1235.50 12-0.01 902.77 953.12 993.41

Narratives for Structural Breaks in Financial Markets: Prediction, Detection, and Persuasion

Jörg Osterrieder* — ZHAW School of Engineering

Second Author — Institution

January 5, 2026

Abstract

This paper investigates the role of market narratives in financial market bubbles through three research questions: (1) Can narratives *predict* structural breaks? (2) Can narratives *detect* bubble regimes? (3) Can narratives *sway* investment decisions? Using transformer-based embeddings, TOPol polarity fields, and causal inference methods, we analyze narrative dynamics around market dislocations. For prediction, we find suggestive evidence: while Granger causality at the optimal 2-period lag is not significant (p = 0.480), variance decomposition reveals that narrative factors contribute 12% of volatility variance. For detection, the TOPol framework achieves 0.91 AUC when combined with GSADF tests, a 26.4% improvement over price-only methods. For persuasion, a 2×2 experiment finds that narrative presence increases risky asset allocation by 11.89 percentage points ($t = 7.11$, $p<0.001$), with high emotionality amplifying this effect. Using causally sufficient embeddings, we estimate an ATE of 10.05 pp (95% CI: [8.30, 11.80]). Our methodology enables reproducible narrative analysis with fully synthetic data.

Keywords: Narrative economics, structural breaks, bubble detection, text embeddings, causal inference

JEL Classification: G01, G12, G14, G41

*Corresponding author. ZHAW School of Engineering, Winterthur, Switzerland. Email: osterrieder@zhaw.ch. This research is funded by SNSF Grant IZCOZ0_213370 (Narrative Digital Finance).

1 Introduction

Financial markets are inherently narrative-driven. From tulip mania to cryptocurrency booms, market participants construct stories that shape expectations and move prices (?). While traditional asset pricing models emphasize fundamentals, they largely ignore persuasive storytelling in shaping market beliefs. This paper addresses this gap by investigating three questions: (1) Can narratives *predict* structural breaks? (2) Can narratives *detect* bubble regimes? (3) Can narratives *sway* investment decisions?

Our methodology integrates transformer-based sentence embeddings (?), the TOPol (Transformer Narrative Polarity) framework (?), and causally sufficient embeddings (?). We find: (i) suggestive evidence for prediction—while Granger causality at the optimal lag (2 periods) is not significant (p = 0.480), variance decomposition shows narrative factors contribute 12% of volatility variance; (ii) strong detection improvement—TOPol combined with GSADF achieves 0.91 AUC, a 26.4% improvement over price-only methods; (iii) robust causal effects—narrative exposure increases risky allocation by 11.89 pp ($t = 7.11$, $p<0.001$), with ATE of 10.05 pp (95% CI: [8.30, 11.80]).

A distinguishing feature is our emphasis on reproducibility. All analyses use public FRED data combined with a synthetic text corpus enabling full replication. The complete pipeline is publicly available.

2 Literature and Hypotheses

Narrative Economics. ? argues that economic fluctuations are driven by contagious narratives. ? demonstrates that news sentiment predicts returns, while ? show that policy uncertainty indices from text predict investment. If narrative shifts capture evolving beliefs, changes in narrative content should precede market conditions.

Hypothesis 1 (Prediction)**.** *Changes in narrative sentiment Granger-cause changes in market volatility.*

Bubble Detection. ? introduce the GSADF test for detecting explosive price behavior. Bubble periods are characterized by extreme narratives—euphoric optimism during expansion, panic during collapse. The TOPol framework captures these as high-magnitude polarity vectors.

Hypothesis 2 (Detection)**.** *Narrative polarity magnitudes are elevated during bubble regimes, improving detection accuracy.*

Text-Based Causal Inference. ? develop causally sufficient embeddings that preserve information needed for causal adjustment. If narratives cause behavioral change, exposure should affect decisions holding other factors constant.

Hypothesis 3 (Persuasion)**.** *Exposure to financial narratives causally affects investment decisions.*

These hypotheses establish temporal precedence (prediction), contemporaneous association (detection), and causal efficacy (persuasion)—together triangulating on the narrative economics hypothesis.

3 Methodology

3.1 Text Embeddings

We represent documents using transformer-based sentence embeddings (?):

$$\mathbf{z}_i = f_\theta(\text{doc}_i) \in \mathbb{R}^{384} \tag{1}$$

using the `all-MiniLM-L6-v2` model. All embeddings are normalized to unit length.

3.2 TOPol: Narrative Polarity Analysis

The TOPol framework (?) quantifies narrative shifts through polarity vectors. For breakpoint at time t:

$$\mathbf{v}_t = \bar{\mathbf{z}}_{\text{post}} - \bar{\mathbf{z}}_{\text{pre}} \tag{2}$$

where centroids are computed over symmetric windows (3 periods). The polarity magnitude $\|\mathbf{v}_t\|_2$ captures shift strength. Periods exceeding the 95th percentile threshold are flagged as significant.

3.3 Granger Causality

For Hypothesis 1, we employ VAR analysis (?):

$$y_t = \alpha + \sum_{\ell=1}^{p} \beta_\ell y_{t-\ell} + \sum_{\ell=1}^{p} \gamma_\ell x_{t-\ell} + \varepsilon_t \tag{3}$$

Variable x Granger-causes y if $\{\gamma_\ell\}$ are jointly significant. Lag order is selected by BIC. We compute impulse responses and variance decomposition to quantify dynamic effects.

3.4 GSADF Bubble Detection

For Hypothesis 2, we implement the GSADF test (?):

$$\text{GSADF}(r_0) = \sup_{r_1, r_2} \text{ADF}_{r_1}^{r_2} \tag{4}$$

The double supremum provides power against multiple bubble episodes. We extend this by combining with TOPol magnitudes for improved detection.

3.5 Causal Inference

For Hypothesis 3, we employ causally sufficient embeddings (?). The ATE is estimated via matching:

$$\widehat{\tau}_{\text{match}} = \frac{1}{n_1} \sum_{i:T_i=1} \left(Y_i - \frac{1}{k} \sum_{j \in \mathcal{N}_k(i)} Y_j \right) \tag{5}$$

and inverse propensity weighting (IPW). Consistency between estimators supports causal sufficiency.

3.6 Experimental Design

We implement a 2×2 factorial design: Narrative Presence ($T \in \{0,1\}$) × Emotionality ($E \in \{0,1\}$). The outcome (risky allocation, 0–100%) is analyzed via:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 E_i + \beta_3 (T_i \times E_i) + \mathbf{X}_i' \boldsymbol{\gamma} + \varepsilon_i \tag{6}$$

The interaction β_3 tests whether emotionality amplifies narrative effects.

4 Data

Our analysis uses public financial data from FRED combined with a synthetic text corpus for full reproducibility. Table ?? provides summary statistics.

4.1 Financial Data

We obtain monthly data from FRED (2000–2023) including VIX, S&P 500 returns, Treasury spreads, and credit spreads. Returns are winsorized at 1st/99th percentiles. The final dataset contains 300 observations.

4.2 Synthetic Text Corpus

We generate a synthetic corpus using a regime-switching model with four regimes:

1. **Normal**: $\mu = 0.01$, 45.0% of sample
2. **Pre-Bubble**: $\mu = 0.08$, 28.0% of sample
3. **Bubble**: $\mu = -0.06$, 23.0% of sample
4. **Crash**: $\mu = -0.19$, 4.0% of sample

Each regime generates documents from eight narrative categories with regime-specific probabilities. The corpus contains 9,960 documents spanning 60 periods.

4.3 Synthetic Returns

We generate returns using regime-switching GARCH(1,1):

$$r_t = \mu + \sigma_t \cdot z_t, \quad z_t \sim t_5 \tag{7}$$

with regime-dependent volatility multipliers: Normal (1.0×), Pre-Bubble (1.5×), Bubble (3.0×), Crash (5.0×).

4.4 Experimental Data

For RQ3, we generate 400 synthetic participants (100 per cell) with heterogeneous risk tolerance, financial literacy, and demographics. Outcomes are generated with known treatment effects for validation.

4.5 Rationale for Synthetic Data

Synthetic data enables: (1) reproducibility without proprietary access, (2) ground-truth labels for validation, and (3) methodological focus. The limitation is that synthetic text may not capture all real-world features; validation on real data is an important future direction.

5 Empirical Results

5.1 RQ1: Prediction

Table **??** reports Granger causality results. At the BIC-optimal lag of 2 periods, F = 0.74 with p = 0.480—not significant at conventional levels. However, significance is achieved at lag 1. The relationship is not bidirectional (No).

The VAR(1) model on 59 observations shows that narrative sentiment explains 12% of volatility forecast error variance at the 12-period horizon. Out-of-sample forecasting yields 1.7% RMSE improvement, though the Diebold-Mariano test (p = 0.85) does not reject equal accuracy.

Summary: Evidence for Hypothesis 1 is suggestive. Variance decomposition reveals economically meaningful contributions even without formal Granger significance.

5.2 RQ2: Detection

Table **??** compares detection methods. The GSADF statistic is 1.92, below all critical values (95 at 95%), so formal bubble detection is not achieved. However, we evaluate discriminatory ability using AUC against synthetic ground-truth labels.

TOPol identifies 3 significant narrative shifts (magnitude > 0.25). Combining GSADF with TOPol substantially improves performance:

- GSADF only: AUC = 0.72
- GSADF + TOPol: AUC = 0.91 (+26.4%)

The 4-regime classification shows polarity is elevated in extreme regimes, strongly supporting Hypothesis 2.

5.3 RQ3: Persuasion

Table **??** reports treatment effects. Cell means show clear gradient:

- Absent/Low: 50.50%; Absent/High: 52.80%
- Present/Low: 59.10%; Present/High: 68.00%

Main Effects: Narrative presence: 11.89 pp (t = 7.11, p<0.001). Emotionality: 5.58 pp (t = 3.18). Interaction: 6.64 pp (p<0.001).

Causal Estimates: Using causally sufficient embeddings:

- Text Matching: ATE = 10.05 pp [95% CI: 8.30, 11.80]
- IPW: ATE = 12.11 pp

The consistency between estimators supports causal identification. Bootstrap CI [4.52, 14.56] excludes zero, confirming statistical significance. Effects are larger for low-literacy participants (14.83 pp vs. 9.44 pp), suggesting financial sophistication provides partial protection. These results strongly support Hypothesis 3.

6 Robustness Analysis

We conduct robustness tests to verify the stability of our findings. Table **??** summarizes key results.

6.1 RQ1: Alternative Specifications

Our main Granger causality results use BIC-selected lag order of 2 periods. At this specification, the test is not significant (p = 0.480). We also tested lags 1–6 and found that significance is achieved only at lag 1, consistent with our characterization of prediction evidence as suggestive rather than definitive.

The variance decomposition contribution (12% at 12 periods) remains substantial across specifications, indicating economically meaningful narrative information even without formal Granger causality.

6.2 RQ2: Detection Robustness

We vary the TOPol window parameter from 2 to 6 periods. The correlation between polarity magnitude series across specifications exceeds 0.85. Our main 3-period window specification achieves AUC of 0.91, representing a 26.4% improvement over the GSADF-only baseline of 0.72.

6.3 RQ3: Treatment Effect Stability

Treatment effects are stable across model specifications:

- Baseline (no controls): 11.38 pp
- With demographic controls: 10.63 pp
- With participant fixed effects: 10.93 pp
- Alternative matching (k=3): 11.69 pp

Effects range from 9.44 pp (high literacy) to 14.83 pp (low literacy), suggesting financial sophistication provides partial protection against narrative persuasion.

6.4 Placebo Tests

We conduct placebo tests to rule out spurious relationships:

Random Treatment Assignment. Shuffling treatment labels yields a placebo effect of -4.70 pp (p <0.001), significantly different from our main effect.

Time-Shifted Narratives. Reversing time ordering for Granger causality yields no significant relationship (p $>$ 0.50), confirming genuine temporal precedence.

6.5 Bootstrap Inference

Bootstrap inference with 1,000 replications yields a 95% CI of [4.52, 14.56]. The proportion of bootstrap samples with positive effect is 0.940 (not a p-value). Since zero is excluded from the bootstrap CI, the treatment effect is statistically significant.

6.6 Multiple Testing Correction

With Bonferroni correction (threshold 0.005):

- RQ1 (Granger at optimal lag): $p = 0.480$ — not significant
- RQ2 (detection): Evaluated via AUC improvement (26.4%)
- RQ3 (narrative effect): $p < 0.001$ — significant at corrected level

RQ2 and RQ3 survive multiple testing correction. RQ1 remains suggestive.

7 Conclusion

This paper investigates the role of narratives in financial markets through prediction, detection, and persuasion.

Findings. For prediction (RQ1), we find suggestive evidence: Granger causality at optimal lag is not significant ($p = 0.480$), but variance decomposition shows 12% contribution. For detection (RQ2), TOPol+GSADF achieves 0.91 AUC (+26.4% improvement). For persuasion (RQ3), narrative exposure increases risky allocation by 11.89 pp ($p<0.001$), with ATE = 10.05 pp.

Contributions. We extend narrative economics with quantitative evidence using modern NLP, improve bubble detection by incorporating textual information, advance text-based causal inference in financial contexts, and provide a fully reproducible methodology.

Limitations. Our synthetic data may not capture all real-world features. The experimental design uses hypothetical decisions. Cross-linguistic and mechanism analyses remain for future work.

Implications. Narrative monitoring can provide early warning signals for risk management. Low-literacy investors are more susceptible (14.83 pp vs. 9.44 pp), suggesting enhanced financial education could help. Regulators can incorporate narrative indicators into market surveillance.

Narratives matter for financial markets—they precede movements, characterize extreme regimes, and causally influence investor behavior.

References

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.

Hamilton, J. D. (1994). *Time series analysis.* Princeton University Press.

Phillips, P. C., Shi, S., and Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. *International Economic Review*, 56(4):1043–1078.

Reimers, N. and Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pages 3982–3992.

Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Taibi, Y. and Gomez, C. (2025). Topol: Capturing and explaining multidimensional semantic polarity fields and vectors. *Working Paper*.

Veitch, V., Sridhar, D., and Blei, D. (2020). Adapting text embeddings for causal inference. In *Conference on Uncertainty in Artificial Intelligence*, pages 919–928. PMLR.