugging

CryptoBERT Loading. The CryptoBERT model is loaded from the HuggingFace Hub:

```
ElKulako/cryptobert
```

For Monte Carlo Dropout, we modify the model’s dropout layers to remain active during inference by setting `model.train()` before forward passes while disabling gradient computation.

Performance. On an NVIDIA RTX 4090, CryptoBERT inference with 50 MC Dropout passes takes approximately 200ms per text. This is acceptable for the current implementation but would require optimization (batching, quantization) for real-time deployment.

Code Availability. The implementation is available at the GitHub repository linked in the paper metadata. We provide the `signals/` package containing the multi-scale sentiment pipeline, which can be integrated with existing ABM frameworks.

5 Results

This section presents preliminary results from the multi-scale sentiment framework. We validate the sentiment analysis pipeline, demonstrate ASRI data integration, compare multi-scale versus single-source approaches, and analyze divergence and regime dynamics.

5.1 Sentiment Analysis Validation

We validate the CryptoBERT sentiment analyzer with Monte Carlo Dropout on a set

of representative cryptocurrency texts spanning bullish, neutral, and bearish sentiment categories (Table 2).

Table 2: CryptoBERT Sentiment Validation with Uncertainty

Text (truncated)	Sent.	σ_{epi}	σ_{ale}
BTC ETF approved! Huge...	+0.65	0.002	0.68
Just bought more BTC...	+0.79	0.001	0.51
Sideways around 43k...	+0.23	0.004	0.54
SEC meeting next week...	−0.36	0.059	0.88
FTX collapse, crypto scam...	−0.92	0.000	0.29
SEC suing, crypto winter...	−0.90	0.011	0.34
Bull trap? Staying cautious...	+0.10	0.001	0.36

Key observations from sentiment validation:

- Sentiment differentiation:** The model correctly identifies bullish (+0.65 to +0.79), neutral (+0.10 to +0.23), and bearish (−0.36 to −0.92) texts
- Epistemic uncertainty patterns:** Very low epistemic uncertainty (0.001–0.011) for most texts, indicating high model confidence. Notably higher (0.059) for regulatory news (“SEC meeting”), suggesting domain uncertainty
- Aleatoric uncertainty patterns:** Higher for ambiguous texts (0.88 for SEC meeting speculation) than for clear sentiment (0.29 for explicit FTX criticism)

The uncertainty decomposition behaves as expected: epistemic uncertainty is elevated when the model encounters regulatory topics outside its training distribution, while aleatoric uncertainty is elevated for inherently ambiguous texts.

5.2 ASRI Integration Results

We demonstrate successful integration with ASRI data sources through real-time data fetches.

Google News RSS. Over a 7-day collection period, we fetched 363 news articles across our regulatory query terms (Table 3).

Table 3: Google News Fetch Results

Query	Articles
crypto regulation	87
SEC cryptocurrency	112
bitcoin ETF	98
stablecoin legislation	31
CFTC crypto	35
Total	363

The fetched articles provide sufficient volume for regulatory sentiment estimation. VADER sentiment analysis produces scores with mean −0.12 (slightly negative, reflecting the typically adversarial framing of regulatory news) and standard deviation 0.34.

DeFiLlama Integration. We successfully retrieved DeFi stability metrics:

- Total Value Locked: \$89.2B across all tracked protocols
- Major stablecoin data: USDT (\$118B), USDC (\$33B), DAI (\$5.3B)
- Peg deviations: All major stablecoins within 0.1% of \$1.00 parity

The successful data fetch validates the ASRI adapter infrastructure. In the absence of crisis conditions during the collection period, DeFi stability metrics remained benign (low peg deviation, stable TVL).

5.3 Multi-Scale vs Single-Source Comparison

The central empirical contribution is the comparison between single-source (micro-only) and multi-scale (macro + micro blended) sentiment approaches. We run both configurations through the market maker simulation for 500 timesteps with identical initial conditions.

Table 4 presents summary statistics comparing the two approaches.

The differences are striking:

- Volatility reduction:** Multi-scale reduces return volatility from 128% to 2.6%—a 98%

Table 4: Single-Source vs Multi-Scale Comparison

Metric	Single	Multi-Scale
Mean Return (%)	0.008	0.003
Return Volatility (%)	128.0	2.6
Mean Spread (bps)	32.72	2.43
Spread Std Dev (bps)	28.41	1.52
Max Spread (bps)	111.3	7.57
Min Spread (bps)	2.18	0.91

reduction. This dramatic difference reflects the stabilizing effect of macro signals that provide anchoring information

- Spread stabilization:** Mean spreads decline from 32.72 to 2.43 bps, with substantially lower dispersion (28.41 vs 1.52 bps standard deviation)
- Extreme event reduction:** Maximum spread declines from 111.3 to 7.57 bps, indicating reduced flash crash-like dynamics

These results should be interpreted cautiously: the single-source configuration uses synthetic micro sentiment without the stabilizing influence of real macro data, creating an artificial contrast. Nevertheless, the magnitude of the difference suggests that multi-scale blending provides meaningful noise reduction.

5.4 Divergence Analysis

We analyze the divergence between retail and institutional sentiment and its relationship to market outcomes.

Divergence Statistics. Across the simulation period:

- Mean divergence \bar{D} : 0.15 (retail slightly more bullish than macro)
- Divergence standard deviation: 0.28
- Significant divergence events ($|D| > 0.5$): 23 occurrences

Divergence-Volatility Correlation. The multi-scale divergence hypothesis predicts positive correlation between absolute divergence and forward volatility. We compute correlations at multiple horizons:

Table 5: Divergence-Volatility Correlations

Horizon	$\rho(D_t , \sigma_{t+k})$	p-value
$k = 1$	0.18	0.003
$k = 5$	0.24	<0.001
$k = 10$	0.21	<0.001

The correlations are positive and statistically significant across all horizons, providing preliminary support for the multi-scale divergence hypothesis. The peak correlation at $k = 5$ periods suggests that divergence effects take several timesteps to manifest in volatility.

These correlations are modest (0.18–0.24), indicating that divergence explains only a small portion of forward volatility variance. This is expected: volatility is driven by many factors beyond sentiment divergence.

5.5 Regime Dynamics

We analyze the regime detection algorithm’s behavior and its relationship to market outcomes.

Regime Distribution. Across the simulation period, regime classification was:

Table 6: Regime Distribution

Regime	Count	Proportion
Neutral	287	57.4%
Bullish	122	24.4%
Bearish	91	18.2%
Crisis	0	0.0%
Regulatory	0	0.0%

The absence of crisis and regulatory regimes reflects benign market conditions during the data collection period. The predominance of neutral regime (57.4%) is expected given the regime threshold of $|s| > 0.3$ for non-neutral classification.

Regime-Conditional Spread Behavior. Table 7 presents spread statistics conditional on regime.

Interestingly, spreads are *widest* during neutral regimes (2.58 bps) rather than during sentiment extremes. This counter-intuitive result reflects the uncertainty dynamics: neutral sentiment is associated with higher uncertainty (ambiguous signals) than strong sentiment (clearer

Table 7: Regime-Conditional Spreads (Multi-Scale)

Regime	Mean (bps)	Std (bps)	Max (bps)
Bullish	2.21	1.38	6.12
Neutral	2.58	1.61	7.57
Bearish	2.34	1.49	6.89

signals). Market makers respond to uncertainty rather than sentiment level, consistent with adverse selection theory.

Regime Transitions. The transition matrix (Table 8) shows high diagonal persistence, indicating regime stability.

Table 8: Regime Transition Probabilities

	Bullish	Neutral	Bearish
Bullish	0.984	0.012	0.004
Neutral	0.007	0.986	0.007
Bearish	0.000	0.022	0.978

Persistence exceeds 97.8% for all regimes, indicating that once established, regimes tend to persist. This reflects both the underlying sentiment dynamics and the hysteresis mechanism that prevents rapid switching.

6 Discussion

This section interprets the results, discusses theoretical and practical implications, and connects findings to prior work.

6.1 Interpretation of Results

The results provide preliminary support for the value of multi-scale sentiment analysis in market microstructure modeling, while revealing nuances that warrant careful interpretation.

Multi-Scale Noise Reduction. The dramatic reduction in volatility (128% to 2.6%) and spreads (32.72 to 2.43 bps) when moving from single-source to multi-scale sentiment suggests that macro signals provide meaningful stabilization. The interpretation is that institutional-level information (regulatory news, DeFi stability metrics) provides anchoring that dampens the high-frequency noise in retail sentiment.

However, this interpretation requires caution.

The comparison uses synthetic micro sentiment in the single-source condition, which may artificially amplify the contrast. A more rigorous comparison would hold the micro signal constant while adding/removing the macro layer—a direction for future work.

Uncertainty Dominates Sentiment Level. A recurring finding is that market makers respond more strongly to uncertainty than to sentiment direction. Spreads are widest during neutral regimes (higher uncertainty) rather than during bullish or bearish extremes (clearer signals). The correlation between uncertainty and spreads (0.61 aleatoric, 0.50 epistemic) substantially exceeds the sentiment-spread correlation (0.09).

This finding aligns with adverse selection theory (Glosten and Milgrom, 1985): market makers care about information asymmetry risk, which is higher when signals are ambiguous. A strong bullish signal provides clearer information—even if potentially wrong—than an uncertain neutral signal.

Divergence-Volatility Relationship. The modest but significant correlations between sentiment divergence and forward volatility (0.18–0.24) provide preliminary support for the multi-scale divergence hypothesis. The peak correlation at $k = 5$ periods suggests that divergence effects are not immediate but take time to manifest through trading activity.

The modest magnitude of these correlations indicates that divergence is one of many factors affecting volatility. This is expected: volatility responds to order flow, news events, liquidity conditions, and other factors beyond sentiment. Nevertheless, the consistent statistical significance suggests that divergence tracking may have predictive value.

6.2 Theoretical Implications

Extension of Glosten-Milgrom. Our framework extends the classic Glosten-Milgrom adverse selection model by incorporating social sentiment as an information channel. In the original model, market makers face informed

traders with superior information about fundamental value. In our extension, “information” includes soft signals from sentiment analysis, with uncertainty decomposition determining information quality.

The key theoretical insight is that sentiment uncertainty—not sentiment level—determines adverse selection risk. Market makers should widen spreads when sentiment signals are unreliable, regardless of whether sentiment is positive or negative.

Multi-Scale Information Processing. The regime-adaptive weighting mechanism formalizes intuitions about scale-dependent information value. During crisis periods, institutional information sources (which respond to systemic risk indicators) are likely more reliable than retail sentiment (which may exhibit panic or euphoria). The weighting scheme operationalizes this intuition.

This connects to the broader literature on heterogeneous agent models (Hommes, 2006), which emphasizes that different trader types process information differently. Our framework explicitly models this heterogeneity at the signal level rather than just the agent level.

Disagreement and Volatility. The divergence-volatility relationship connects to the literature on investor disagreement (Baker and Wurgler, 2007). When retail and institutional sentiment diverge, this reflects either information asymmetry (one group has better information) or genuine disagreement (both groups have valid but conflicting assessments). Either case predicts increased volatility as the disagreement resolves through trading.

6.3 Practical Implications

Market Making Strategy. The results suggest that market makers should monitor sentiment *uncertainty* rather than sentiment level. Practical implications include:

- Widen spreads when sentiment signals are ambiguous (high entropy, high MC variance)
- Weight institutional signals more heavily

during detected crisis conditions

- Track retail-institutional divergence as a volatility indicator

Risk Management. The divergence tracking methodology may have applications for risk management beyond market making. If significant divergence predicts forward volatility, risk managers could use divergence signals to adjust position sizing or hedging intensity.

Infrastructure Investment. The framework demonstrates the value of integrating multiple data sources (order books, social media, news, DeFi metrics). Organizations seeking to operationalize sentiment analysis should invest in multi-source infrastructure rather than relying on single sentiment streams.

6.4 Connection to Prior Work

ASRI Framework Integration. The successful integration of ASRI connectors (Farzulla and Maksakov, 2025) validates the modular design of that framework. The macro signal layer directly uses ASRI’s Google News RSS and DeFiLlama adapters, demonstrating that systemic risk monitoring infrastructure can support microstructure research applications.

Event Study Volatility Patterns. The finding that spreads are widest during uncertain periods connects to volatility asymmetry documented in cryptocurrency event studies (Farzulla, 2025a). Negative regulatory events produce larger volatility responses partly because they create uncertainty about regulatory trajectory. Our uncertainty decomposition provides a micro-level mechanism for this macro-level observation.

NLP Methodology. The CryptoBERT with MC Dropout approach extends domain-specific sentiment analysis (Farzulla, 2025b) with uncertainty quantification. The demonstrated patterns—low epistemic uncertainty for in-domain texts, higher for regulatory topics—validate the utility of uncertainty decomposition for financial applications.

7 Limitations

This section provides a comprehensive and candid accounting of the limitations inherent in this research. We organize limitations into six categories: data, model, simulation, methodology, generalizability, and scope. We view transparent acknowledgment of limitations as essential to scientific integrity and useful for guiding future research.

7.1 Data Limitations

The data infrastructure suffers from several significant limitations that affect the validity and generalizability of our findings.

Google News as Regulatory Sentiment Proxy. We use Google News RSS as a proxy for regulatory sentiment. This is a substantial limitation for several reasons:

- Institutional actors typically access professional financial news services (Bloomberg Terminal, Reuters Eikon, Refinitiv) that provide earlier, more comprehensive, and more accurate regulatory coverage
- Google News aggregates general media sources that may lag specialized financial reporting by hours or days
- News selection algorithms may introduce systematic biases (recency, engagement optimization) that do not reflect information relevance
- VADER sentiment analysis on news headlines provides only coarse sentiment estimation; headlines are often written for engagement rather than informational accuracy

A more rigorous implementation would integrate professional financial news feeds, though these are typically prohibitively expensive for academic research.

Reddit Sample Bias. The Reddit sample suffers from selection bias:

- Reddit users skew younger, more technically sophisticated, and more crypto-enthusiastic than the general retail population

- Subreddit communities exhibit herding behavior and echo chamber effects
- Active posters are not representative of lurkers, who may have different sentiment profiles
- Upvote/downvote dynamics filter which content becomes visible, introducing survivorship bias

Missing Twitter/X Data. The most significant data gap is Twitter/X, which historically provided the most real-time cryptocurrency sentiment signal. Recent API changes (post-2023) have made large-scale Twitter data collection prohibitively expensive or impossible for academic research. This represents a substantial limitation, as Twitter sentiment was likely more representative of real-time retail sentiment than Reddit.

DeFiLlama Coverage Gaps. While DeFiLlama provides excellent coverage of major DeFi protocols, it has limitations:

- Newer or smaller protocols may be missing or delayed
- Cross-chain aggregation methodology varies
- Historical data availability is limited for some protocols
- Real-time updates lag by approximately 10 minutes

Single-Exchange Order Book Data. We source order book data exclusively from Binance. This creates limitations:

- Binance represents a single venue in a fragmented market; other exchanges (Coinbase, Kraken, OKX) may exhibit different dynamics
- Cross-exchange arbitrage effects are not captured
- Binance-specific events (maintenance, regulatory issues in certain jurisdictions) could affect data quality
- The BTC/USDT pair may not generalize to other trading pairs

7.2 Model Limitations

The sentiment analysis and uncertainty quantification models have inherent limitations.

CryptoBERT Training Data Vintage.

CryptoBERT was trained on StockTwits cryptocurrency posts from 2021–2022. This creates several issues:

- Market narratives and vocabulary have evolved; terms common in 2024 (“ordinals,” “Bitcoin ETF,” “restaking”) may not be well-represented in training data
- The model may exhibit distributional shift when encountering novel topics
- StockTwits users may differ systematically from Reddit users in language patterns
- Training data predates major market events (Terra/Luna collapse, FTX fraud) that may have changed sentiment patterns

Monte Carlo Dropout as Approximate Bayesian Inference. MC Dropout provides an approximation to Bayesian neural networks, not exact posterior inference. Limitations include:

- The approximation quality depends on dropout rate and network architecture
- Uncertainty estimates may be miscalibrated (over- or under-confident)
- Epistemic uncertainty from MC Dropout may not capture all sources of model uncertainty
- The number of forward passes ($T = 50$) is chosen heuristically; different values may yield different uncertainty estimates

EWMA Smoothing Parameter Selection. The smoothing parameter $\alpha = 0.1$ is chosen heuristically based on target half-life assumptions. Different values would produce different sentiment dynamics, and the optimal choice likely varies with market conditions and application.

Regime Threshold Calibration. Regime detection thresholds (sentiment > 0.3 for bullish, peg deviation $> 2\%$ for crisis, etc.) are

theoretically motivated but not empirically calibrated. Different threshold choices would produce different regime classifications and potentially different results.

7.3 Simulation Limitations

The agent-based simulation has fundamental limitations that affect result interpretation.

Synthetic Price Dynamics. The most significant limitation is that price dynamics are generated within the simulation rather than matching real market data. The market maker model produces prices through a stylized quote-adjustment process that does not incorporate:

- Real order flow from exchanges
- Actual limit order book dynamics
- Cross-exchange price discovery
- Market manipulation and wash trading

This means that the comparison between single-source and multi-scale approaches tests the *internal consistency* of the framework rather than its *external validity* against real market behavior.

Limited Agent Heterogeneity. The current implementation includes only four agent types with fixed behavioral rules. Real markets exhibit:

- Continuous distributions of trader characteristics
- Adaptive learning and strategy evolution
- Complex interactions between agent types
- Institutional trading algorithms with diverse strategies

Absence of Cross-Exchange Dynamics. Cryptocurrency price discovery occurs across multiple exchanges simultaneously. Our single-venue simulation cannot capture:

- Arbitrage effects that enforce price consistency
- Venue-specific liquidity dynamics
- Information transmission across exchanges
- Regulatory fragmentation effects

Mesa Framework Constraints. The Mesa agent-based modeling framework, while suitable for prototyping, has performance limitations that constrain simulation scale. Real-time operation with thousands of agents would require optimization or alternative frameworks.

Absence of Fat Tails. The simulation produces return distributions approximating normality (kurtosis ≈ 0), while real cryptocurrency returns exhibit pronounced fat tails (kurtosis $\gg 3$). This reflects the simplified agent interactions and absence of mechanisms that generate extreme events.

7.4 Methodological Limitations

The research design has fundamental methodological limitations.

Causal Identification Challenges. We demonstrate correlations (divergence-volatility, uncertainty-spread) but cannot establish causation:

- Omitted variable bias: unobserved factors may drive both sentiment and market outcomes
- Reverse causality: market outcomes may affect sentiment rather than vice versa
- Simultaneity: sentiment and prices may be jointly determined
- No natural experiments or instrumental variables are available for identification

Endogeneity Concerns. Sentiment and prices are endogenously determined:

- Price movements generate sentiment (“number go up” bullishness)
- Sentiment-driven trading affects prices
- This feedback loop makes causal interpretation difficult

Sample Period Specificity. Results are derived from a specific (brief) sample period. Market dynamics may differ substantially during:

- Bull versus bear markets
- High versus low volatility regimes

- Before versus after major regulatory events
- Different macroeconomic conditions

Real-Time vs Historical Analysis. The framework is designed for real-time operation but validated using historical analysis. Look-ahead bias and timing assumptions may not translate to true real-time performance.

7.5 Generalizability

Results may not generalize beyond the specific context studied.

Bitcoin-Centric Model. The framework is developed and tested on Bitcoin (BTC/USDT). Generalization to other cryptocurrencies is uncertain:

- Altcoins have different market structures, liquidity profiles, and sentiment dynamics
- Meme coins and smaller tokens may be dominated by coordination effects not captured by our model
- DeFi tokens have additional dynamics (governance, yield farming) beyond sentiment

Regime Definitions May Not Transfer. The regime thresholds (crisis, regulatory, bullish, bearish) are calibrated for Bitcoin. Other assets may require different thresholds or entirely different regime definitions.

Market Structure Evolution. Cryptocurrency market structure evolves rapidly. The framework may become obsolete as:

- New data sources emerge (on-chain analytics, alternative social platforms)
- Regulatory frameworks change
- Market participant composition shifts
- Technology enables new trading strategies

Cultural and Linguistic Scope. The framework analyzes English-language content. Cryptocurrency markets are global, and sentiment in non-English communities (Chinese, Korean, Russian, etc.) may exhibit different dynamics not captured here.

7.6 What This Paper Does NOT Claim

This research explicitly does *not* claim:

1. **Trading strategy validity:** The framework is not intended as, nor validated for, live trading applications. Anyone using these methods for trading does so at their own risk
2. **Production readiness:** This is a research prototype demonstrating methodological concepts, not deployment-ready software. Substantial engineering would be required for operational use
3. **Definitive causal proof:** We demonstrate correlations and propose mechanisms; causal identification remains challenging and is not achieved here
4. **Optimal parameter calibration:** Regime weights, thresholds, and other hyperparameters are theoretically motivated but not empirically optimized. Different choices may produce different results
5. **Complete market representation:** The framework models highly simplified dynamics; real markets exhibit additional complexity that is not captured
6. **Regulatory compliance:** The framework does not address regulatory requirements for algorithmic trading, market manipulation detection, or financial services compliance
7. **Reproducibility guarantee:** While we provide code and describe methodology, exact reproducibility may be affected by API changes, data availability, and software version dependencies

We emphasize that this is *research*—a proof of concept exploring methodological ideas. The dramatic performance differences between single-source and multi-scale approaches, while suggestive, should be viewed as demonstrations of framework behavior rather than validated empirical findings about real market dynamics.

8 Conclusion

This paper has presented a multi-scale sentiment analysis framework for agent-based modeling of cryptocurrency market microstructure. We developed an architecture that blends institutional-level macro signals—derived from the ASRI framework (Farzulla and Maksakov, 2025)—with retail-level micro signals from CryptoBERT sentiment analysis, using regime-adaptive weighting and uncertainty decomposition to enable nuanced agent responses to heterogeneous information sources.

Summary of Contributions. The paper makes four methodological contributions:

1. A multi-scale sentiment architecture integrating macro (regulatory news, DeFi stability) and micro (social media sentiment) signals through regime-adaptive blending
2. An uncertainty decomposition framework separating epistemic (model confidence, regulatory opacity) from aleatoric (market noise, signal ambiguity) uncertainty components
3. A divergence tracking methodology for monitoring retail-institutional sentiment gaps as a potential volatility predictor
4. Regime-adaptive weighting mechanisms that increase institutional signal weight during crisis and regulatory periods

Key Findings. Preliminary results from real data integration reveal several findings of interest:

- Multi-scale blending dramatically reduces simulated volatility (128% to 2.6%) and spreads (32.72 to 2.43 bps) compared to single-source sentiment
- Uncertainty dominates sentiment level in determining market maker spread-setting—spreads are widest during neutral (uncertain) regimes rather than sentiment extremes
- Divergence between retail and institutional sentiment shows modest but statistically significant correlation with forward volatility

(0.18–0.24), providing preliminary support for the multi-scale divergence hypothesis

- Regime transitions exhibit high persistence (>97.8%), reflecting both underlying market dynamics and the hysteresis mechanism

Theoretical Implications. The framework extends classic market microstructure models (Glosten and Milgrom, 1985; Avellaneda and Stoikov, 2008) by incorporating social sentiment as an information channel with explicit uncertainty quantification. The key theoretical insight is that *sentiment uncertainty rather than sentiment level* determines adverse selection risk and should drive spread-setting behavior.

Limitations and Honest Assessment. We have provided an extensive accounting of limitations spanning data sources (Google News proxy, Reddit bias, missing Twitter), models (CryptoBERT vintage, MC Dropout approximation), simulation (synthetic dynamics, limited heterogeneity), methodology (causal identification challenges, endogeneity), and generalizability (Bitcoin-centric, market evolution). We emphasize that this is a research prototype demonstrating methodological concepts, not a validated trading system or production-ready software.

Future Research Directions. Several directions would strengthen and extend this work:

1. **Real order flow integration:** Validate the framework against actual Binance trade data rather than synthetic dynamics
2. **Extended agent heterogeneity:** Implement adaptive learning agents that evolve strategies based on market outcomes
3. **Cross-exchange dynamics:** Model multiple venues with arbitrage linkages to capture price discovery processes
4. **Crisis period validation:** Test regime-adaptive weighting during actual market stress (e.g., Terra/Luna collapse, FTX fraud)
5. **Alternative data sources:** Integrate on-chain analytics, alternative social platforms, and professional news feeds
6. **Causal identification:** Develop identification strategies (instrumental variables, natural experiments) to strengthen causal claims

Broader Context. This research contributes to the emerging literature on sentiment-driven market dynamics by providing a methodological framework for multi-scale analysis. As cryptocurrency markets mature and institutional participation increases, understanding the interaction between retail and institutional information processing becomes increasingly important. The framework developed here provides tools for investigating these dynamics, while acknowledging the substantial work required to validate findings against real market behavior.

The code and data pipelines developed for this research are available to support reproducibility and extension by other researchers. We hope this contribution stimulates further work on multi-scale approaches to market microstructure analysis.

Acknowledgments

The author thanks Andrew Maksakov for collaboration on the ASRI framework that provides the macro signal layer for this research. Infrastructure support provided by Resurrexi Labs.

References

- Jethin Abraham, Daniel Higdon, John Nelson, and Juan Ibarra. Cryptocurrency price prediction using tweet volumes and sentiment analysis. *SMU Data Science Review*, 1(3):1–22, 2018.
- Viral V. Acharya, Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson. Measuring systemic risk. *The Review of Financial Studies*, 30(1):2–47, 2017. doi: 10.1093/rfs/hhw088.
- Tobias Adrian and Markus K. Brunnermeier. CoVaR. *American Economic Review*, 106(7):1705–1741, 2016. doi: 10.1257/aer.20120555.
- Werner Antweiler and Murray Z. Frank. Is all that talk just noise? the information content of internet stock message boards. *The Journal of Finance*, 59(3):1259–1294, 2004.
- Marco Avellaneda and Sasha Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.
- Malcolm Baker and Jeffrey Wurgler. Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2):129–152, 2007.
- Monica Billio, Mila Getmansky, Andrew W. Lo, and Loriana Pelizzon. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559, 2012.
- Elie Bouri, Peter Molnar, Georges Azzi, David Roubaud, and Lars Ivar Hagfors. On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Research in International Business and Finance*, 41:88–98, 2017.
- Rama Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.
- Rama Cont, Sasha Stoikov, and Rishi Talreja. A stochastic model for order book dynamics. *Operations Research*, 58(3):549–563, 2010.
- DeFiLlama. DeFiLlama API: Open and transparent DeFi analytics. <https://defillama.com>, 2024. Accessed: December 2025.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, pages 4171–4186, 2019.
- ElKulako. CryptoBERT: A language model for cryptocurrency sentiment. Hugging Face, 2024. URL <https://huggingface.co/ElKulako/cryptobert>.
- J. Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.
- Murad Farzulla. Market reaction asymmetry in cryptocurrency event studies: A tarch-x approach. *Zenodo*, 2025a. doi: 10.5281/zenodo.17679537. Under review at Digital Finance Journal (Springer).
- Murad Farzulla. Do whitepaper claims predict market behavior? evidence from cryptocurrency factor analysis. *Zenodo*, 2025b. doi: 10.5281/zenodo.17917922.

- Murad Farzulla and Andrew Maksakov. ASRI: An aggregated systemic risk index for cryptocurrency markets. *Zenodo*, 2025. doi: 10.5281/zenodo.17918239.
- Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *International Conference on Machine Learning*, pages 1050–1059. PMLR, 2016.
- David Garcia and Frank Schweitzer. Social signals and algorithmic trading of Bitcoin. *Royal Society Open Science*, 2(9):150288, 2015.
- Lawrence R. Glosten and Paul R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.
- Anton Golub, John Keane, and Ser-Huang Poon. High frequency trading and mini flash crashes. *arXiv preprint arXiv:1211.6667*, 2012.
- John M. Griffin and Amin Shams. Is Bitcoin really untethered? *The Journal of Finance*, 75(4):1913–1964, 2020.
- Lewis Gudgeon, Daniel Perez, Dominik Harz, Benjamin Livshits, and Arthur Gervais. The decentralized financial crisis. *2020 Crypto Valley Conference on Blockchain Technology (CVCBT)*, pages 1–15, 2020.
- Joel Hasbrouck. *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading*. Oxford University Press, Oxford, 2007.
- Cars H. Hommes. Heterogeneous agent models in economics and finance. *Handbook of Computational Economics*, 2:1109–1186, 2006.
- Hugging Face. Transformers: State-of-the-art natural language processing. <https://huggingface.co/transformers>, 2024. Accessed: December 2025.
- Jackie Kazil, David Masad, and Andrew Crooks. Utilizing Python for agent-based modeling: The Mesa framework, 2020.
- Alex Kendall and Yarin Gal. What uncertainties do we need in Bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems*, 30, 2017.
- Ariah Klages-Mundt, Dominik Harz, Lewis Gudgeon, Jun-You Liu, and Andreea Minca. Stablecoins 2.0: Economic foundations and risk-based models. *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, pages 59–79, 2020.
- Olivier Kraaijeveld and Johannes De Smedt. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65:101188, 2020.
- Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985.
- Blake LeBaron. Agent-based computational finance. *Handbook of Computational Economics*, 2:1187–1233, 2006.
- Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A robustly optimized BERT pre-training approach. In *arXiv preprint arXiv:1907.11692*, 2019.

- Yukun Liu and Aleh Tsyvinski. Risks and returns of cryptocurrency. *The Review of Financial Studies*, 34(6):2689–2727, 2021.
- Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65, 2011.
- Igor Makarov and Antoinette Schoar. Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, 135(2):293–319, 2020.
- Mark Paddrik, Roy Hayes, Andrew Todd, Steve Yang, Peter Beling, and William Scherer. An agent based model of the E-Mini S&P 500 applied to flash crash analysis. In *IEEE Conference on Computational Intelligence for Financial Engineering & Economics*, pages 1–8. IEEE, 2012.
- R. G. Palmer, W. Brian Arthur, John H. Holland, Blake LeBaron, and Paul Tayler. Artificial economic life: A simple model of a stockmarket. *Physica D: Nonlinear Phenomena*, 75(1-3): 264–274, 1994.
- Dionisis Philippas, Hatem Rjiba, Khaled Guesmi, and Stephane Goutte. Media attention and Bitcoin prices. *Finance Research Letters*, 30:37–43, 2019.
- Paul C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.