Financial-Condition-Guided Retrieval for Adaptive Bayesian Decision Forecasting Under Structural Breaks

Lieu Hoai Phuc[1,2], Tran Nguyen Y Binh[1,2], Nguyen Thi Anh Ngoc[1,2], Pham Le Bao Kim[1,2], Phan Nguyen Tri Tin[1,2], and Nguyen Phuc Son[1,2]

[1] University of Economics and Law, Viet Nam
[2] Vietnam National University, Ho Chi Minh City, Viet Nam
{phuclh22414c, binhtny23413e, ngocnta23414a, kimplb23413e, tinpnt23413e}@st.uel.edu.vn, sonnp@uel.edu.vn

Abstract. Bayesian Dynamic Linear Models provide coherent probabilistic forecasting, yet their sequential learning can adapt too slowly after abrupt policy-driven regime shifts, producing transient bias precisely when accurate predictions are most valuable. We propose FCI-Retrieval-DLM, an extension that retrieves historically similar macro-financial episodes via the Financial Conditions Impulse on Growth (FCI-G) index and applies a reliability-gated correction to the baseline predictive mean within event windows. A Shuffled-Retrieval Control, which preserves the full intervention pipeline while destroying semantic alignment, serves as a placebo-style negative control to isolate the contribution of context-guided matching. In walk-forward NASDAQ forecasting, FCI-Retrieval-DLM yields modest directional improvements over Base-DLM in both point and distributional accuracy, with gains concentrated in onset-centered post-event windows. However, differences relative to the shuffled control remain small under the current configuration, indicating that evidence for a distinct semantic retrieval advantage is not yet conclusive.

Key words: Policy-sensitive forecasting; FCI-G; Dynamic Linear Model.

1 Introduction

Macroeconomic policy interventions, whether monetary or fiscal, expansionary or contractionary, arrive as discrete, forward-looking events that can abruptly reshape the relationship between financial conditions and asset returns. Yet these policy windows, precisely when forecasts are most valuable for trading and risk management, are also when models are most prone to failure. Identical actions may trigger opposite market reactions depending on the prevailing macro-financial regime. This state dependence implies that policy effects are inherently context-specific rather than uniform over time. Bayesian state-space models with time-varying parameters and stochastic volatility [21,18] have become standard tools for forecasting under structural instability because they update sequentially while preserving probabilistic coherence. However, their stability under gradual change can become a weakness around discrete shocks. After prolonged calm periods, posterior uncertainty contracts and learning slows, generating adaptation lag in the immediate aftermath of policy announcements. Empirical evidence further indicates that policy transmission is regime-dependent, with market responses varying across financial conditions and crisis states [12,6,9]. A forecasting system should therefore adapt not only continuously over time, but selectively and more rapidly when economically consequential events occur.

To address this challenge, we propose an event-aware, context-guided Bayesian forecasting framework. When a policy event occurs, the model does not relearn from scratch. Instead, it retrieves historically similar macro-financial environments using the Financial Conditions Impulse on Growth (FCI-G) index [2] as a compact context signal. By identifying past episodes in which similar conditions led to systematic forecast errors, the framework temporarily accelerates learning and applies a conservative adjustment to the baseline forecast. This mechanism reduces adaptation lag while preserving Bayesian coherence. We evaluate the approach in walk-forward forecasting of equity returns around recurrent, precisely dated policy events, including Federal Reserve 25-basis-point rate changes and major tax announcements. The proposed method improves both point and density forecasts, with gains concentrated in the immediate post-event window where conventional state-space models adapt most slowly. More broadly, the framework introduces a policy-sensitive mechanism that enhances the responsiveness of Bayesian state-space models to discrete structural disruptions.

2 Related Work

2.1 Forecasting Under Structural Instability

Forecasting financial time series under structural breaks is a central challenge in macro-finance. When the data-generating process shifts, models that impose parameter constancy suffer substantial forecast deterioration, as shown in [18]; standard autoregressive forecasts perform poorly in small samples unless breaks are explicitly modeled. Regime-switching models [10] capture abrupt transitions via discrete latent states, but real-time regime identification is difficult and discrete switching may create artificial discontinuities when adjustments are gradual. These limitations motivate adaptive frameworks in which parameters evolve continuously over time.

2.2 Bayesian State-Space Models and Adaptation Lag

State-space time-varying parameter models offer a coherent alternative. Dynamic Linear Models and TVP-VARs allow coefficients to evolve stochastically, typically as random walks, while maintaining Bayesian consistency through sequential updating [21,18]. However, adaptation to abrupt shocks can remain gradual: as posterior uncertainty contracts during stable periods, new information receives less weight, creating a trade-off between responsiveness and forecast stability in Bayesian learning [5].

2.3 Financial Conditions and Regime-Dependent Dynamics

Financial Conditions Indexes (FCIs) summarize macro-financial states by aggregating credit spreads, funding conditions, and asset prices, and have been shown to improve volatility and risk forecasts, particularly during stress episodes [15]. The Financial Conditions Impulse on Growth (FCI-G) provides a forward-looking measure of how past financial conditions affect future economic activity [2]. Although FCIs are widely used as predictors, their role in shaping the learning dynamics of Bayesian state-space models remains limited, making their integration into adaptive filtering a natural extension of the literature.

3 Problem Statement

We consider probabilistic forecasting of financial time series in environments where structural breaks and policy-driven shocks may abruptly alter market dynamics. A common baseline in this setting is the Dynamic Linear Model with time-varying parameters and stochastic volatility (DLM-TVP-SV), which performs recursive Bayesian filtering while allowing coefficients and volatility to evolve over time [21,18].

However, this gradual learning mechanism becomes problematic when the data-generating process shifts abruptly. After long periods of stability, the model becomes increasingly confident in its parameter estimates. As a result, new observations receive relatively little weight in the updating step, slowing the adjustment of the predictive mean and creating temporary forecast bias [16]. At the same time, variance discounting reacts to sudden shocks by increasing predictive uncertainty. Although this mechanism protects the model from overconfidence during turbulent periods, the resulting uncertainty often decays slowly, leaving predictive intervals overly wide even after market conditions begin to stabilize [5].

These two effects jointly reduce forecasting reliability during policy-driven regime shifts. The predictive mean adapts too slowly to the new environment, while predictive uncertainty remains inflated longer than necessary. Consequently, both point forecasts and probabilistic forecasts deteriorate precisely during the periods when accurate predictions are most valuable for real-time decision-making.

4 Methodology

4.1 Data and Preprocessing

Target Variable We evaluate the proposed framework using U.S. equity data, with the NASDAQ index serving as an empirical testbed. The index captures the dynamics of technology-oriented and growth-focused equities and provides a representative high-frequency market environment. Daily observations from 1994 to 2025 are employed to analyze market responses to macroeconomic and policy-driven events with clearly identified announcement dates.

Let y_t denote the daily logarithmic return at time t, defined as

$$y_t = \log(P_t) - \log(P_{t-1}) \tag{1}$$

where P_t is the closing price on trading day t. Working with daily returns allows precise detection of abrupt expectation adjustments around policy events and structural changes.

Conditioning Variable: FCI-G as Market State To condition return dynamics on the broader macro-financial environment, we incorporate the Financial Conditions Impulse on Growth (FCI-G) as the sole contextual variable. The role of FCI-G in this framework is not that of a direct macroeconomic predictor, but rather as a quantitative proxy for the prevailing market state.

FCI-G is a model-based composite index constructed from multiple financial components, including interest rates, credit spreads, asset prices, and exchange rates. Unlike individual macro indicators such as GDP growth, inflation, or policy rates, which each provide only partial and indirect perspectives on economic conditions, FCI-G aggregates these channels into a unified measure of financial tightness or ease. Its explicit lag structure captures how past financial conditions generate forward-looking headwinds or tailwinds for growth. As such, FCI-G provides a coherent summary of the macro-financial regime under which market participants form expectations.

From an econometric perspective, incorporating multiple macro variables separately would substantially expand the state dimension in a Bayesian state-space framework. Given the recursive nature of Dynamic Linear Model (DLM) estimation, such dimensional expansion may introduce multicollinearity, weaken parameter identification, and inflate posterior uncertainty. In contrast, FCI-G offers a principled dimensionality-reduction strategy that preserves informational richness while maintaining statistical tractability and filtering stability. It therefore serves as an economically grounded and parsimonious conditioning variable for regime-aware adaptation.

Event Identification and Validation Within the Bayesian DLM framework, structural change is treated as an adaptive filtering problem rather than as an explicit breakpoint search. Following West and Harrison [21], abrupt shifts can be accommodated through two mechanisms: intervention and discounting (variance inflation). Intervention modifies the prior state distribution at economically meaningful points, allowing discrete shifts in level or coefficients. Discounting increases prior state uncertainty, enabling faster adaptation without disrupting sequential Bayesian updating. These mechanisms form the baseline adaptation structure.

To operationalize when adaptation should be activated, we introduce an event-driven triggering layer. Candidate intervention events consist of three macroeconomic policy categories with clear transmission channels to asset valuation: (i) Federal Reserve rate hikes (25bps), (ii) Federal Reserve rate cuts (25bps), and (iii) U.S. tax-reduction announcements. These events are precisely time-stamped, forward-looking in their implications for discount rates and growth expectations, and recurrent across policy cycles, making them natural candidates for regime shifts.

However, not every policy announcement generates an observable discontinuity in market data. To avoid spurious activation of adaptive mechanisms, each candidate event date is subjected to statistical validation. For a given event time t_0, balanced pre- and post-event windows are constructed, and distributional differences are evaluated using a Cramér–von Mises statistic within a permutation-based testing framework [4]. Under the null hypothesis of no structural break, pre- and post-event labels are treated as exchangeable. An event is classified as a validated breakpoint if its associated p-value falls below a predefined significance threshold.

Only statistically confirmed breakpoints are incorporated as conditional triggers for intervention or discounting. This design ensures that structural adaptation is activated exclusively when both economic relevance and empirical evidence of discontinuity are present.

4.2 Bayesian Dynamic Linear Model with Time-Varying Parameters and Adaptive Volatility (Base-DLM)

The baseline model is a Bayesian Dynamic Linear Model (DLM) with time-varying regression coefficients, following Davy and Miles (1977) and West and Harrison (1997) [21,7]. It provides a real-time probabilistic forecasting framework for NASDAQ returns under an information-available-at-time-t protocol, using only contemporaneous inputs, including ragged-edge macro series, as-of aligned macro-financial indicators, and delayed policy events. The state-space formulation allows the relationship between returns and macro-financial conditions to evolve smoothly over time while maintaining probabilistic coherence.

To enhance adaptability in policy-sensitive environments, Base-DLM incorporates event-dependent discounting in the state evolution and online calibration of predictive variance, yielding well-calibrated predictive distributions across heterogeneous regimes [17]. Despite this flexibility, the model remains structurally agnostic to event type: adaptation is entirely reactive through Bayesian filtering and driven by realized forecast errors, without explicit recognition of recurring patterns associated with monetary policy shocks, financial stress episodes, or other macroeconomic interventions. Accordingly, Base-DLM serves as a stable benchmark for evaluating the incremental contribution of proposed extensions.

The model produces multi-step forecasts

$$(y_{t+1}, \dots, y_{t+H}) \tag{2}$$

for horizons $H = 2$–5, using only information available up to time t.

State–Space Representation The observation equation is

$$y_t = F_t' \theta_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, V_t), \tag{3}$$

where

$$F_t = (1, x_{t1}, \ldots, x_{tp})' \tag{4}$$

includes an intercept and lagged macro-financial predictors, θ_t is a vector of time-varying coefficients, and V_t denotes observation variance. State evolution follows

$$\theta_t = G_t \theta_{t-1} + u_t, \qquad u_t \sim \mathcal{N}(0, W_t), \tag{5}$$

with

$$G_t = I_{p+1}, \tag{6}$$

implying independent random-walk dynamics and a first-order Markov structure that allows smooth parameter drift [7].

Discounting and Predictive Learning Rather than specifying W_t directly, parameter evolution is governed by the discount-factor formulation of Simonoff (1986) [19]. Let (m_{t-1}, C_{t-1}) denote the posterior moments of θ_{t-1}. The prior at time t is

$$\theta_t \mid D_{t-1} \sim \mathcal{N}(a_t, R_t), \qquad a_t = G_t m_{t-1}, \qquad R_t = \delta_t^{-1} C_{t-1}, \tag{7}$$

with $\delta_t \in (0, 1]$. In information form, the prior precision satisfies

$$P_{t|t-1} = \delta_t P_{t-1}, \tag{8}$$

which is algebraically equivalent to the standard Kalman filter representation under random walk state evolution [21]. This formulation guarantees strictly positive definite covariance matrices while controlling the effective adaptive gain of the filter.

The one-step-ahead predictive distribution is

$$y_t \mid D_{t-1} \sim t_{\nu_{t-1}}(f_t, q_t), \tag{9}$$

where

$$f_t = F_t' a_t, \qquad q_t = F_t' R_t F_t + S_{t-1}. \tag{10}$$

Here S_{t-1} is learned online via a Normal–Gamma conjugate updating scheme, yielding heavy-tailed predictive distributions with closed-form recursion [21].

Event Sensitivity and Calibration To increase responsiveness around policy shocks, the discount factor varies over time,

$$\delta_t = \begin{cases} \delta_{\text{event}}, & \text{if event\_day}_t = 1, \\ \delta_{\text{normal}}, & \text{otherwise}, \end{cases} \tag{11}$$

inflating prior state uncertainty during event periods without modifying the observation equation [19]. Predictive uncertainty is further calibrated using recent forecast errors

$$e_t = y_t - f_t, \tag{12}$$

with variance inflation

$$\tilde{q}_t = \lambda_t q_t, \tag{13}$$

where λ_t is a smoothed, regime-dependent factor estimated online [17]. Multi-step forecasts are obtained by iterating the discounted state evolution forward from time t, yielding coherent predictive distributions for $H = 2$–5.

Theoretical Validity Under this specification, Base-DLM satisfies the core requirements of a coherent Bayesian state-space DLM, including explicit measurement and transition equations, random-walk identity evolution, first-order Markov structure, independent disturbances, and positive-definite covariance propagation [21,7]. Time-varying discounting enables accelerated adaptation during event periods, while conjugate variance learning ensures well-calibrated, heavy-tailed predictive uncertainty [17].

4.3 Retrieval-Based Mean Adaptation (FCI-Retrieval-DLM) and Identification Strategy

As described in Section 4.2, Base-DLM provides a coherent real-time probabilistic forecast via sequential Bayesian filtering. However, because Base-DLM learns only from incoming observations, prolonged stability can reduce effective adaptive gain and lead to transient directional bias after abrupt, event-driven shifts. To localize responsiveness without changing the state evolution or predictive uncertainty structure, we introduce a retrieval-based mean adaptation layer (FCI-Retrieval-DLM) implemented on top of Base-DLM.

FCI-G Context and Retrieval At each time t, FCI-Retrieval-DLM constructs a macro-financial context vector $c_t \in \mathbb{R}^d$ from information available in real time. In this study, the context vector c_t includes the level of financial conditions (FCI-G), its short-term change $\Delta\text{FCI-G}_t$, and its deviation from a recent moving average $\text{FCI-G}_t - \text{MA}_{22,t}$. A dynamic memory $\mathcal{D} = \{(c_\tau, e_\tau)\}$ stores historical contexts and their associated one-step-ahead signed forecast errors from Base-DLM, where e_τ is the forecast error realized at time $\tau + 1$.

Given the current context c_t, historical analogs are retrieved by cosine similarity between c_t and past contexts c_τ. Let $\mathcal{N}_k(t)$ denote the top-k most similar contexts. Their errors are aggregated into a similarity-weighted signal

$$\tilde{e}_t = \sum_{\tau \in \mathcal{N}_k(t)} w_\tau e_\tau, \tag{14}$$

which summarizes systematic bias observed in historically similar regimes. This construction is closely related to kernel and nearest-neighbour regression estimators [14,20].

Reliability Gating and Mean Correction To avoid spurious adjustments, FCI-Retrieval-DLM separates detection (Trigger) from intervention (Apply). A correction is applied only if reliability conditions are satisfied: (i) retrieval is triggered, (ii) effective matching confidence exceeds a minimum threshold, (iii) the retrieved signal strength is sufficiently large, and (iv) the time index lies in a predefined event window. The apply indicator is

$$\mathbf{1}_{\text{apply},t} = \mathbf{1}_{\text{trigger},t} \cdot \mathbf{1}_{c_t^{\text{eff}} \geq c_{\min}} \cdot \mathbf{1}_{\text{gain}_t \geq g_{\min}} \cdot \mathbf{1}_{\text{event-window}}. \tag{15}$$

When Apply holds, the bias-mode correction is

$$\Delta_t^{\text{raw}} = \gamma\, \tilde{e}_t\, c_t^{\text{eff}}, \qquad \Delta_t = \text{clip}(\Delta_t^{\text{raw}}, -\Delta_{\max}, \Delta_{\max}), \tag{16}$$

and the predictive mean becomes

$$\hat{y}_t = \hat{y}_t^{(0)} + \mathbf{1}_{\text{apply},t} \Delta_t, \tag{17}$$

where $\hat{y}_t^{(0)}$ is the Base-DLM prediction. Conceptually, FCI-Retrieval-DLM borrows information from historically similar macro-financial regimes to anticipate the direction of adjustment before sufficient new observations have accumulated, thereby accelerating post-event mean adaptation while preserving the Bayesian filtering backbone; see Eqs. (14)–(17).

Identification via Shuffled Retrieval Placebo We evaluate the contribution of semantic (context-consistent) retrieval using a controlled design with three variants: Base-DLM without mean adaptation; FCI-Retrieval-DLM with the retrieval and gated mean correction described above; and Shuffled-Retrieval Control, which preserves the full intervention pipeline (retrieval, gating, and correction opportunity) but destroys semantic alignment by permuting retrieval associations. This placebo-style negative control [13,3,1] ensures that any advantage of FCI-Retrieval-DLM over the shuffled control is interpretable as mechanism-specific evidence for context-guided matching, rather than a generic effect of adding an adjustment layer. The next section details the walk-forward evaluation protocol, event windows, and ablation settings used in the empirical study.

4.4 Experimental Design

All methods are evaluated under walk-forward (online) updating with date-aligned scoring: model parameters are updated sequentially using information available up to time $t-1$, and predictions are scored at time t without look-ahead.

Event Windows We consider three onset-centered post-event windows, denoted $w0_1$, $w0_3$, and $w0_5$, corresponding to days since event onset in $\{0, 1\}$, $\{0, \ldots, 3\}$, and $\{0, \ldots, 5\}$, respectively. These choices balance two competing considerations: narrower windows focus on immediate post-onset adaptation but have limited sample size, whereas wider windows increase statistical power but dilute the event-onset signal.

Ablation Design We compare three matched conditions that differ only in matching semantics:

- No-adaptation control: matching-driven mean correction is disabled.
- Semantic event matching (proposed): retrieval preserves event-context similarity.
- Shuffled-Retrieval Control (negative control): intervention mechanics are preserved, but semantic alignment is destroyed by permutation of retrieved targets within the neighbor set.

The Shuffled-Retrieval Control functions as a placebo test for semantic retrieval [13,3,1]. It retains the intervention pipeline (triggering, gating, and the opportunity to apply corrections) while removing semantic alignment between the current event context and retrieved analogs. Under this design, a measurable advantage of semantic matching over the shuffled control is interpretable as mechanism-specific evidence rather than a generic side effect of adding an adjustment layer.

We report results at two scopes: (i) an all-scope aggregate over the full evaluation sample, and (ii) an EventWin-scope aggregate restricted to onset-centered post-event windows.

4.5 Evaluation Metrics

We report point accuracy (RMSE/MAE) and distributional accuracy using strictly proper scoring rules (LPD and CRPS). These metrics evaluate the quality of the predictive distribution used for decision support (rather than the realized utility of a particular policy). Let $\hat{y}_t$ denote the predictive mean and $p_t(\cdot)$ / $\mathcal{F}_t(\cdot)$ the predictive density / CDF. The log predictive density is

$$\mathrm{LPD} = \frac{1}{T}\sum_{t=1}^{T} \log p_t(y_t). \tag{18}$$

CRPS is a strictly proper scoring rule for distributional forecasts [8,11]:

$$\mathrm{CRPS}(\mathcal{F}_t, y_t) = \int_{-\infty}^{\infty} \left(\mathcal{F}_t(z) - \mathbf{1}\{y_t \le z\}\right)^2 dz, \qquad \mathrm{CRPS} = \frac{1}{T}\sum_{t=1}^{T} \mathrm{CRPS}(\mathcal{F}_t, y_t). \tag{19}$$

Unless stated otherwise, lower is better for CRPS/RMSE/MAE and higher is better for LPD.

5 Results

We evaluate the proposed retrieval mechanism in a policy-event setting, reporting both all-scope metrics (full sample) and onset-centered post-event windows around policy announcement dates (Fed rate decisions and major tax announcements), where rapid regime adjustment is most decision-relevant.

Main Comparison Tables 1 and 2 report the core ablation under the main setting. For both tables, lower CRPS/RMSE/MAE indicate better point or distributional accuracy, while higher LPD indicates better log predictive density; comparisons are made within each metric column.

Table 1. Core ablation in the main setting (All-scope metrics).

Setting	CRPS (all)	RMSE (all)	MAE (all)	LPD (all)
Base-DLM	0.010522	0.015530	0.012356	2.545900
FCI-Retrieval-DLM (apply=0)	0.010436	0.015530	0.012356	2.545900
FCI-Retrieval-DLM (proposed)	**0.010383**	**0.015411**	**0.012270**	2.552260
Shuffled-Retrieval Control	0.010398	0.015457	0.012345	**2.553044**

Setting	Trigger (%)	Apply (%)	Mean conf.
Base-DLM	0.000	0.000	0.000000
FCI-Retrieval-DLM (apply=0)	41.026	0.000	0.107913
FCI-Retrieval-DLM (proposed)	38.462	38.462	0.101543
Shuffled-Retrieval Control	50.000	50.000	0.147124

Notes: Trigger = candidate retrieval rate; Apply = gated intervention rate; Mean conf. = average matching confidence (triggered steps). Lower is better for CRPS/RMSE/MAE; higher is better for LPD.

Table 2. Event-window metrics for the same core ablation (window $w0_1$).

Setting	CRPS	RMSE	MAE	LPD
Base-DLM	0.008723	0.012587	**0.011180**	2.694646
FCI-Retrieval-DLM (apply=0)	**0.008616**	0.012587	**0.011180**	2.694646
FCI-Retrieval-DLM (proposed)	0.008618	**0.012570**	0.011229	2.695382
Shuffled-Retrieval Control	0.008641	0.012694	0.011355	**2.697144**

Notes: Event-window = $w0_1$ (days since onset in $\{0, 1\}$); interpreted as a decision-relevant post-announcement slice. Best value per metric is in bold (ties are jointly bold).

In all-scope evaluation (Table 1), enabling retrieval-based adaptation modestly improves distributional and point accuracy relative to Base-DLM (CRPS: 0.010383 vs 0.010522; RMSE: 0.015411 vs 0.015530; LPD: 2.5523 vs 2.5459). However, differences relative to Shuffled-Retrieval Control are small and mixed in sign, indicating near-parity between semantic and randomized retrieval at this scope.

Within the tightest onset window $w0_1$ (Table 2), improvements over Base-DLM are more pronounced on CRPS and RMSE (0.008618 vs 0.008723; 0.012570 vs 0.012587), consistent with the hypothesis that retrieval-based correction is most effective immediately after event onset. The gap relative to Shuffled-Retrieval Control remains narrow, suggesting that the current FCI-G context may not yet provide sufficient discriminative power for semantic matching. No formal significance test is reported; these differences are therefore interpreted as directional evidence.

Robustness to Event-Window Definition Table 3 evaluates sensitivity across wider onset neighborhoods ($w0_1$, $w0_3$, $w0_5$), where broader windows trade increased sample size against dilution of the immediate onset signal.

Table 3. Robustness to the event-window definition.

Panel A: Performance (CRPS and RMSE)

Window	CRPS (FCI-retrieval)	dCRPS (retrieval-shuf)	dCRPS (retrieval-Base)	RMSE (FCI-retrieval)	dRMSE (retrieval-shuf)	dRMSE (retrieval-Base)
w0_1	0.010262	-0.000001	-0.000178	0.015443	-0.000004	-0.000115
w0_3	0.008954	+0.000000	-0.000066	0.014632	-0.000002	-0.000060
w0_5	0.008741	+0.000000	-0.000055	0.013950	-0.000002	-0.000043

Panel B: Mechanism diagnostics (event-window only)

Window	FCI-Retrieval-DLM			Shuffled-Retrieval Control		
Window	Trigger (%)	Apply (%)	Mean conf.	Trigger (%)	Apply (%)	Mean conf.
$w0_1$	37.500	37.500	0.065082	37.500	37.500	0.065166
$w0_3$	18.750	18.750	0.032541	18.750	18.750	0.032583
$w0_5$	12.500	12.500	0.021694	12.500	12.500	0.021722

Notes: $w0_1/w0_3/w0_5$ = days 0–1/0–3/0–5 post-onset (policy-sensitive evaluation slices). "FCI-retrieval" = FCI-Retrieval-DLM; Base = Base-DLM. Differences are FCI-retrieval minus comparator.

Relative to Base-DLM, FCI-Retrieval-DLM remains directionally better on CRPS and RMSE across all three windows (e.g., ΔCRPS: -0.000178, -0.000066, -0.000055), while differences relative to Shuffled-Retrieval Control are near zero. Mechanism diagnostics (Panel B) show that trigger rate, apply rate, and matching confidence decline monotonically from $w0_1$ to $w0_5$, consistent with the onset-focused gating design. Overall, retrieval-based adaptation yields consistent directional gains over the unadjusted baseline, but semantic retrieval advantage beyond the shuffled placebo remains limited under the present configuration.

6 Conclusion

We propose a combined financial forecasting method to enhance the accuracy of short-term forecasts in contexts where policy events trigger abrupt shifts in market structure. We introduce the FCI-Retrieval-DLM model, an extension of Base-DLM that incorporates FCI-G-guided data retrieval and a forecast-adjustment mechanism during event periods. Empirical results show that the most noticeable improvements occur immediately after event onsets, while overall performance remains stable across the full sample. However, because results remain close to those of the Shuffled-Retrieval Control group, the true benefit of historical context matching is not yet clearly distinguishable under the current configuration. Future research will focus on enhancing the robustness of these findings and extending the framework to broader market settings.

References

1. Abadie, A., Diamond, A., Hainmueller, J.: Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. Journal of the American Statistical Association **105**(490), 493–505 (2010)
2. Ajello, A., Cavallo, M., Favara, G., Peterman, W.B., Schindler, J., Sinha, N.R.: A new index to measure U.S. financial conditions. FEDS Notes (2023)
3. Bojinov, I., Shephard, N.: Time series experiments and causal estimands: Exact randomization tests and trading. Journal of the American Statistical Association **114**(528), 1665–1682 (2019)
4. Bugni, F.A., Li, J., Li, Q.: Permutation-based tests for discontinuities in event studies. Quantitative Economics **14**(1), 37–70 (2023)
5. Cogley, T., Sargent, T.J.: Drifts and volatilities: monetary policies and outcomes in the post WWII US. Review of Economic Dynamics **8**(2), 262–302 (4 2005)
6. Davig, T., Gerlach, J.R.: State-dependent stock market reactions to monetary policy. International Journal of Central Banking (2006)
7. Davy, P., Miles, R.E.: Sampling theory for opaque spatial specimens. Journal of the Royal Statistical Society Series B: Statistical Methodology **39**(1), 56–65 (9 1977)
8. Gneiting, T., Raftery, A.E.: Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association **102**(477), 359–378 (2007)
9. Guo, H., Hung, C.H.D., Kontonikas, A.: The fed and the stock market: A tale of sentiment states. Journal of International Money and Finance **128** (11 2022)
10. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica **57**(2), 357 (3 1989)
11. Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting **15**(5), 559–570 (2000)
12. Kontonikas, A., MacDonald, R., Saggu, A.: Stock market reaction to fed funds rate surprises: State dependence and the financial crisis. Journal of Banking and Finance **37**(11), 4025–4037 (11 2013)
13. Lipsitch, M., Tchetgen, E.T., Cohen, T.: Negative controls: A tool for detecting confounding and bias in observational studies. Epidemiology **21**(3), 383–388 (2010)
14. Nadaraya, E.A.: On estimating regression. Theory of Probability and Its Applications **9**(1), 141–142 (1964)
15. Opschoor, A., van Dijk, D., van der Wel, M.: Predicting volatility and correlations with financial conditions indexes. Journal of Empirical Finance **29**, 435–447 (2014)
16. Pesaran, M.H., Timmermann, A.: Small sample properties of forecasts from autoregressive models under structural breaks. Journal of Econometrics **129** (2005)
17. Prado, R., Ferreira, M.A.R., West, M.: Time Series : Modeling, Computation, and Inference, Second Edition. Chapman and Hall/CRC (2021)
18. Primiceri, G.E.: Time varying structural vector autoregressions and monetary policy. American Economic Review **95**(3), 821–852 (2005)
19. Simonoff, J.S.: Jackknifing and bootstrapping goodness-of-fit statistics in sparse multinomials. Journal of the American Statistical Association **81**(396), 1005–1011 (1986)
20. Watson, G.S.: Smooth regression analysis. The Indian Journal of Statistics, Series A **26**(4), 359–372 (1964)
21. West, M., Harrison, J.: Bayesian Forecasting and Dynamic Models, Second Edition. Springer (1997)