

1. Background

Over the past decade, the Reserve Bank of India (RBI) has placed increasing emphasis on communication—via Monetary Policy Statements (MPS), Monetary Policy Committee (MPC) minutes, and governor speeches—to guide markets and anchor expectations. Prior work (e.g., Shrimali Ahmad 2025) shows that the “tone” of RBI’s MPS (measured by a custom “Net-Hawkish” dictionary) predicts the direction of the next policy rate move and influences equity-market volatility. However, until now:

- Most studies focus solely on the MPS text rather than the more detailed MPC minutes or governor speeches.
- The standard “Net-Hawkish” measure produces a discrete (hawkish vs. dovish) signal tied to rate shifts, but does not capture gradations of RBI’s risk aversion (degree of tightening) or willingness to provide “policy liberty” (degree of loosening) beyond a simple cut/hold/hike.
- There is limited evidence on how these richer communication channels map to real-time market indicators of risk (e.g., credit spreads, VIX-NSE) or to quantitative measures of liquidity and credit conditions.

2. Research Gap

Use of Minutes & Speeches:

MPS (released on policy days) are carefully crafted, but they omit much of the internal dialogue and dissenting viewpoints found in MPC minutes (released days later) and governor QA transcripts. Those sources often contain forward-looking nuances—“some members were concerned about inflation surprises” versus “some worried about growth headwinds.” We lack a systematic way to translate those finer shades into a continuous index of “risk aversion” or “policy liberty.”

Continuous Stance Measure:

Existing text-based indices (Net-Hawkish, LM dictionary, FS dictionary) yield either a normalized ratio of hawkish/dovish counts or a broad positive/negative sentiment score. None produces a continuous gauge of how “tightly” or “loosely” RBI is positioned along a spectrum. This matters because two “hawkish” statements might differ in intensity (“moderately hawkish” vs. “very hawkish”), with correspondingly different market and policy effects.

Linking to Market & Liquidity Outcomes

We know that a hawkish MPS tends to push up VIX or widen bond spreads, but it is unclear how signals contained in MPC minutes or governor speeches—once distilled into a “risk-aversion index”—track daily/weekly changes in (a) credit spreads, (b) swap spreads, (c) term-premium proxies, and (d) actual liquidity aggregates (e.g., M3, CD ratio). A more complete mapping between communication-derived stance and quantitative tightening/loosening variables does not yet exist for India.

3. Problem Statement

“Can we construct a continuous ‘RBI Risk-Aversion Index’—ranging from strongly dovish (policy-liberty) to strongly hawkish (risk-averse)—by mining the text of MPS, MPC minutes, and governor speeches? If so, to what extent does this index (1) predict the magnitude and direction of subsequent policy moves (e.g., repo-rate changes, CRR adjustments, OMO operations), and (2) correspond with real-time market indicators of risk aversion (e.g., VIX-NSE, credit spreads, swap spreads) and liquidity measures (e.g., M3 growth, CD ratio, call-money rates) in India?”

4. Research Objectives & Questions

Objective 1 (Text-Based Stance Construction):

- Build a custom, field-specific dictionary and algorithm to assign each policy episode (MPS, subsequent MPC minutes, and governor speeches) a continuous “Risk-Aversion Score” ($\text{StanceScore} \in [-1, +1]$, where $+1$ = extremely hawkish/risk-averse, -1 = extremely dovish/liberty-oriented).
- Refine existing “Net-Hawkish” methodology by adding new keywords/modifiers drawn from minutes and speeches (e.g., “some members expressed concerns over inflation” vs. “some worried about growth headwinds”).

Objective 2 (Policy Prediction):

- Test whether StanceScore_t (from the days of MPS/minutes/speeches) predicts the magnitude of the next repo-rate shift (in bps) or CRR change at meeting $t+1$ (and $t+2$).

- Compare predictive accuracy of the text-derived StanceScore to a baseline model that uses only quantitative variables (inflation gap, output gap, OIS-implied rate expectations).

Objective 3 (Market & Liquidity Linkages):

- Estimate daily regressions linking StanceScore_t to (a) $\Delta\text{VIX NSE}_t$, (b) $\Delta(10\text{Y G-sec yield})_t$, (c) $\Delta(5\text{Y AAA corporate} - 5\text{Y G-sec spread})_t$, (d) $\Delta(1\text{Y IRS} - \text{OIS})_t$, and (e) daily/weekly changes in M3 growth or CD ratio.
- Determine whether a more hawkish (positive) StanceScore systematically raises market risk indicators (e.g., vol spiked, spreads widened) and whether a dovish (negative) StanceScore compresses those indicators.

Objective 4 (Composite Monetary Condition Index):

- Combine StanceScore with key quantitative elements (ΔRepo , ΔCRR , ΔSLR , $\Delta\text{OIS spread}$, M3 growth) into a single “Indian Monetary Condition Index ($\text{MCI}_{\text{India}}$)” that continuously tracks tightening (positive values) versus loosening (negative values).
- Validate the $\text{MCI}_{\text{India}}$ by showing it (a) correlates with cyclical turning points in QGDP growth and CPI inflation, and (b) outperforms purely quantitative MCIs when predicting below-trend inflation or growth.

5. Methodology Overview

Text Collection & Cleaning (Weeks 1–2):

Collect MPS PDFs, MPC minutes PDFs, and full governor-speech transcripts (all RBI publications) from July 2010–October 2021. Strip headers/footers, convert to plain UTF-8 text, remove stopwords, normalize synonyms (e.g., “inflationary pressure” “inflation risk”).

Dictionary Construction & StanceScore Computation (Weeks 2–4):

Extend the existing RBI field-specific dictionary (36 keywords and hawkish/dovish modifiers) to include new minute/speech-specific terms (e.g., “downside risk,” “upside inflation pressure,” “growth headwinds,” “spillover risks”). For each document, count hawkish vs. dovish modifiers within ± 5 -word windows around each keyword; compute

$$\text{StanceScore}_t = \frac{\#\text{Hawkish}_t - \#\text{Dovish}_t}{\#\text{Hawkish}_t + \#\text{Dovish}_t}.$$

If multiple sources pertain to the same policy episode (e.g., MPS on t , Minutes on $t + 10$, Speech on $t+10$), either keep them separate or average them to form a composite StanceScore.

Orthogonalization (Optional Robustness, Weeks 4–5):

Regress raw StanceScore $_t$ on contemporaneous lagged inflation gap, output gap, repo $_{t-1}$, OIS $_{t-1}$, log(NSE $_{t-1}$), log(REER $_{t-1}$), and any global uncertainty indices. Use the residual $\varepsilon_{\text{Stance},t}$ as the “pure communication shock” that cannot be explained by macro/financial variables.

Policy Prediction Regressions (Weeks 5–6):

Dependent variable A: ΔRepo_{t+1} in basis points (continuous). Run OLS:

$$\Delta \text{Repo}_{t+1} = \alpha + \beta \text{StanceScore}_t + \gamma_1(\pi_{t-1} - \pi^*) + \gamma_2(y_{t-1} - y^*) + \delta \text{Repo}_{t-1} + \varepsilon.$$

Dependent variable B: Ordered (-1, 0, +1) classification of “cut/hold/hike” at $t + 1$. Run ordered-probit, with StanceScore $_t$ and standard Taylor-rule controls. Assess: Significance and economic magnitude of β . Does a +0.2 increase in StanceScore today predict a +20bp hike next meeting?

Market & Liquidity Regressions (Weeks 6–7):

Equity Volatility (VIX) Model:

$$\Delta \text{VIX}_t = \alpha + \beta \text{StanceScore}_t + \gamma |\Delta \text{OIS}_t| + \delta \Delta \text{VIX}_{t-1} + u_t.$$

Credit-Spread Model:

$$\Delta(5Y \text{ AAA}-5Y \text{ G-sec})_t = \alpha + \beta \text{StanceScore}_t + \gamma \Delta(10Y \text{ G-sec})_t + \phi \Delta \text{VIX}_{US,t} + \varepsilon_t.$$

Liquidity (M3/CD) Model (Weekly Frequency):

$$\Delta M3_w = \alpha + \beta \text{StanceScore}_w + \gamma \Delta(\text{CallRate})_w + \phi \Delta(\text{CRR})_w + \epsilon_w.$$

Interpretation: A positive β in the VIX model indicates that more hawkish language raises equity-market fear. A positive β in the credit-spread model implies that hawkish tone widens corporate spreads. A negative β in the M3 model would imply hawkish tone slows money-supply growth.

Composite MCI_{India} Construction (Weeks 7–8):

Standardize each component (repo gap vs. neutral, Δ CRR, Δ OIS spread, Δ M3, Δ CD ratio, and StanceScore) to z -scores. Combine via a simple average (or weighted average with weights derived from a principal-component analysis):

$$MCI_{India,t} = \frac{1}{6} [z(\text{RepoGap}_t) + z(\Delta\text{CRR}_t) + z(\Delta\text{OISSpread}_t) + z(\Delta\text{M3}_t) + z(\Delta\text{CDRatio}_t) + z(\text{StanceScore}_t)].$$

Plot MCI_{India} from mid-2010 to late-2021; identify peaks (tightening episodes) and troughs (easing episodes).

Validation & Robustness:

- **Out-of-Sample Test:** Calibrate the models on July 2010–December 2018; test on January 2019–October 2021.
- **Dictionary Variants:** Compare StanceScore against a simpler Net-Hawk dictionary and against a purely quantitative MCI (without the text component).
- **Linguistic Controls:** Include readability (Flesch) and lexical-diversity (CTTR) scores as regressors to ensure StanceScore is not just capturing language complexity.

6. Expected Contributions

- **Novel Continuous Stance Measure:** Deliver the first continuous “RBI Risk-Aversion Index” that seamlessly integrates MPS, MPC minutes, and governor speeches into a single number capturing policy “tightness” vs. “liberty.”
- **Policy Forecasting:** Show that this index significantly improves the prediction of repo-rate changes and other liquidity operations, beyond standard macro-financial controls.
- **Market Impact Mapping:** Provide empirical evidence that shifts in StanceScore lead (Granger-style) to changes in equity-market fear (VIX), credit spreads, swap spreads, and money-supply growth—validating that RBI language truly moves real-world market risk appetite.
- **Composite Monetary Condition Index (MCI_{India}):** Create an integrated, high-frequency gauge of India’s tightening/loosening stance that combines both quantitative policy variables and central-bank communication signals.

7. Significance & Next Steps

- **For Academics:** Fills a gap by incorporating richer RBI communication channels (minutes and speeches) into a continuous, text-based framework. Provides a blueprint for future work on emerging-market central bank communication.
- **For Practitioners (Traders, Economists, Credit Strategists):** Offers a real-time, text-augmented gauge of RBI's risk aversion—helping them anticipate rate moves, adjust bond positions, and price credit risk more dynamically.
- **Next Steps:** Once the index is constructed and validated, one could extend the analysis by (a) decomposing StanceScore into sub-dimensions (inflation-focused words vs. growth-focused words), or (b) applying a similar framework to other EM central banks (e.g., Brazil, South Africa) for cross-country comparison.