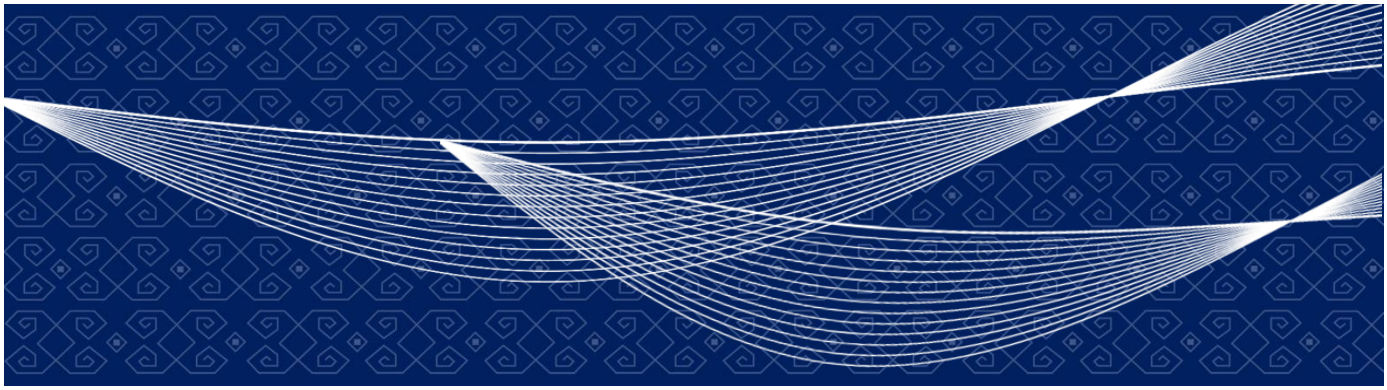
BANK INDONESIA
BANK SENTRAL REPUBLIK INDONESIA



WP/19/2025

WORKING PAPER

ECHOES OF POLICY: LEVERAGING AI/ML TO SUPPORT CENTRAL BANK COMMUNICATION STRATEGIES

Rudy Marhastari, Cicilia Anggadewi Harun, Retno Muhardini, Agatha Silalahi, Annes Nisrina Khoirunnisa, Rheznandya Arkaputra Azis, Sintia Aurida, Rahardian Luthfan Ihtifazhuddin, Citra Ayu Rossi Wulandari, Alvin Andhika Zulen, Amin Endah Sulistiawati

2025

ECHOES OF POLICY: LEVERAGING AI/ML TO SUPPORT CENTRAL BANK COMMUNICATION STRATEGIES

Rudy Marhastari, Cicilia Anggadewi Harun, Retno Muhardini, Agatha Silalahi, Annes Nisrina Khoirunnisa, Rheznandya Arkaputra Azis, Sintia Aurida, Rahardian Luthfan Ihtifazhuddin, Alvin Andhika Zulen, Amin Endah Sulistiawati, Citra Ayu Rossi Wulandari

Abstract

This study evaluates the effectiveness of Bank Indonesia's communication strategy by integrating computational linguistics, media sentiment analytics, and macroeconomic diagnostics within a unified empirical framework. Using advanced Natural Language Processing (NLP) techniques, BI's press releases from 2019–2024 are transformed into quantitative indicators capturing clarity, sentiment, comprehensiveness, consistency, and economic appropriateness. In parallel, news articles on inflation and exchange rate developments are analyzed to assess how policy messages are transmitted or amplified through media channels. These linguistic features are further enriched using Named Entity Recognition to identify stakeholder-specific resonance and potential pathways of narrative distortion within the public communication ecosystem. To assess macroeconomic implications, a VARX model links communication characteristics to intermediary channels, market expectations, and macroeconomic outcomes under both normal and anomalous conditions. Complementing this analysis, an Early Warning System (EWS) employing a 12-month rolling window and IsolationForest anomaly detection identifies periods of inflation and exchange-rate stress, providing a diagnostic foundation for anticipating heightened communication demands. The findings show that central bank communication functions not only as an information conduit but also as an active policy instrument that shapes expectations and influences market behavior. Building on these insights, the study proposes a three-pillar framework: Features, Timing, and Channels; to strengthen clarity, responsiveness, and coherence in central bank communication. This research advances the literature by integrating AI/ML-based diagnostics with policy communication analysis, offering an empirically grounded approach to enhancing communication effectiveness, transparency, and expectation management.

Keywords: Central Bank Communication, Communication Feature, Sentiment Analysis, Communication Impact Analysis, Early Warning System

1. Introduction

1.1 Background

Central bank communication has long been recognized as a critical instrument in supporting the effectiveness of monetary policy (Bernanke, 2007; Draghi, 2012; Woodford, 2005). Through well-crafted communication strategies, central banks can signal the direction of policy, provide insights into economic conditions, and share projections on inflation and interest rates, thereby shaping market expectations (Binder, 2017; Blinder et al., 2022). The clarity of messages and the consistency of conveyed information are key to maintaining policy stability and credibility (Mishkin, 2007; Svensson, 2010). Hence, central bank communication serves not only as a vehicle for information dissemination but also as the "echoes of policy" that influence public perception and economic decision-making more broadly.

The advancement of Artificial Intelligence (AI) and Machine Learning (ML), particularly Natural Language Processing (NLP) techniques, presents opportunities to overcome these challenges (Hansen & McMahon, 2016; Hansen et al., 2018; Marhastari, 2024). NLP methods can efficiently extract content from text, detect sentiment, identify keywords, and reveal hidden patterns within unstructured data (Ardia et al., 2019; BIS, 2025). These capabilities enable central banks to conduct more comprehensive assessments of the quality, consistency, and impact of their communication.

Beyond technical challenges, strategic considerations in central bank communication must also be addressed. The choice of communication strategy; such as the level of transparency versus ambiguity; can influence market reactions and macroeconomic stability (Dincer & Eichengreen, 2013; Ehrmann & Talmi, 2020). Furthermore, central banks should design key features and map out communication strategies consistently (Ehrmann & Talmi, 2020; Marhastari, 2024) to maintain information stability and prevent disinformation that may undermine institutional credibility (Ehrmann & Talmi, 2020; BIS, 2022). Therefore, efforts to maximize the resonance of policy messages through AI/ML technologies must be integrated with a deep understanding of the essence, objectives, and governance of central bank communication itself.

Periods of economic anomaly such as sudden exchange rate pressures, global financial turmoil, or abrupt inflation surges present additional challenges. During such conditions, the risk of misinterpretation increases, media narratives become more volatile, and public expectations can shift rapidly. Crisis communication frameworks such as the Situational Crisis Communication Theory (Coombs, 2007) highlight the importance of instructive, adjusting, and stabilizing narratives to maintain public confidence. Furthermore, the role of media as an intermediary is amplified during turbulent episodes, reinforcing the need for communication that is timely, targeted, and contextually grounded.

Against this backdrop, the availability of digital analytics; including Natural Language Processing (NLP), sentiment analysis, and anomaly detection systems; provides new opportunities for central banks to evaluate the effectiveness of their communication and anticipate risks. This study leverages linguistic modeling, media analysis, and dynamic visualization tools to assess the resonance and macroeconomic implications of Bank Indonesia's press releases, particularly during periods of economic anomaly.

1.2 Problem Statement and Research Question

Although Bank Indonesia consistently communicates its policy decisions through official press releases, several communication challenges persist. These include variability in message clarity and consistency across policy cycles; divergence between the tone intended in official communication and the tone interpreted and amplified by media narratives; limited empirical evidence on how linguistic features embedded in policy statements influence macroeconomic variables; and the absence of systematic

tools to detect and anticipate communication risks during periods of economic anomalies.

These challenges converge into the central problem of this study: to what extent do the linguistic characteristics and communication behaviors of Bank Indonesia shape media narratives as intermediary channels, influence public and stakeholder expectations, and ultimately affect macroeconomic outcomes, particularly under anomaly conditions? Based on this core problem, the study formulates the following research questions:

1. What are the linguistic characteristics of Bank Indonesia's press releases, and how do these features evolve over time?
2. How does media and stakeholders sentiment respond to Bank Indonesia's communication, and what forms of narrative drift or amplification arise in the process?
3. How do specific communication features; clarity, sentiment, comprehensiveness, consistency, and economic appropriateness; affect intermediary channels, stakeholder expectations, and macroeconomic outcomes such as inflation and the exchange rate?
4. Can an Early Warning System based on linguistic, communication, and macroeconomic indicators effectively detect anomaly periods when enhanced communication is required?
5. What communication strategies and channel configurations are most effective across different economic environments, particularly between normal and anomaly conditions?

1.3 Research Objective

Based on the research questions, this study aims to achieve the following objectives.

1. To identify and analyze the linguistic characteristics of Bank Indonesia's press releases and examine how these features evolve across different policy cycles and economic conditions.
2. To evaluate media sentiment dynamics and narrative transformations arising from the dissemination of Bank Indonesia's communication, including the detection of narrative drift, amplification, and tone divergence.
3. To assess the influence of key communication features, including clarity, sentiment, comprehensiveness, consistency, and economic appropriateness, on intermediary channels such as media tone, stakeholder responses, and macroeconomic outcomes, particularly inflation and the exchange rate.
4. To develop and test an Early Warning System (EWS) that integrates linguistic, communication, and macroeconomic indicators to detect emerging anomaly periods that require enhanced or adaptive communication strategies.
5. To formulate effective communication strategies and channel configurations tailored to different economic environments.

1.4 Research Gap and Novelty

Despite the extensive literature on central bank communication, several gaps remain, particularly in the Indonesian context. First, existing studies tend to focus on general communication principles without systematically quantifying linguistic characteristics and linking them to macroeconomic variables or media behavior. Second, there is limited empirical research that distinguishes central bank communication performance under normal and anomaly conditions, even though communication needs differ significantly between the two regimes. Third, few studies incorporate multi-layered analytical tools; combining NLP-based feature extraction, media sentiment analysis, impact analysis modeling, and Early Warning Systems; to provide a holistic evaluation of communication effectiveness.

This research contributes to filling these gaps through several innovations:

1. Linguistic–Macroeconomic Integration:
   The study links linguistic features of BI press releases with inflation, exchange rate movements, media sentiment, and stakeholder expectations, providing an integrated empirical framework.
2. Empirical Evaluation of Media Resonance:
   Through sentiment mapping and NER-based stakeholder analysis, the study uncovers how different audience segments interpret and respond to BI's communication.
3. Anomaly-Specific Communication Analysis:
   The research differentiates communication effects of anomaly periods, offering insights on how central banks should adapt their messaging during stress conditions.
4. Development of a Communication-Based Early Warning System:
   Using signaling approach, the study constructs macroeconomic anomaly detection dashboard that supports risk-sensitive communication planning.
5. Integration of Communication Channel Strategy (ATL–BTL–TTL):
   The study combines communication theory with empirical findings to design a comprehensive channel strategy tailored for central bank contexts.

Together, these contributions provide a novel and comprehensive framework for evaluating, optimizing, and modernizing central bank communication.

2. Literature Review

2.1 Central Bank Communication

In recent monetary policy frameworks, communication has become a strategic instrument alongside conventional tools such as interest rate policy and open market operations. Beyond conveying policy decisions, communication guides expectations and strengthens monetary transmission by shaping how households, firms, and financial markets interpret economic developments (Blinder et al., 2008; Woodford, 2005). Well-anchored expectations are essential for macroeconomic stability, as credible central banks reduce the likelihood that economic agents react excessively to short-term shocks (Bernanke, 2007). In open economies such as Indonesia, effective communication also helps stabilize external perceptions, capital flows, and exchange rate dynamics by reducing uncertainty and speculative behavior (Reis, 2013).

Expectations are the core of modern macroeconomic behavior, influencing decision on economic activities, such as consumption, investment, and pricing (Mishkin, 2007). Because monetary policy operates with lags, its effectiveness depends critically on how well agents anticipate future conditions. Expectations are actively shaped by policy credibility, ongoing communication, and macroeconomic signals. When expectation become unanchored, inflation can rise due to pre-emptive adjustments by firms and workers, even in the absence of real economic pressures (Svensson, 2010). Conversely, if public trust in central bank ability to contain inflation deteriorates, even minor shocks may trigger excessive market reactions and macroeconomic instability (Carney, 2019). Consequently, managing expectations is as important as implementing policy itself.

Effective central bank communication depends not on frequency alone, but on specific characteristics that determine its influence on expectations, trust, and policy transmission. To shape expectations effectively, communication must adhere to certain principles, such as clarity, consistency, and coherency. Clarity ensures messages are accessible and interpretable with minimal ambiguity (Casiraghi & Perez, 2022; Marhastari, 2024), while comprehensiveness provides the underlying rationale, framework, and transmission mechanisms needed for agents to form accurate expectations (Simorangkir, 2014). Sentiment awareness—the alignment of tone and emotional framing with macroeconomic objectives—further enhances credibility and

stabilizes expectations. Together, these characteristics shape predictable and stable behavioral responses.

Central bank communication also varies by type and audience. Public communication targets broad audiences through press conferences, official statements, and digital platforms, while segmented communication addresses specific groups such as analysts or market participants via closed meetings or focused discussions (Bholat, 2019; Liu, 2022). A further distinction exists between formal communication structured and institutionalized outputs and informal communication, such as interviews or unscheduled remarks, which offer flexibility and timeliness (Haan, 2019). Relatedly, the literature distinguishes between broadly accessible "above-the-line" communication and selective "below-the-line" messaging, underscoring the need for tailored strategies that account for heterogeneous information processing among economic agents.

Beyond its formats and characteristics, the strategy behind central bank communication determines its real-world impact. One of the most critical strategies involves simplification and relatability, which are the ability to explain complex economic decisions in terms that can be understood by the general public (Bholat, 2019). In more dynamic context, central banks may shift their strategy to emphasize current conditions over future projections. (Gati, 2022) argues that in volatile environments, focusing on the present can reduce the uncertainty associated with forward-looking signals. Similarly, during financial crisis, central banks often adopt crisis-specific strategies, such as justification, differentiation, and corrective messaging, to manage institutional reputation and preserve market stability (Gyamfi, 2024). Ultimately, effective central bank communication must remain adaptive, coherent, and firmly grounded in institutional credibility and the broader monetary policy framework.

2.2 Global Practices in Central Bank Communication

Central bank communication has undergone a profound transformation over the past three decades, shifting from a paradigm of secrecy toward one of strategic transparency. Historically, ambiguity was regarded as beneficial for preserving policy flexibility and limiting market overreaction, resulting in minimal public disclosure by institutions such as the Federal Reserve and the Bank of England prior to the 1980s. This approach began to change with the emergence of inflation-targeting regimes in the late 1980s and early 1990s, pioneered by the Reserve Bank of New Zealand in 1990, followed by the Bank of England's institutionalization of communication tools such as Inflation Reports and the publication of MPC minutes after gaining operational independence in 1997 (King, 2005). The establishment of the European Central Bank in 1998 further entrenched transparency norms through systematic communication, including regular bulletins and press conferences. The Federal Reserve, initially adhering to "constructive ambiguity" under Alan Greenspan, progressively modernized its communication framework in the 2000s under Ben Bernanke, introducing the Summary of Economic Projections, regular press conferences, and formally adopting an inflation target in 2012 (Bernanke, 2013). In Asia, communication reforms were more gradual: the People's Bank of China expanded external communication in response to deeper global financial integration despite limited institutional independence (RBA, 2019), while the Bank of Japan shifted from vague statements toward more structured guidance through Outlook Reports and post-meeting press conferences, with leadership changes shaping communication clarity and tone (CEPR, 2021). Collectively, these developments reflect a historical shift in which communication has evolved from a passive policy byproduct into a core instrument of macroeconomic management.

Despite a shared emphasis on communication, central banks differ in how messaging authority is structured, broadly following collegial or individualistic models. The Federal Reserve illustrates a hybrid approach, combining collective decision-making within the FOMC with individual public communication by its members, a practice that enhances transparency but risks message fragmentation (Ehrmann & Fratzscher, 2007). To limit this risk, central banks impose institutional constraints, such as blackout periods around policy meetings at the Fed and the Bank of England, or the ECB's

preference for highly collegial communication that discourages publicly visible dissent despite some differentiation in press conferences (Issing, 2005). These arrangements reflect a trade-off between transparency and coherence: uniform messaging reduces uncertainty but may conceal internal debate, while individual expression strengthens accountability yet can undermine credibility if left unchecked. Beyond formal rules, communication also varies in tone and accessibility, with the Bank of England encouraging MPC members to explain their votes, and the Bank of Japan enhancing clarity under Governors Kuroda and Ueda, although markets still find its long-term policy stance difficult to interpret (Japan Times, 2025; Reuters, 2024).

Crises often force central banks to move beyond routine communication and adopt more proactive, emotionally responsive, and multifaceted strategies, as illustrated by the Global Financial Crisis and the Covid-19 pandemic. During these periods, central banks relied on structured messaging frameworks informed by theory and experience, including the Situational Crisis Communication Theory adapted for central banking, which emphasizes sequential messaging through justification, differentiation, and corrective action (Gyamfi, 2024), as well as the Three-Tiered Crisis Communication Model that distinguishes immediate response, narrative management, and expectation reorientation. Empirical evidence from the Federal Reserve shows a marked increase in communication frequency, greater reliance on press briefings, and heightened visibility of the chair, while the ECB adopted a more assertive tone, most notably in Mario Draghi's "whatever it takes" speech in 2012, demonstrating how credible, sentiment-driven communication can stabilize markets and restore confidence (De Grauwe & Ji, 2013).

3. Methodology and Data

3.1 Methodology

This study applies an integrated methodological framework to analyze Bank Indonesia's communication strategies using Natural Language Processing (NLP). Press releases and inflation- and exchange rate–related news articles from 2019–2024 are analyzed to quantify communication quality in terms of clarity, completeness, economic appropriateness, consistency, and sentiment. These indicators are combined with Named Entity Recognition (NER) to map public perception and identify message resonance or distortion. By integrating computational linguistics with established central bank communication theories, the framework ensures both analytical rigor and policy relevance. Furthermore, Communication impact analysis is conducted using a machine learning–based approach to evaluate how central bank communication shapes expectations and influences economic behavior. The analysis links communication features, clarity, consistency, sentiment, comprehensiveness, and economic appropriateness, with intermediary channels (credibility and media sentiment), market expectations, and macroeconomic outcomes such as inflation and exchange rates. An Early Warning System (EWS) is then developed to detect potential anomalies by monitoring exchange rate volatility and inflation.

3.1.1 Methodology of Press Release Communication Feature Analysis

Natural Language Processing (NLP) constitutes the analytical core of this study by transforming qualitative communication attributes into quantifiable indicators. Through standardized preprocessing and linguistic modeling, NLP enables objective assessment of clarity, comprehensiveness, sentiment, economic appropriateness, and consistency in Bank Indonesia's press releases and related news coverage. The resulting standardized dataset ensures analytical reliability and supports robust feature extraction. This section details the methodology for evaluating the linguistic and thematic dimensions of central bank communication, with each core feature operationalized using quantitative measures grounded in the monetary communication and computational linguistics literature.

3.1.1.1 Clarity

Clarity represents the extent to which communication is linguistically accessible while maintaining technical precision. It combines readability and economic content density into a composite score.

The Indonesian Flesch–Kincaid Readability Score, denoted as F, is calculated using the formula:

$$F = 206.835 - 1.015 \times ASL - 84.6 \times ASW$$

where ASL is the Average Sentence Length (total words divided by total sentences), and ASW is the Average Syllables per Word (total syllables divided by total words). Both are adjusted for Indonesian morphological structures as validated by Nababan (1991). A higher readability score indicates a text that is easier to comprehend. The Economic Density (ED) reflects the proportion of economic terms in the text, computed as:

$$ED = \frac{\sum_{i=1}^{n} w_i f_i}{N}$$

where $f_i$denotes the frequency of the i^{th}economic term, $w_i$represents the thematic weight derived from the validated economic lexicon, and $N$is the total word count.

To obtain a unified clarity measure, both readability and economic density are normalized and combined into a Composite Clarity Score (CS) using the formula:

$$CS = 0.7 \times F_{norm} + 0.3 \times ED_{norm}$$

The weighting scheme (70:30) follows Blinder et al. (2008) and Hansen and McMahon (2016), who emphasize that linguistic clarity accounts for a larger portion of effective communication, while technical precision remains necessary to maintain policy credibility.

3.1.1.2 Economic Appropriateness

Economic appropriateness assesses whether the balance between linguistic simplicity and technical complexity aligns with the expected communication standards of a central bank. It is constructed as a two-step measure that integrates readability and economic density in relative terms.

The Base Appropriateness Score (BA) is derived from the normalized readability score, where higher readability corresponds to higher accessibility. This base score is then adjusted using the position of the document's economic density within the corpus distribution:

$$EA = BA + \Delta_Q$$

where $\Delta_Q$denotes the quartile-based adjustment determined by the economic density's placement within the corpus. Documents within the interquartile range (Q1–Q3) receive positive adjustments, as they exhibit balanced levels of linguistic and technical content. Conversely, documents below Q1 or above Q3 are adjusted downward. This approach is adapted from Bulíř, Čihák, and Jansen (2013), who highlight that both excessive simplicity and over-technicality can undermine communication effectiveness.

3.1.1.3 Comprehensiveness

Comprehensiveness measures how extensively and deeply each press release covers relevant policy issues and provides contextual reasoning. The index incorporates three subcomponents—topic coverage, content depth, and information completeness—aggregated into a single weighted indicator:

$$COMP = 0.4 \times TC + 0.35 \times CD + 0.25 \times IC$$

The weighting rationale follows Hansen and McMahon (2016) and Dincer et al. (2019), where broader topic coverage is prioritized as the primary driver of comprehensiveness, followed by analytical depth and structural completeness.

Topic coverage (TC) quantifies the proportion of thematic categories from the economic lexicon represented in the document, ensuring that a broad range of relevant policy areas are addressed, including those related to monetary policy, inflation, exchange rates, and financial stability. *Content depth (CD)* captures the frequency of analytical expressions and reasoning markers normalized by sentence length, while

information completeness (IC) evaluates whether structural sections such as background context, data discussion, and policy analysis are present. A higher comprehensiveness score reflects the ability of communication to convey both breadth and analytical substance.

3.1.1.4 Sentiment

The sentiment feature measures the tonal orientation of Bank Indonesia's policy communication and captures whether the conveyed message reflects optimism, neutrality, or caution. This feature adopts a hybrid modeling framework that combines a lexicon-based economic sentiment model and a transformer-based contextual sentiment model to ensure both interpretability and contextual depth. The final sentiment consensus score integrates both components as:

$$S_{\text{consensus}} = \alpha S_{\text{lexicon}} + (1-\alpha) S_{\text{ML}}$$

where α=0.5 ensures balanced weighting between lexical precision and contextual interpretation. The resulting score is categorized as positive, neutral, or negative using symmetric thresholds of ±0.05. These thresholds are not derived directly from any specific literature but are calibrated to reflect small yet meaningful tonal shifts in the text, consistent with the conceptual framework advanced by Shapiro and Wilson (2022), who emphasize that even subtle changes in communication tone can significantly affect market expectations.

This hybrid approach thus aligns with recent developments in text-based analysis of monetary policy communication. By integrating rule-based and contextual perspectives, the sentiment feature captures both explicit emotional tone and implicit linguistic cues, allowing for a more nuanced and empirically grounded assessment of the Bank's communication stance.

3.1.1.5 Consistency

Consistency represents the stability and coherence of policy communication over time, reflecting the credibility of the institution. It is evaluated through two complementary methods: lexical consistency and semantic consistency. Lexical consistency measures vocabulary overlap between documents using Cosine Similarity on Term Frequency–Inverse Document Frequency (TF-IDF) vectors:

$$C_{lexical} = \frac{\vec{A} \cdot \vec{B}}{\| \vec{A} \| \times \| \vec{B} \|}$$

where $\vec{A}$and $\vec{B}$ represent vectorized word frequency profiles of two documents. Higher similarity values indicate greater consistency in terminology use.

Semantic consistency evaluates the similarity in meaning even when different wording is employed. It uses Soft Cosine Similarity, which considers semantic relationships between words:

$$C_{semantic} = \frac{\sum_{i,j} S_{ij} A_i B_j}{\sqrt{\sum_{i,j} S_{ij} A_i A_j} \times \sqrt{\sum_{i,j} S_{ij} B_i B_j}}$$

where $S_{ij}$is the semantic similarity between words $i$and j, and $A_i, B_j$are their corresponding vector representations. The Overall Consistency Score (CON) is obtained as:

$$CON = \frac{C_{lexical} + C_{semantic}}{2}$$

This integrated measure captures both lexical and conceptual coherence, ensuring that while expressions may vary, the fundamental policy message remains consistent across time.

3.1.2 Methodology of Media and Stakeholder Sentiment Analysis

The analysis of news-based sentiment serves as a complementary approach to the evaluation of Bank Indonesia's communication effectiveness. The methodology consists of three main components: (i) hybrid sentiment analysis, (ii) stakeholder identification using Named Entity Recognition (NER), and (iii) sentiment–stakeholder integration. These stages are designed to capture both the tone and the agent of sentiment in a structured, data-driven manner, enabling the analysis of resonance patterns between institutional communication and media narratives.

This produces a sentiment distribution across stakeholder groups, revealing asymmetries in perception between government actors, financial regulators, market analysts, and the general public. Consistent with Hansen et al. (2018) and Thorsrud (2020), these findings help identify how external interpretations of policy signals may amplify or distort the intended messages of central bank communication.

3.1.3 Methodology of Communication Impact Analysis

The purpose of communication impact analysis is to assess how effectively a policy message especially in formal communication documents shapes expectations and influences behaviors in both the market and the real economy. An effective policy message not only delivers information in a clear manner but also produces measurable effects on how the public and market participants respond. To capture this, the study employs a combination of quantitative approaches and machine learning techniques, allowing for an evaluation of communication impacts across different timeframes and complexities of reaction.

Vector Autoregressive Exogenous variables (VARX) models are employed to analyze the short- and long-term effects of policy communication on key macroeconomic indicators such as inflation, exchange rates, and inflation expectations.

3.1.4 Methodology of Early Warning System

In central bank communication, building an Early Warning System aims to quickly detect potential risks or economic instability caused by market perceptions of the information or policies being communicated. Specifically, in relation to the third research question, this study focuses on the development of an Early Warning System (EWS) designed to detect anomaly characterized by the depreciation of the Rupiah exchange rate and rising inflation.

a. Inflation

In this study, an analysis was conducted to compare actual inflation anomaly periods with various indicators using several methods, including standard deviation, the Kalman filter, and the Chow test, with the Isolation Forest identified as the most effective. The anomaly detection process using the Isolation Forest algorithm is carried out in two main stages, the training stage and the evaluation stage.

Overall, the implementation of the Isolation Forest algorithm in this study proved effective in detecting anomalies in economic time series data. The method successfully identifies periods of deviation from historical patterns, serving as an early warning signal for potential inflationary pressures. With its computational efficiency and flexibility across various types of data, Isolation Forest is considered a relevant approach for developing an early warning system to monitor inflation dynamics.

b. Rupiah Exchange Rate

One of the key processes in developing an early warning system or other detection systems is the accurate identification of anomalies in recorded data. To identify periods during which the rupiah exchange rate exhibits abnormal movements, this study employs a univariate statistical approach based on a dynamic threshold (rolling threshold) (Alnami et al., 2025). This method is selected for its simplicity, interpretability, and compatibility with the early warning system framework, which relies on binary signaling. The same principle is applied to the indicators designated as early warning

signals. Anomaly detection for each leading indicator utilizes the same thresholding method as that used for the exchange rate.

3.1.4.1 Correlation Test

The lead-lag correlation analysis was conducted to examine the temporal relationship between the anomaly indicator and its variables, which reflect public perception or sentiment dynamics regarding movements in the exchange rate and inflation. Technically, the indicator's variable was lagged up to six previous months ($t-1\ to\ t-6$) to construct lagged time series features. In the following correlation analysis, the Pearson correlation serves as the analytical foundation. The resulting correlation matrix represents the strength and direction of the linear relationship between the Inflation and Rupiah exchange rate to each lag of the indicator variable.

3.1.4.2 Probit Test

This analysis employs a binary probit model to estimate the probability of exchange rate anomalies based on changes in leading indicators over several preceding periods. The probit model is chosen because it is appropriate for data with a dichotomous dependent variable (anomaly and non-anomaly) and is based on the assumption that the probability of occurrence follows a cumulative normal distribution function. Mathematically, the model represents the relationship between a latent variable that reflects exchange rate pressures and the observed indicators, thereby allowing the estimation of the likelihood of exchange rate stress in subsequent periods. The model parameters are estimated using the Maximum Likelihood Estimation (MLE) method, and the statistical significance of the coefficients is evaluated using the z-statistic and p-value.

3.1.4.3 Signaling Approach

In an Early Warning System (EWS), determining the appropriate crisis signal is a crucial step. To transform predicted probabilities into accurate signals, an EWS applies a specific threshold value. In this study, the threshold used to categorize an event as a signal is defined as isolation forest and the rolling mean plus one standard deviation (rolling mean + 1 std). Based on this threshold, the system will issue an alarm whenever the predicted pre-crisis probability exceeds the cut-off value (τ). The compatibility between the signals generated by the EWS and the actual outcomes constitutes what is known as the contingency matrix. From this matrix, various measures of the EWS's predictive performance can be derived.

3.2 Data

This study utilizes a several complementary datasets that collectively support the overall analytical framework. The data include communication policy data (press releases), expectation data (consumer surveys), coverage data (news articles and Google Trends), credibility indices (monetary, macroprudential, and payment system policy), and macroeconomic data (inflation and exchange rate). These diverse data sources enable a comprehensive analysis of how communication impacts macroeconomic outcomes.

Table 3.31. Data

Together, these datasets enable a structured and measurable analysis of two complementary aspects of policy communication how Bank Indonesia formulates and conveys its policy messages formally through press releases, and how these messages are subsequently received, interpreted, and disseminated across media channels and stakeholder groups. By linking both the institutional voice and public perception dimensions, the data framework allows for an integrated understanding of communication resonance in the context of monetary policy.

4. Result and Analysis

This chapter summarizes the study's key results across three analytical stages. The first stage examines the linguistic and semantic features of Bank Indonesia's press releases, focusing on five dimensions, clarity, comprehensiveness, sentiment, economic appropriateness, and consistency. These indicators provide a structured overview of how policy messages are formulated and how their characteristics evolve over time. The second stage analyzes how these communications resonate externally by assessing media and stakeholder sentiment. It highlights how institutional messages are transmitted, reframed, or amplified, revealing patterns of narrative drift and the broader interpretive dynamics surrounding monetary policy communication. The third stage offers preliminary evidence on how communication attributes relate to market expectations and selected macroeconomic indicators. The analysis points to potential links between clarity, tone, and consistency with movements in inflation expectations and exchange rates, setting the foundation for deeper econometric investigation.

Figure 4. 1 Integrated Result Framework Linking Communication Features, Media Sentiment, and Econometric Impact Analysis

4.1 Press Release Communication Feature Analysis

An analysis of seventy-two Bank Indonesia press releases from 2019–2024 provides a quantitative overview of five key communication dimensions: clarity, comprehensiveness, economic appropriateness, sentiment, and consistency. The time-series approach tracks how these features evolve, showing how the institution formulates and adjusts its policy messages in response to shifting economic conditions while preserving deliberate consistency, including during the COVID-19 period and subsequent recovery.

4.1.1 Clarity

Clarity is constructed from two main components: linguistic readability and economic density. The readability dimension is measured using the Indonesian Flesch–Kincaid formula, which captures sentence length and word complexity to evaluate linguistic accessibility. Economic density represents the proportion of economic and financial terms contained in the text, as identified through the economic lexicon of 14 categories. Together, these elements produce a composite clarity score that reflects the balance between linguistic transparency and technical precision.

Figure 4.2 presents the time-series trend of the overall clarity score from 2019 to 2024. This visualization provides an overview of how the readability and linguistic precision of Bank Indonesia's press releases evolved over time, serving as a basis for identifying shifts or stability in communication style.

Figure 4. 2 Time-Series Trend of Overall Clarity Scores (2019–2024)

The composite clarity score, which combines readability and linguistic appropriateness, remains consistently stable across the 2019–2024 period, with an overall mean of 66.5. The Indonesian Flesch–Kincaid readability score averages 56.8, indicating a moderate reading difficulty suitable for readers with an advanced or professional background. Meanwhile, the average economic density reaches 9.6 percent, with a narrow range between 8.5 and 11.2 percent, suggesting a high and uniform concentration of economic terminology.

Over time, the clarity feature exhibits remarkable steadiness, with the three-document rolling average maintaining a near-flat trajectory throughout the five-year period. The limited fluctuation of individual document scores implies that Bank Indonesia consistently employs a well-structured and deliberate communication strategy. The institution prioritizes technical accuracy and terminological precision, maintaining a writing style that is accessible to policy professionals and analysts without

sacrificing depth or formality. This enduring clarity reinforces its credibility and reinforces communication effectiveness among specialized audiences.

4.1.2 Comprehensiveness

Comprehensiveness is built upon three core subcomponents: topic coverage, content depth, and information completeness. Topic coverage measures the breadth of policy themes addressed within each press release, derived from the 14-category economic lexicon. Content depth quantifies the frequency of analytical reasoning or explanatory statements that enhance interpretive richness, while information completeness evaluates the presence of contextual sections such as background, policy rationale, and data interpretation. These metrics are aggregated into a single weighted index that reflects both scope and analytical depth.

Figure 4.3 presents the time-series trend of the comprehensiveness score between 2019 and 2024. The visualization offers an overview of how the breadth and informational depth of Bank Indonesia's press releases have changed over time, providing a contextual basis for interpreting variations in policy coverage and thematic scope.

Figure 4. 3 Time-Series Trend of Economic Appropriateness Scores (2019–2024)

The average comprehensiveness score of 49.4 indicates that while Bank Indonesia's press releases generally provide relevant policy content, the depth and breadth of coverage vary considerably across documents. This variability reflects differences in topic scope, with some releases focusing narrowly on operational updates, while others deliver broader analytical context linking policy rationale to macroeconomic conditions.

The time-series visualization highlights a relatively high degree of volatility in this feature compared to clarity. The three-document rolling average tends to hover below the 50-point mark but displays alternating peaks and dips over time. Periods of broader coverage typically correspond to major monetary policy announcements, while lower scores are observed in short or event-specific releases. This variation suggests that comprehensiveness is context-dependent, responding to the thematic scope of each communication rather than a fixed institutional standard. Such adaptive flexibility ensures relevance and timeliness but may occasionally allow room for interpretive gaps among general audiences.

4.1.3 Economic Appropriateness

Economic appropriateness combines normalized readability and economic density to assess whether each communication achieves the right balance between accessibility and technical sophistication. The composite measure adjusts for documents that fall within or outside the interquartile range of the corpus' density distribution, ensuring an optimal linguistic–technical equilibrium.

Figure 4.4 presents the time-series trend of the economic appropriateness score between 2019 and 2024. The visualization illustrates how effectively Bank Indonesia's press releases balance linguistic simplicity with technical precision, reflecting the institution's effort to ensure that policy communication remains both analytically rigorous and accessible to its intended audience.

Figure 4. 4 Time-Series Trend of Economic Appropriateness Scores (2019–2024)

Economic appropriateness, which measures the balance between linguistic simplicity and technical sophistication, records an average score of 75.9, with values largely ranging between 65 and 88. This consistently high score confirms that Bank Indonesia's communication achieves an optimal equilibrium between accessibility and analytical rigor.

The temporal pattern demonstrates minimal variation, with the three-document rolling average line remaining steady throughout the sample period. The close clustering

of document-level scores suggests that the institution applies a uniform communication standard calibrated to its primary target audience of market participants and financial analysts. This linguistic consistency aligns with international benchmarks for effective central bank communication, emphasizing clarity without oversimplification.

4.1.4 Sentiment

Sentiment is quantified through a hybrid consensus score derived from two complementary approaches: a lexicon-based economic sentiment model and a transformer-based contextual sentiment model. The lexicon model assesses the frequency and polarity of sentiment-bearing terms, while the machine learning model (based on contextual embeddings) captures tone and nuance beyond explicit wording. The final sentiment score represents a weighted average between these two components, ensuring both interpretability and contextual sensitivity.

Figure 4.5 shows the time-series trend of the sentiment consensus score from 2019 to 2024. This visualization provides an overview of the tone dynamics within Bank Indonesia's press releases, depicting how the balance between positive, neutral, and negative expressions evolved over time. It also helps identify how the overall communication mood responded to shifts in macroeconomic conditions and policy contexts.

Figure 4. 5 Time-Series Trend of Sentiment Consensus Scores (2019–2024)

The sentiment consensus score averages 0.713 on a 0–1 scale, confirming a predominantly positive tone in Bank Indonesia's policy communication. This finding is consistent with the institution's strategic emphasis on confidence-building and policy reassurance.

The time-series trend reveals that sentiment remains consistently above the neutral midpoint of 0.5 across all periods, reflecting a sustained tone of optimism even during challenging economic conditions. A slight decline appears in early 2020, coinciding with heightened uncertainty at the onset of the COVID-19 pandemic. However, the sentiment quickly recovers in subsequent periods, reinforcing the effectiveness of Bank Indonesia's communication in maintaining market confidence and anchoring expectations. This persistence of positive sentiment across time underlines the central bank's role in guiding narratives of stability and resilience in the public domain.

4.1.5 Consistency

Consistency is assessed through two complementary metrics: lexical similarity and semantic similarity. Lexical consistency is computed using TF–IDF cosine similarity to evaluate the overlap of key terms across documents, while semantic consistency applies the Soft Cosine Similarity to capture conceptual coherence even when different wording is used. These two measures are averaged into an overall consistency index that reflects the stability of the institution's narrative over time.

Figures 4.6 and 4.7 display the time-series trends for lexical and semantic consistency scores across the 2019–2024 period. These visualizations jointly demonstrate the degree of uniformity in Bank Indonesia's policy narratives, highlighting the extent to which linguistic coherence and semantic alignment are maintained across successive press releases. Together, they provide a comprehensive view of how consistent wording, phrasing, and thematic framing contribute to the overall stability and credibility of institutional communication.

Figure 4. 6 Time-Series Trend of Lexical Consistency Scores (2019–2024)

Figure 4. 7 Time-Series Trend of Semantic Consistency Scores (2019–2024)

Consistency stands out as the most stable feature among all five dimensions. The average semantic similarity score reaches 0.922, while lexical similarity averages 0.864. These values indicate a high degree of uniformity in both language and meaning across all press releases.

Throughout the 2019–2024 period, the trend line for consistency remains nearly flat, and document-level variations are minimal. This stability suggests that Bank Indonesia's communication maintains a coherent policy narrative regardless of short-term economic developments or thematic shifts in content. The persistence of linguistic and conceptual consistency strengthens the institution's credibility and fosters trust among stakeholders, reinforcing its role as a steady anchor for market expectations.

4.1.6 Robustness Test Result

The robustness analysis confirms that the indicators of communication features are statistically sound and methodologically valid. The summary of the results is presented in Table 4.1.

Table 4.1 Robustness Test Result

The correlation results indicate that document length does not systematically affect the clarity, comprehensiveness, or economic appropriateness scores, with all correlations close to zero. These results confirm that the metrics are not structurally biased and effectively measure intrinsic linguistic characteristics. The strong correlation between the overall clarity and readability components ($\rho = 0.970$) and the moderately strong relationship for comprehensiveness ($\rho = 0.753$) validate the internal coherence of these composite indicators. This suggests that the metrics successfully capture the conceptual intent of each feature.

For consistency, the low correlation between lexical and semantic measures ($\rho = 0.214$) confirms that the two indicators reflect distinct yet complementary aspects of textual stability: the first focusing on word choice repetition, and the second on conceptual alignment across documents. The Fleiss' Kappa result of 0.4086 demonstrates moderate agreement among human annotators, consistent with the interpretive range established by Landis and Koch (1977). This confirms that the expert-labeled sentiment data provide a reliable benchmark for validating model-generated sentiment classifications. Overall, the robustness tests establish that the communication feature indicators are stable, coherent, and reliable. These results strengthen the validity of the study's findings and support the use of these indicators as credible tools for assessing the effectiveness of Bank Indonesia's communication strategies over time.

The consistent linguistic patterns observed in Bank Indonesia's press releases demonstrate a deliberate and well-structured communication approach that prioritizes clarity, coherence, and credibility. These internal communication features serve not only as indicators of policy transparency but also as signals transmitted to the broader information environment. To understand how these institutional messages are interpreted, amplified, or reshaped once they reach the public domain, the following section extends the analysis to the media and stakeholder level, examining how sentiment and narrative responses evolve in reaction to Bank Indonesia's communication signals.

4.2 Media and Stakeholder Sentiment Analysis

The sentiment analysis of online media provides an empirical basis for understanding how Bank Indonesia's communication is framed across the public information space. From over 190,000 news articles published between 2019 and 2024, a stratified sample of 7,000 articles was generated to maintain proportional sentiment representation 3,492 positive, 2,161 negative, and 1,347 neutral, ensuring balanced modeling and preserving the structure of the full dataset. The coverage is predominantly positive, with favorable articles highlighting economic stability, monetary discipline, and

effective inflation control. Negative sentiment tends to emerge during episodes of external volatility, such as commodity price shocks or exchange-rate pressures, and is often driven by market commentary and analyst criticism. Neutral sentiment largely reflects factual, event-driven reporting such as meeting summaries or policy announcements. Each sampled article was then processed using an integrated sentiment–stakeholder identification framework, linking sentiment polarity to specific entities mentioned in the text. This enables a structured classification of sentiment across stakeholder groups, including government institutions, regulators, analysts, businesses, and the public. The aggregated results establish a foundation for examining how sentiment varies across stakeholders and overtime. This connection between textual sentiment and stakeholder attribution offers insight into how Bank Indonesia's communication interacts with broader media narratives, shaping market perceptions and policy credibility. The next section provides a temporal and comparative assessment of these dynamics. The following section presents a detailed temporal and comparative analysis of these stakeholder-level dynamics.

4.2.1 Stakeholder Sentiment Trends Over Time

Figure 4. 8 Temporal Dynamics of Stakeholder Sentiment in Economic News (2019-2024)

The temporal analysis reveals notable month-to-month fluctuations in sentiment among analysts, government representatives, and regulators. While their trends generally move in parallel, clear divergences emerge at key moments. During early 2020 and mid-2022, when inflationary pressures intensified, sentiment among analysts dropped more sharply and recovered more slowly than sentiment linked to regulators or government officials.

Across the 2019–2024 period, the average sentiment range (–0.4 to 0.8) shows that media coverage rarely reaches extreme polarity but still varies meaningfully over time. These patterns indicate that Bank Indonesia's policy communication resonates unevenly across stakeholder groups. Even when official messages are clear and consistent, media narratives that incorporate diverse stakeholder viewpoints can reinforce or counterbalance them, producing alternating periods of sentiment alignment and divergence. These dynamic underscores the complex interplay between institutional communication and public interpretation.

4.2.2 Sentiment Distribution by Analyst and Institution

Figure 4. 9 Sentiment Distribution among Domestic Analysts and Research Institutions

The sentiment distribution among frequently cited domestic analysts and research institutions shows that Bank Indonesia's communication is generally received constructively. Positive and neutral tones dominate, especially among general analysts and research institutions, indicating broad confidence in the bank's policy direction.

Negative sentiment, while smaller in share, remains significant at roughly one-third of coverage. These critiques, often voiced by independent economists and firms such as Lembaga Riset Sekuritas&Bank and Bahana Sekuritas, introduce analytical caution by highlighting policy risks and market uncertainties, thereby enriching the public discourse.

Anonymized figures (Ekonom_1 to Ekonom_7) complement institutional commentary by providing individual expert perspectives. Together, these patterns underscore the media's role as an interpretive intermediary, where the mix of supportive, neutral, and critical views shapes perceptions of policy credibility and market expectations.

4.2.3 Media Platforms with the Most Frequent Critical Coverage

Analysis of the media distribution reveals that several online platforms serve as key channels for disseminating analytical commentary with a critical or cautious tone toward Bank Indonesia's policy communication. Among the most active sources are Liputan6.com, which recorded around 317 articles quoting analysts with negative sentiment, followed by Bisnis.com with 255, Bisnis Indonesia with 223, and Harian Kontan with 197. These outlets play a central role in shaping how alternative interpretations of monetary policy are circulated within the financial information ecosystem.

Figure 4. 10 Top Media Platforms with High Frequency of Critical or Negative Coverage

The prominence of diverse media platforms indicates that audiences relying on media coverage may encounter more varied and occasionally skeptical narratives than those following official releases directly, reflecting distinct communication pathways with differing tones and depth. While Bank Indonesia's official statements provide a consistent and constructive anchor for policy credibility, media framing and stakeholder quotations introduce interpretive diversity, particularly during periods of uncertainty. Analysts and research institutions play a key intermediary role in reframing policy signals, which are further amplified by media outlets. This layered information environment underscores the importance of clear, accessible, and transparent central bank communication to ensure public narratives remain aligned with intended policy messages.

4.3Communication Impact Analysis

Figure 4. 11 Framework Communication Impact Analysis

To assess the transmission of press-release features, we use monthly data from 2019–2024. We begin with preliminary exploration, calculating exchange-rate growth and conducting descriptive statistics, to map distributions and initial trends. We then estimate a Vector Autoregressive model with exogenous variables (VARX), preceded by standard assumption tests, to capture dynamic interactions across variables. Impulse Response Functions (IRF) are used to evaluate how shocks propagate over time, and robustness checks ensure the stability of the results. All findings are cross-validated with relevant literature to strengthen interpretation. The dataset includes the Monetary Central Bank Credibility Index, Monthly Number of Articles, Average News Sentiment, and key macroeconomic indicators (inflation and the exchange rate), covering January 2019 to December 2024. Visualizations and descriptive statistics are presented to illustrate the evolution and variability of these variables.

Figure 4. 12 Periodic Macroeconomic Data

Figure 4. 13 Periodic Number and Sentiment of News Articles

The chart shows how media attention and sentiment toward central bank press releases evolved from 2019–2024. The top panels reveal a rising trend in news coverage within 1, 3, 7, and 30 days after each release, with peaks in 2022–2023 when inflation and exchange-rate pressures intensified. This indicates that press releases attract

greater and more persistent attention during periods of macroeconomic uncertainty, with coverage often lasting up to a month. The lower panels show that sentiment around inflation and exchange-rate news fluctuates near neutral, occasionally dipping negative during periods of heightened economic stress, especially in 2022–2023. Although coverage volume rises over time, sentiment remains relatively stable and only weakly correlated with article counts (–0.25), suggesting that more coverage does not necessarily imply more positive or negative tone. In addition, the data highlights two dynamics: central bank communications increasingly sustain public discussion, and sentiment responds mainly to economic conditions rather than communication intensity. These patterns warrant further analysis, particularly in relation to inflation and exchange-rate movements.

Figure 4. 14 Monetary Central Bank Credibility Index Dynamics

The Monetary Central Bank Credibility Index from 2019–2024 shows broadly stable dynamics. The Overall index fluctuates within a narrow range and remains mean-reverting, indicating steady perceptions of policy effectiveness. The Coordination index, however, exhibits a clear regime shift during the COVID-19 period, rising to a higher and more stable level that reflects strengthened inter-agency coordination in crisis conditions. These identifiable shifts are useful for later robustness checks and interpretation of pandemic-related impulse responses. Before estimating the VAR model, we test all variables for stationarity to ensure valid dynamic analysis. Several series, news articles (1, 3, 7, and 30 days), inflation, and the Credibility Expectation index, were non-stationary and were differenced to achieve stationarity. After re-testing and confirming stationarity, the dataset is ready for VARX estimation and subsequent assumption testing.

Table 4.2 VARX result for Inflation Rate Model

The VARX results in Inflation Rate modelling point two clear channels. First, inflation tends to respond positively to the Number of Articles 30 Days after a press release and negatively to the Monetary Credibility Index (Overall). Interpreted through the expectations channel, more sustained coverage a month after announcement likely keeps the policy narrative salient and can tilt short-run inflation expectations upward, while higher perceived credibility anchors expectations and dampens price pressures. This pattern is consistent with evidence that credible and predictable central-bank communication stabilizes macro expectations and market reactions (Alan Blinder et al., 2008; Jansen et al., 2005). It also fits the literature showing that policy communication disseminated via news media moves expectations and near-term pricing (Kurav, 2021).

Furthermore, the results also uncover upstream links that help explain the dynamics of credibility and coverage. Thirty Days Number of Articles (coverage) itself rises with (higher) consistency in press release. In other words, the more consistent central bank communication (press release), the more able the news in gathering and talking about the information in the communication. Meanwhile, the dynamics of Monetary Credibility Index (Overall) significantly influenced by sentiment positively. Simply put, favorable tone is associated with stronger credibility which is known to reduce uncertainty and improve policy traction (Alan Blinder et al., 2008; Kurov & Stan, 2018)(Alan Blinder et al., 2008; Kurov & Stan, 2018). The net inflation outcome will reflect the balance of these forces, how intensively the message is covered and how credibly it lands.

Figure 4. 15 IRF Function for Inflation Modelling

The impulse response results (Figure 4.15) show that the Monetary Credibility Index and the Number of Articles 30 Days After Press Release significantly affect inflation, particularly at lag 1. A positive shock to credibility reduces inflation by about 0.05

standard deviations, indicating that stronger credibility helps anchor inflation expectations. In contrast, increased media coverage raises inflation in the short term, suggesting that heightened attention may amplify inflation expectations or market sensitivity. These results underscore the dual role of central bank credibility in stabilizing inflation and media exposure in shaping short-term inflation dynamics. This result aligns with evidence from Alan Blinder et al. (2008) and Ehrmann & Fratzscher (2005), who found that central bank credibility and communication significantly affect market reactions and inflation expectations. Similarly, recent studies such as Kurov & Stan (2018) demonstrate that monetary communication transmitted through news media can influence asset prices and inflation expectations in the short term.

To further improve the credibility of our model, we also do robustness tests to ensure the consistency of our findings. There are several tests we conducted, which are unit root, residual autocorrelation, and robustness test with several scenarios. These tests were arranged based on Kabundi and Mlachila (2018) and Nsor-Ambala and Anarfo (2022). The scenarios are a model with different lags (lagup), a model with different variable (proxy), and a model with different time period using COVID-19 period as a subsample for our models (pre). The results of robustness test are shown below.

Figure 4. 16 Stability Test in Inflation Model

The Lagrange-Multiplier test results for inflation suggest that there is no significant autocorrelation at lag 1, as the probability value (0.12324) is above the typical significance threshold (10%). This finding supports the assumptions of no autocorrelation in the residuals at lag 1, which is an important assumption for VARX modelling. Furthermore, the stability test for the roots of the companion matrix (as shown in Figure 4.16) indicates that all the roots lie inside the unit circle, which confirms that the inflation model is stable and well-specified. Stability in this context ensures that the system will not produce explosive or unrealistic responses to shocks, and that the VARX model for inflation is correctly estimated. The absence of instability or autocorrelation in this setup strengthens the validity of the results and supports proceeding with the estimation of the VARX model for inflation.

Figure 4. 17 Robustness Tests in Inflation Modelling

The impulse response results across all scenarios reveal a consistent pattern, with shocks to the Credibility Index influencing the Inflation Rate, which generally fades by the fifth lag. While there is a slight variation in the significance at lag 2, the overall response remains robust across different model specifications, indicating a stable relationship between credibility index shocks and inflation. The subsample analysis conducted for the period of the COVID-19 pandemic shows more divergence, where the credibility index's effect becomes more uncertain and volatile, suggesting that during periods of high economic uncertainty, the inflationary response to monetary credibility shocks is less predictable. This heightened uncertainty could be attributed to the unprecedented nature of the pandemic, where typical relationships between central bank communications and inflationary expectations might have been disrupted.

Table 4.3 VARX results for Exchange Rate Volatility Model

The VARX results for the Exchange Rate Volatility block indicate a clear stabilization channel driven by institutional coordination and information tone. Exchange-rate volatility decreases after positive shocks to the Monetary Credibility Index (Coordination) and after improvements in news sentiment. Stronger policy coordination enhances predictability and reduces uncertainty-driven hedging, while more positive media tone limits shock amplification through attention and narrative effects. These findings align with evidence that central bank communication and news sentiment shape currency-market volatility.

The upstream relationships clarify these mechanisms: economic appropriateness in press releases increases monetary credibility, and positive press-release sentiment leads to more favorable media tone. Together, economically appropriate communication raises credibility and lowers volatility, while more positive sentiment improves market tone and further stabilizes the exchange rate. These channels are consistent with empirical work showing that communication quality strengthens credibility, anchors expectations, and enhances policy transmission to financial markets.

Figure 4. 18 IRF of Exchange Rate Modelling

The impulse response results in Figure 4.18 reveal that both the Central Bank Monetary Credibility Index: Coordination and the Sentiment of the Press Release of Exchange Rate have significant influence on the Exchange Rate Volatility, particularly at lag 1. A negative response is observed following a positive shock in the Credibility Index, with the exchange rate declining more than 0.05 standard deviations, indicating that stronger central bank credibility effectively moderates exchange rate volatility. This suggests that stronger central bank credibility plays a crucial role in moderating exchange rate fluctuations, stabilizing expectations, and reducing market uncertainty. Similarly, a positive shock in the Number of Articles variable yields a decrease in exchange rate volatility beyond the same lag, suggesting that heightened sentiment about exchange rate following press releases may anchoring public expectations of exchange rate or signaling a reduced to monetary communication.

The findings in Figure 4.18 align with established literature on the role of central bank credibility and media sentiment in shaping exchange rate volatility. Central bank credibility has long been understood to stabilize financial markets by reducing uncertainty and anchoring expectations, as highlighted in studies such as Blinder et al. (2008), which demonstrate how clear central bank communication can lead to more stable market reactions. Similarly, the role of media sentiment in influencing market perceptions is well-documented in studies like Bholat et al. (2015), which suggest that news sentiment can affect the volatility of financial variables, including exchange rates. The decrease in volatility observed aftershocks to both the Credibility Index and Press Release Sentiment is consistent with these findings, reinforcing the importance of central bank communication and media coverage in mitigating market reactions and ensuring exchange rate stability.

To further improve the credibility of our model, we also do robustness tests to ensure the consistency of our findings. There are several tests we conducted, which are unit root, residual autocorrelation, and robustness tests with several scenarios. These tests were arranged based on Kabundi and Mlachila (2018) and Nsor-Ambala and Anarfo (2022). The scenarios are a model with different lags (lagup), a model with different variable (proxy), and a model with different time period using COVID-19 period as a subsample for our models (pre). The results of robustness test are shown below.

Figure 4. 19 Stability Test for Exchange Rate Model

Similarly, the Lagrange-Multiplier test for exchange rate volatility also shows no significant autocorrelation at lag 1, with the probability value (0.22671) indicating that the null hypothesis of no autocorrelation cannot be rejected. This suggests that the residuals from the VARX model for exchange rate volatility are well-behaved and that the model specification remains valid. Additionally, stability test for the roots of the companion matrix also indicates that the exchange rate volatility model is stable, with all roots inside the unit circle. This ensures that, like the inflation model, the exchange rate volatility model produces stable, non-explosive responses to shocks, and that the results of the VARX model for exchange rate volatility are reliable. The combination of no autocorrelation and stability in the system for exchange rate volatility reinforces the robustness of the model used for analysis.

Figure 4. 20 Robustness Test in Exchange Rate Volatility Modelling

The impulse response results in Figure 4.20 indicate that positive shocks to the Central Bank Monetary Credibility Index (Coordination) generally reduce exchange rate volatility, with effects fading by the fourth lag. In the base and proxy scenarios, volatility initially declines before rebounding, suggesting temporary and occasionally delayed effects of credibility shocks. The pre-COVID scenario shows more unstable responses, highlighting how crisis conditions can weaken the stabilizing role of credibility.

The analysis also examines which press release features drive media coverage. VARX results show that article volume responds differently across horizons: in the short term (1–3 days), coverage is mainly driven by press release sentiment, where more negative tones attract greater media attention, consistent with negativity bias. In the longer term (7–30 days), communication consistency becomes the main driver, sustaining media attention beyond the initial news cycle. Press release sentiment also positively influences average news sentiment, indicating that official tone anchors downstream reporting. Overall, the findings suggest a two-step mechanism in media transmission: sentiment shapes how coverage is framed, while consistency determines how long the issue remains salient.

Table 4.4 Linear Regression Result for Further NER Analysis

Linear regression results from the NER analysis show that press-release sentiment significantly improves perceptions among analysts, officials, regulators, and businesses ($p<0.05$), while having no measurable effect on public perception. This suggests that expert groups are more responsive to communication tone, whereas the broader public may lack the context to interpret nuanced policy signals. These findings highlight the need for communication strategies that better bridge the gap between specialist audiences and the general public to strengthen overall policy understanding and credibility.

The analyses further show that communication features transmit differently under economic stress. Clarity becomes especially influential during periods of inflation and exchange-rate anomalies, helping anchor expectations and reduce uncertainty, consistent with research emphasizing the heightened value of clear communication in turbulent times. While both Economic Appropriateness and Clarity appear linked to increased short-term volatility, Clarity ultimately reduces exchange-rate fluctuations and supports currency strength by aligning market expectations. This is consistent with evidence showing that transparent, context-appropriate communication can raise confidence even if it triggers temporary volatility.

Comprehensiveness, however, has adverse effects under anomaly conditions, lowering monetary credibility and dampening news sentiment. This supports the view that overly detailed communication can overwhelm audiences in times of uncertainty, making concise, targeted messaging more effective than exhaustive explanations. In addition, the VARX and IRF results reinforce the central role of communication channels, credibility, news volume, and sentiment in shaping inflation and exchange-rate dynamics. Higher credibility lowers inflationary pressures, while intense media attention can amplify expectations. For exchange-rate volatility, stronger credibility and more positive sentiment help stabilize market behavior. These findings underscore that effective communication, clear, contextually appropriate, and appropriately concise, is a key instrument for managing expectations and maintaining macroeconomic stability.

4.4Early Warning System

Figure 4. 21 The Architecture of the Early Warning System Development

4.4.1 Inflation

This section aims to analyze the relationship between public communication, societal behavior, and macroeconomic indicators in relation to actual inflation dynamics in Indonesia. Through empirical analysis, this study evaluates the extent to which indicators such as Google Trends "Inflasi", the Consumer Confidence Index (CCI), the Consumer Expectation Index (CEI), the number of news articles "Inflasi", the Monetary Credibility Index, and the Fuel Sales Index are correlated with and influence potential inflationary pressures, as reflected in the correlation values among variables. Furthermore, the ability of each indicator to provide early signals of inflationary changes is examined using probit regression to assess the significance of their relationships.

In addition, this section identifies variables that have the potential to serve as key indicators within the signaling approach to support the development of an inflation early warning system. The analysis is further reinforced through a robustness test to ensure the consistency and reliability of the proposed model. The overall results serve as the foundation for developing a more comprehensive framework for an inflation early warning system that takes into account behavioral, perceptual, and policy related dimensions.

4.4.1.1 Correlation Test

Correlation analysis was conducted to examine the extent to which two variables move in the same or opposite directions over time. In the context of this study, the correlation value is used to measure the strength and direction of the relationship between variables. The results of the correlation analysis provide an essential basis for understanding intervariable relationships and for assessing the potential role of each indicator as an early warning signal of inflation dynamics.

Table 4.5 Correlation Value Analysis (Inflation)

The analysis results indicate that several indicators have a strong correlation with inflation. The Number of News Articles "Inflasi" shows the highest correlation, 0.60 at lag 4. The Google Trends "Inflasi" indicator and the Current Economic Condition Index (IKE) both exhibit a moderate positive correlation of 0.48 at lag 2, while the Fuel Sales Index correlates 0.48 at lag 6 and the Consumer Confidence Index (IKK) 0.46 at lag 2. In contrast, the Monetary Credibility Index shows a negative correlation of -0,47 at lag 4. Overall, these indicators from behavior, perception, market activity, media, to credibility provide an effective framework for predicting and monitoring inflation trends.

Based on the correlation values presented in Table 4.5, Figure 4.22 illustrate the comparison between actual inflation data and the indicators with the highest correlation to inflation. These figures illustrate the relationship patterns and the dynamic movements of both variables over time.

Figure 4. 22 Comparison of Inflation Anomaly Periods

Isolation Forest results show that anomalies across communication and economic indicators consistently align with, often even precede, major shifts in Indonesia's inflation, highlighting their value as early warning signals. Spikes in Google Trends "Inflasi" appear concurrently with or slightly ahead of inflation increases, especially during early 2020 and mid–late 2022. Consumer survey indicators (IKK and IKE) also register anomalies in mid-2020 to 2021, well before the 2022 inflation surge, signaling that changes in household sentiment can anticipate rising price pressures. Likewise, anomalous increases in news coverage on "Inflasi" during mid–late 2022 reflect heightened public attention that typically accompanies turning points in inflation dynamics. Fuel Sales Index anomalies in early 2020, 2024, and late 2025 suggest that rising fuel consumption, an indicator of increased mobility and demand, can foreshadow demand-driven inflation. The Monetary Credibility Index shows concentrated anomalies

in 2022, coinciding with rapid inflation, indicating that weakening credibility can act as an early signal of policy-related inflation risks. Overall, anomalies in Google Trends, consumer sentiment indices, inflation-related news volume, fuel sales, and monetary credibility strongly track actual inflation movements. These multidimensional indicators, spanning behavioral, perceptual, and fundamental factors, prove effective for building an early warning system that is timely, comprehensive, and robust across economic conditions. Integrating them enables policymakers to detect emerging inflation pressures earlier and respond more precisely to maintain macroeconomic stability.

4.4.1.2 Probit Test

Following the correlation analysis that examined the relationships between the indicator variables and actual inflation to identify their interconnections, the next step is to evaluate how effectively these indicators can generate significant signals of inflationary changes. To achieve this, a probit regression approach is applied, allowing for a probabilistic assessment of the relationship between a binary dependent variable such as the occurrence or non-occurrence of inflation anomalies and several independent variables that previously exhibited strong correlations. This method helps estimate the probability of each indicator issuing an early warning signal of inflation anomalies, thereby providing empirical evidence on the reliability and effectiveness of these indicators as early warning instruments.

The results of the correlation analysis are summarized in the Table 4.6, which presents the variables with the strongest associations with actual inflation. The relatively high correlation values suggest that these indicators have a meaningful relationship with inflation dynamics, either by reflecting current economic conditions or by serving as early predictors capable of anticipating future inflationary trends.

Table 4.6 Probit Analysis (Inflation)

Probit results in Table 4.6 show that both communication-based and economic indicators significantly affect the probability of inflation anomalies. Google Trends "Inflasi" (lag 2) has a positive and significant effect, indicating that rising public search intensity serves as an early warning signal. The Consumer Confidence Index (IKK) shows a significant negative effect, implying that declining consumer optimism increases inflation risk, while the Consumer Expectation Index (IKE) has a positive effect, suggesting that excessive optimism can fuel demand-driven inflation. The Fuel Sales Index (lag 6) also positively influences inflation anomalies, reflecting increased economic activity, while inflation-related news coverage (lag 4) provides an additional informational signal. In contrast, the Monetary Credibility Index (lag 4) has a strong negative effect, indicating that credible and tight monetary policy reduces inflation risk. In addition, the results underscore the importance of combining behavioral, communication, and fundamental indicators to build an effective inflation Early Warning System (EWS).

4.4.1.3 Signaling Approach

Based on the results of the probit regression estimation, several variables were identified as statistically significant indicators for the early warning system. The next step is to analyze the effectiveness of these independent variables as indicators within the early warning framework. In this stage, the signaling approach is applied to evaluate the performance of each indicator in providing early signals of inflation changes. Through this approach, each indicator is assessed based on its ability to accurately identify periods of inflation anomalies (true signals) while minimizing false signals, thereby determining which indicators are the most reliable and consistent in detecting potential inflationary pressures in the future. By assessing the ability of each indicator to distinguish between inflationary pressures and normal conditions, this analysis confirms their effectiveness in supporting an early warning system for inflation dynamics. The following presents the performance results of each indicator within the signaling approach as an early warning tool for detecting inflation anomalies.

Figure 4. 23 Confusion Matrix of the Signaling Approach Results

The analysis shows substantial variation in indicator effectiveness for detecting inflation anomalies. The Monetary Credibility Index and the Number of News Articles on "Inflasi" emerge as the most effective indicators. Declines in monetary credibility often precede rising inflation, making it a key signal of policy-related inflation risk, while surges in inflation-related media coverage closely track or even precede anomaly periods, yielding high detection accuracy with few missed signals. Google Trends "Inflasi" and the Consumer Confidence Index (IKK) display moderate performance. Google Trends captures rising public concern ahead of inflation episodes but is influenced by external information shocks, while IKK reflects household expectations that do not always align with actual inflation, leading to false negatives. The Fuel Sales Index also performs moderately, as its response to inflation is lagged and affected by non-price factors. In contrast, the Current Economic Condition Index (IKE) shows the weakest performance, with delayed and inaccurate signals driven by divergence between perceptions and actual price movements. Therefore, Indicators based on policy credibility and public information are more reliable for timely inflation detection, while consumer perception indicators serve as complementary inputs. Among all indicators, inflation-related news coverage stands out as the most effective early warning signal, supported by its high predictive power and very low Noise-to-Signal Ratio, indicating strong accuracy in capturing inflationary pressures.

Table 4.7 Predicted Value and NSR (Inflation)

Conversely, the Monetary Credibility Index and the Number of News Articles "Inflasi" emerge as the most effective indicators showing strong predictive ability and relatively low NSR values. The Google Trends "Inflasi" variable and the Consumer Confidence Index (IKK) show moderate performance, with relatively higher NSR values. The Fuel Sales Index demonstrates more limited predictive capability. Meanwhile, the Current Economic Condition Index (IKE) exhibits the weakest performance, characterized by low predictive accuracy and high NSR values, making it less suitable as a standalone indicator in an inflation early warning system.

4.4.1.4 Robustness Test

After evaluating the effectiveness of the indicators in providing early signals of inflationary pressures through the signaling approach, the next step is to assess the robustness or reliability of the results. The robustness test aims to ensure that the indicators' ability to detect inflation anomalies is not coincidental or limited to specific periods but remains consistent across different conditions. To achieve this, several evaluation methods are employed, including the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC), which assess the accuracy and consistency of each indicator in distinguishing between normal and anomalous inflation periods. Additionally, usefulness and loss score measures are applied to evaluate the economic value and cost of misclassification associated with each indicator. Through this comprehensive robustness analysis, the study ensures that the indicators incorporated into the inflation early warning system provide accurate, consistent, and reliable signals under varying economic circumstances.

Figure 4. 24 ROC Curve and AUC Value

Google Trends "Inflasi" shows the strongest performance, with an AUC of 0.71, indicating high accuracy and timeliness in detecting inflation anomalies. The Fuel Sales Index (0.67) and Current Economic Condition Index (IKE, 0.68) perform moderately, capturing early inflation signals but with higher false detections due to slower responses rooted in consumption patterns and perceptions. The Consumer

Confidence Index (IKK) follows closely (0.66), while the Number of News Articles on "Inflasi" (0.58) and the Macro Monetary Credibility Index (0.38) exhibit weak detection capability. Overall, Google Trends is the most responsive indicator, followed by Fuel Sales Index and IKE. Robustness is further evaluated using usefulness and loss scores (Table 4.8), where lower loss values indicate more stable performance, providing a comprehensive assessment of each indicator's suitability for an inflation early warning system.

Table 4.8 Usefulness and Loss Score (Inflation)

The usefulness score captures the informativeness of each indicator, while the loss score measures policymaker costs arising from false alarms (Type I errors) and missed signals (Type II errors). The results indicate that the Macro-monetary Index and the Number of News Articles perform best, providing the most reliable signals in detecting inflation anomalies. Google Trends "Inflation" and the Consumer Confidence Index (IKK) show moderate but still meaningful performance despite some detection errors. The Fuel Sales Index exhibits weaker predictive power due to its lagged response to inflation dynamics, while the Current Economic Condition Index (IKE) ranks lowest, with poor accuracy and high NSR values, limiting its effectiveness as a standalone early warning indicator.

Table 4.9 Summary of Robustness Tests (Inflation)

Robustness test results in Table 4.9 show varying indicator performance in detecting inflation anomalies. Google Trends "Inflasi" achieves the highest AUC (0.71), followed by IKE (0.68), Fuel Sales Index (0.67), and IKK (0.66). In terms of usefulness, News Articles score highest (0.67), followed by the Monetary Credibility Index (0.63), indicating that information-based indicators provide valuable signals despite lower AUC values. Overall, these results confirm the effectiveness of the signaling approach, with information-based indicators being more responsive than survey or monetary measures.The findings also reveal a complementary role among indicators: public attention and news-based indicators provide early signals, consumer perception indicators offer supportive insights, and fundamental indicators supply economic context. Integrating these dimensions into a single early warning system improves the accuracy, consistency, and timeliness of inflation detection.

4.4.2 Exchange Rate

4.4.2.1 Correlation Analysis

Correlation analysis is used to identify relationships among indicators, including internet search intensity, central bank survey results, and news sentiment related to exchange rate depreciation or inflation. By examining how these variables interact with crisis episodes, the central bank can employ them as early warning signals and enhance policy analysis through machine learning. More broadly, this approach supports the development of an early warning system by identifying leading indicators of crisis risk. Heatmap correlation further aids policy analysis by visualizing the strength of relationships between indicators and crises. This study evaluates correlations at lags one to six to identify the strongest association with exchange rate anomalies. Indicator selection for the early warning system is therefore based on correlation analysis, as probit analysis was found to be less effective, yielding weak statistical significance and limited explanatory power.

Table 4.10 Correlation Analysis Result (Exchange Rate)

The analysis identifies several indicators with high correlations (coefficients above 0.5), including Google Trends–based variables (Trend Rupiah Lag 6, JISDOR Lag 1, and Spot Exchange Lag 6) and consumer survey indicators (Consumer Confidence Index Lag

4, Current Economic Condition Index Lag 6, Durable Goods Purchasing Index Lag 1, and Employment Availability Expectation Index Lag 4). Additionally, negative sentiment from press releases emerges as a relevant signal, as its movement aligns with exchange rate depreciation (Figure 4.25), indicating that adverse news sentiment may serve as an early indicator of potential crisis conditions.

Figure 4. 25 Indicators of EWS and Exchange Rate Curve

After identifying the seven factors with the highest correlations, a composite analysis was subsequently conducted to examine the potential for a stronger relationship.

4.4.2.2Probit Test

a. Individual Variable

The probit analysis was conducted to model the probability of a crisis occurrence (crisis = 1; no crisis = 0). This estimation aimed to filter indicators capable of accurately capturing crisis signals. Overall, the probit results for each indicator revealed relatively weak correlation analysis (>0.5) and small Pseudo R^2 values (<0.1), indicating limited predictive performance of the model. Therefore, the author decided to conduct a probit analysis on the exchange rate by constructing a composite variable. The variables selected for inclusion in the composite analysis were those that are theoretically and empirically relevant to the phenomenon or concept being represented.

b. Composite Variable

After identifying the indicators that exhibit a relatively strong correlation with exchange rate movements, a strategy was developed to optimize the extraction and combination of signals from various cyclical indicators. A previous study by Beltran et al. (2024) show that combining multiple indicators into a composite improves performance relative to optimizing individual indicators. Following this approach, this study constructs two composite indicators based on data sources. Consumer survey–based indicators, namely the Current Economic Condition Index (lag 6), Consumer Confidence Index (lag 4), and Durable Goods Purchasing Index (lag 1), are combined into the Socio-Economic Index Indicator. Meanwhile, Google Trends–based indicators, including g Spot Exchange (lag 6) and JISDOR (lag 1), are combined into the Foreign Exchange Market Indicator.

Probit results show that the Composite Socio-Economic Index has a positive and statistically significant effect on anomaly probability ($p|z| = 0.020$), indicating that improvements in socio-economic conditions are associated with a higher likelihood of anomalies, possibly reflecting perception–reality mismatches. The Foreign Exchange Market Indicator is marginally significant ($p|z| = 0.092$), suggesting that exchange rate pressures may also contribute to anomaly emergence. The pseudo-R^2 of 0.2759 indicates that these composite indicators explain a substantial share of variation in anomaly occurrences.

Table 4.11 Probit Regression Result with Composite Variables

Figure 4. 26 Weighted Probit Result

Composite variables are constructed to strengthen analytical signals, consisting of the Socio-Economic Index (S = IKE + IKK + IPBT) and the Foreign Exchange Market Indicator (V = Spot Exchange + JISDOR). The probability of exchange rate anomalies is estimated using a probit model, with parameters obtained via maximum likelihood. The predicted probabilities are smoothed using a moving average and normalized by z-scores

to enable comparison with exchange rate movements, while anomalies are identified using thresholds based on the exchange rate's moving average and standard deviation. The weighted probit results (Figure 4.26) show a pattern that aligns with currency depreciation trends, indicating the model's ability to detect potential crisis signals. The simultaneous rise in probit probabilities and exchange rate pressures highlights the interaction between deteriorating socio-economic conditions and foreign exchange market dynamics, reinforcing the probit model's validity as a predictive indicator and its relevance for inclusion in an Early Warning System (EWS).

4.4.2.3 Signaling Approach

Individual anomaly points are first identified for each indicator to generate signals, which are then evaluated using the signaling approach to assess early warning system accuracy. Following Kaminsky et al. (1998), this study maps signals using a threshold defined as the rolling mean plus one standard deviation. Indicators with correlation values above 0.5 are further analyzed to evaluate their ability to predict future crises. Signal distributions are compared with exchange rate crisis events (Figure 2), and a signal is deemed accurate if it occurs within a six-month window prior to a crisis.

Figure 4. 27 Signal Analysis from each indicator

The alignment between Early Warning System (EWS) signals and actual outcomes is summarized in a contingency matrix, where signal performance depends on the chosen threshold. Lower thresholds reduce missed crises but increase false alarms. Following Alnami et al. (2025), this study determines thresholds using a rolling Z-score based on a 30-day moving average and standard deviation. This approach adapts to local data dynamics, making it well-suited for time-series analysis and consistent with prior machine learning studies that show its effectiveness in anomaly detection.

Figure 4. 28 Confusion Matrix of Exchange Rate

Based on the confusion matrix, performance metrics are computed, including accuracy, precision, sensitivity, specificity, false negative and false positive rates, and the noise-to-signal ratio. The noise-to-signal ratio measures incorrect signals outside crisis periods relative to correct signals within crisis windows. These metrics are then used to further assess the alignment between predicted signals and actual crisis events.

Table 4.12 Signal Result Based on the Signaling Approach Analysis

In the signaling framework, sensitivity (true positive rate) measures the proportion of correctly predicted crises relative to total crisis events. The Noise-to-Signal Ratio (NSR) serves as a key metric to evaluate indicator effectiveness, capturing the balance between false alarms and correct crisis signals. Lower NSR values indicate stronger discriminatory power and higher reliability of the Early Warning System (EWS), while higher values reflect excessive noise. The results show that JISDOR and Negative Press Release Sentiment are the most effective indicators, with NSR values of 0.3485 and 0.0218, respectively, indicating strong early detection with minimal errors. This finding is consistent with Kaminsky et al. (1998), as both indicators generate timely signals while maintaining low noise. JISDOR reflects external pressures on the financial system, while negative media sentiment captures rising market expectations of instability. Trend Rupiah and the Employment Availability Expectation Index also perform moderately well (NSR < 1), signaling foreign exchange dynamics and deteriorating labor market expectations ahead of crises.

In contrast, Spot Exchange, IKE, and IKK exhibit weak performance, with NSR values above 1, indicating more false signals than accurate predictions, likely due to short-term volatility or seasonal effects. The Durable Goods Purchasing Index (IPBT) occupies a borderline position, with an NSR slightly above 1 but still capturing shifts in consumption behavior prior to crises. In addition, the signaling analysis underscores the

importance of a multi-indicator approach. Low-NSR indicators function effectively as primary signals, while moderately performing indicators serve as complementary signals, jointly enhancing detection accuracy and managing the trade-off between false alarms and missed crises within the EWS framework.

4.4.2.4 Robustness Test

a. ROC and AUC

ROC and AUC analyses are used as goodness-of-fit measures to assess model quality and performance, complementing the signaling approach by evaluating correctly predicted crises, missed crises, and false alarms. This assessment is crucial given its implications for policy decisions. Consistent with prior studies Alessi and Detken (2011), Betz et al. (2014), Bussiere & Fratzscher (2006), Davis and Karim (2008), the framework recognizes that missed crises and false alarms carry different policy costs. The ROC–AUC analysis captures this trade-off, justifying the need to evaluate both dimensions separately.

Figure 4. 29 The ROC and AUC Results of Each Indicator in EWS

ROC–AUC results show clear differences in predictive performance across indicators. Negative press release sentiment (AUC = 0.76), JISDOR (0.74), and Trend Rupiah (0.73) perform best, indicating strong early detection of exchange rate anomalies with high sensitivity and low false positives. Confidence and expectation indicators, IKE (0.68), IKK (0.60), and IEKL (0.62), exhibit moderate predictive power, consistent with their lagging nature. The Durable Goods Purchasing Index (IPBT, 0.59) shows weaker performance, while the Spot Exchange indicator (0.53) performs near random. In addition, communication-based and exchange rate–related indicators are most effective for early warning, while survey-based indicators contribute more to monitoring medium-term resilience. Integrating these indicators within a composite Early Warning System enhances robustness, improves crisis detection, and strengthens policy responsiveness to emerging financial risks.

b. Usefulness and Loss Function

In implementing an Early Warning System (EWS), policymakers face a trade-off between missing crises and issuing false alarms. Missing emerging vulnerabilities typically entails higher costs than unnecessary preventive actions. To reflect this asymmetry, policymakers can define a loss function that assigns different weights to each type of error.

Table 4.13 Comparison of the usefulness and loss scores per Indicators

The analysis shows that the indicator with the highest usefulness provides the most accurate crisis signals within a six-month prediction window, conditional on threshold selection (Figure 5). Negative Sentiment Press Release (NSPR) performs best, with the highest usefulness score (0.9010) and the lowest loss value (1.0990), indicating strong efficiency in capturing early crisis signals associated with rupiah depreciation. JISDOR (0.7899), Trend Rupiah (0.4364), and IPBT (0.2440) also demonstrate relatively strong predictive performance, highlighting the relevance of keyword-based indicators in detecting external financial pressures. In contrast, Spot Exchange and IKE show very low usefulness and high loss values, suggesting limited standalone effectiveness and greater suitability within composite indicators. Overall, the inverse relationship between usefulness and loss confirms a trade-off in which more informative indicators yield lower prediction errors and fewer false signals. This finding is consistent with previous literature, such as Sarlin (2013) and Alessi & Detken (2011) which emphasizes the importance of using loss-based evaluation to capture policymakers' preferences towards two types of prediction errors.

4.4.3 Application of digital visualization in early warning systems

Building on the literature on macroeconomic early warning systems, this study employs Power BI with machine learning models to identify key signal indicators of financial stress. The approach also determines optimal thresholds to define crisis onset, aiming to maximize correct detections while minimizing false alarms.

Figure 4. 30 Digital visualization of an early warning system.

The Early Warning System Rupiah Anomaly Signaling Dashboard provides a data-driven framework for monitoring exchange-rate stability and identifying emerging market pressures. Built on the signaling approach of Kaminsky et al. (1998), the system flags early warnings when indicators deviate from statistically defined normal ranges. Its long-table structure, comprising Date, Indicator, and Value, supports flexible normalization and consistent DAX-based processing in Power BI. Rolling means (M) and rolling standard deviations (σ) are calculated for each indicator, allowing anomalies to be detected based on relative deviations rather than absolute levels. This method adjusts to changing volatility conditions and aligns with the literature on composite early warning systems, which demonstrates that multi-indicator models produce more reliable signals and fewer false alarms than single-variable approaches (Kaminsky et al., 1998). Interactive visualizations in Power BI enhance transparency and interpretability, integrating quantitative assessment with intuitive displays. As a result, the dashboard serves as a practical decision-support tool for monetary authorities, enabling timely, evidence-based monitoring of macroeconomic stability and systemic risks.

4.5 Central Bank Communication Response During Anomaly Conditions

Central bank communication is a key instrument for achieving monetary policy objectives by shaping stakeholder expectations. As emphasized by Blinder et al. (2008), effective communication can influence financial markets, enhance policy predictability, and strengthen monetary transmission by providing clear information on policy goals, strategies, and economic outlook. Its effectiveness depends on transparency and accountability, particularly during periods of uncertainty, as reinforced by post-crisis reforms and Bank Indonesia's legal mandates (Law No. 3/2004; Law No. 4/2023). During economic anomalies such as rising inflation or exchange rate depreciation, communication also serves as a tool for expectation management and psychological stabilization. Clear and credible messaging helps anchor inflation expectations, reduce misinterpretation, and mitigate excessive market reactions, thereby reinforcing central bank credibility and policy effectiveness.

To guide communication during anomaly conditions, this study applies Situational Crisis Communication Theory (SCCT) by Coombs (2007), which links communication effectiveness to the level of responsibility attributed to an institution. Central bank crisis communication comprises two stages: basic response and reputational repair. The basic response involves providing instructive and contextual information to reduce uncertainty and public anxiety. Reputational strategies then depend on perceived responsibility, victim, accidental, or preventable, ranging from denial and bolstering, to diminishment with corrective actions, or rebuilding strategies involving apologies and recovery measures to restore credibility.

After defining its crisis communication posture, the central bank must choose effective channels to ensure rapid and accurate information delivery. Communication channels are broadly classified into Above the Line (ATL) and Below the Line (BTL). ATL channels, such as press conferences, policy announcements, and public statements, emphasize broad awareness and reinforce perceptions of stability. In contrast, BTL channels provide more targeted and contextualized communication through stakeholder forums, industry dialogues, academic engagement, and financial literacy programs. During anomaly conditions, the combination of ATL and BTL channels ensures both widespread reassurance and actionable understanding among key economic actors.

Through-the-Line (TTL) communication integrates mass and targeted channels to ensure message consistency across platforms (Pelsmacker et al., 2018). For central banks, TTL combines broad policy announcements with follow-up materials such as digital visuals, analyst briefings, and public education, preserving message clarity during volatile conditions. Effective crisis communication must be stabilizing, expectation-guiding, and coherently integrated across ATL, BTL, and TTL channels, enabling Bank Indonesia to reduce uncertainty, maintain credibility, and reinforce economic stability during periods of heightened volatility.

5. Conclusion

This study offers a comprehensive evaluation of Bank Indonesia's communication by integrating linguistic analysis, media sentiment assessment, stakeholder response mapping, communication impact modeling, anomaly diagnostics, and an Early Warning System. The findings confirm that central bank communication functions as a policy instrument that shapes expectations, influences market reactions, and supports macroeconomic stability. Empirical results show that linguistic features, such as clarity, tone, consistency, and economic relevance, interact with media narratives and stakeholder interpretations, affecting inflation perceptions, exchange rate movements, and overall economic sentiment. Tone and clarity emerge as key drivers of favorable media coverage and reduced uncertainty, while ambiguous or overly complex messages amplify negative sentiment. Consistent communication anchors expectations in normal conditions, whereas concise and targeted messaging proves more effective during economic stress. The study also highlights the media's amplifying role and heterogeneous stakeholder responses, particularly among financial market participants, underscoring the need for proactive media engagement and segmented communication strategies.

The Early Warning System (EWS) developed in this study adds a critical diagnostic layer to communication planning by detecting deviations in linguistic and macroeconomic signals, enabling more timely and preemptive communication during high-risk periods. Its integration with digital visualization tools enhances interpretability and supports both internal decision-making and transparency. The findings highlight that effective central bank communication requires not only timely and accurate information but also strategic message design, channel integration, and audience targeting. Bank Indonesia operates within a dynamic ecosystem where language, media, stakeholder perceptions, and economic signals interact, necessitating a holistic approach that combines linguistic optimization, rapid-response mechanisms, digital tools, and continuous empirical monitoring. Policy recommendations are structured across time horizons: short-term actions focus on improving clarity, tone, frequency, and media alignment; medium-term reforms institutionalize communication templates, audience segmentation, crisis protocols, and monitoring; and long-term priorities embed the EWS into governance, modernize platforms, strengthen public literacy, and establish a dedicated communication research function. Together, these measures enhance the credibility and effectiveness of Bank Indonesia's communication and support its mandate for monetary and financial stability. This study contributes to the literature by demonstrating how linguistic analytics and digital diagnostics complement macroeconomic analysis. Future research may extend this framework through multimodal communication analysis, social network-based perception measures, and experimental testing, reinforcing the role of data-driven communication in effective monetary policy and public trust.

6. References

Aldasoro, I., Borio, C., & Drehmann, M. (2018). Early Warning Indicators of Banking Crises: Expanding the Family. *Journal of Financial Transformation*, *48*(March).

Alessi, L., & Detken, C. (2011). Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy*, *27*(3). https://doi.org/10.1016/j.ejpoleco.2011.01.003

Alnami, H., Mohzary, M., Assiri, B., & Zangoti, H. (2025). An Integrated Framework for Cryptocurrency Price Forecasting and Anomaly Detection Using Machine Learning. *Applied Sciences (Switzerland)*, *15*(4), 1–13. https://doi.org/10.3390/app15041864

Azis, I. J. (2024). Listen and Design: On Micro, Small and Medium Enterprises. In *Listen and Design: On Micro, Small and Medium Enterprises*. https://doi.org/10.1007/978-981-97-3248-7

Bank for International Settlements. (2022). *Central bank communications: Modern challenges and digital transformation* (BIS Papers No. 125).

Barbariol, T., & Susto, G. A. (2022). TiWS-iForest: Isolation forest in weakly supervised and tiny ML scenarios. *Information Sciences*, *610*. https://doi.org/10.1016/j.ins.2022.07.129

Beltran, D. O., Dalal, V. M., Jahan-parvar, M. R., & Paine, F. A. (2024). Optimizing composite early warning indicators ☆. *North American Journal of Economics and Finance*, *74*(August), 102250. https://doi.org/10.1016/j.najef.2024.102250

Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking and Finance*, *45*(1). https://doi.org/10.1016/j.jbankfin.2013.11.041

Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). *Text mining for central bank communication* (Bank of England Working Paper No. 529).

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.-J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4), 910–945. https://doi.org/10.1257/jel.46.4.910

Bulíř, A., Čihák, M., & Jansen, D. (2013). *Does inflation targeting matter? Measuring the impact of monetary policy communication* (IMF Working Paper No. 13/208).

Bussiere, M., & Fratzscher, M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, *25*(6). https://doi.org/10.1016/j.jimonfin.2006.07.007

Carcillo, F., Le Borgne, Y. A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, *557*. https://doi.org/10.1016/j.ins.2019.05.042

Charlet, D., Damnati, G., & Paroubek, P. (2017). SimBow at SemEval-2017 Task 3: Soft-Cosine Semantic Similarity between Questions for Community Question Answering. In *Proceedings of SemEval-2017*, 14–80.

Chow, C. K. (1970). On Optimum Recognition Error and Reject Tradeoff. *IEEE Transactions on Information Theory*, *16*(1). https://doi.org/10.1109/TIT.1970.1054406

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., … Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. *Proceedings of ACL*, 8440–8451.

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics*: A step-by-step approach (2nd ed.). Wiley.

Davis, E. P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, *4*(2). https://doi.org/10.1016/j.jfs.2007.12.004
Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT*, 4171–4186.
Dincer, N. N., Eichengreen, B., & Geraats, P. M. (2019). Transparency of monetary policy in the post-crisis world. *Economic Modelling*, 82, 1–18. https://doi.org/10.1016/j.econmod.2019.08.004
Dudáš, A. (2024). Graphical representation of data prediction potential: correlation graphs and correlation chains. *Visual Computer*, *40*(10). https://doi.org/10.1007/s00371-023-03240-y
Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382. https://doi.org/10.1037/h0031619
Gorodnichenko, Y., Pham, T., & Talavera, O. (2023). The voice of monetary policy. *American Economic Review*, 113(2), 548–584. https://doi.org/10.1257/aer.20220129
Hansen, L. K., Liisberg, C., & Salamon, P. (1997). The error-reject tradeoff. *Open Systems and Information Dynamics*, *4*(2). https://doi.org/10.1023/A:1009643503022
Hansen, S., & McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99, 114–133. https://doi.org/10.1016/j.jinteco.2015.12.008
Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics*, 133(2), 801–870. https://doi.org/10.1093/qje/qjx045
Jablonski, A., & Mendrok, K. (2025). Automatic threshold setting for anomaly detection. *Mechanical Systems and Signal Processing*, *228*(April 2024), 112462. https://doi.org/10.1016/j.ymssp.2025.112462
Kaminsky, G. L., & Reinhart, C. M. (1999). The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, *89*(3). https://doi.org/10.1257/aer.89.3.473
Kaminsky, G., Lizondo, S., & Reinhart, C. M. (1998). Leading Indicators of Currency Crises. *IMF Staff Papers*, *45*(1). https://doi.org/10.2307/3867328
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
Lee, V., Seng, W., Bakar, S. A., & Abidinsa, S. (2024). *Machine learning for anomaly detection in money services business outlets using data by geolocation*. *23*.
Li, Z., Yang, Z., & Sun, M. (2020). Domain-specific named entity recognition via fine-tuned language models: Applications to economic text. *Computational Linguistics Research*, 56(3), 201–218.
Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference*. IEEE.
Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x
Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
Marhastari, R., 2024. *Topics on Non-Expert Economic Agents' Inflation Expectations*, University of York.
Nababan, P. W. J. (1991). *Kedwibahasaan dan pengajaran Bahasa Indonesia*. Gramedia.
Russo, A., Moneta, F., & Gandolfo, A. (2021). Hybrid sentiment analysis models in multilingual economic communication. *Journal of Economic Behavior &*

Organization, 188, 945–960. https://doi.org/10.1016/j.jebo.2021.06.012
Sarlin, P. (2013). On policymakers' loss functions and the evaluation of early warning systems. *Economics Letters*, *119*(1). https://doi.org/10.1016/j.econlet.2012.12.030
Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2010). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *40*(1). https://doi.org/10.1109/TSMCA.2009.2029559
Shapiro, A. H., & Wilson, D. J. (2022). Taking the Fed at its word: A new approach to estimating central bank communication effects. *Journal of Monetary Economics*, 126, 1–15.
Sidorov, G., Gelbukh, A., Gómez-Adorno, H., & Pinto, D. (2014). Soft similarity and soft cosine measure: Similarity of features in vector space model. *Computación y Sistemas*, 18(3), 491–504. https://doi.org/10.13053/cys-18-3-2043
Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
Thorsrud, L. A. (2020). Words are the new numbers: A news-driven indicator of economic activity. *Journal of Business & Economic Statistics*, 38(2), 393–409. https://doi.org/10.1080/07350015.2018.1444989
Vlaar, P. J. G. (2002). *Early warning systems for currency crises : A Markov-switching approach*. 253–

Appendix

Appendix 1: Details of Research Methodology

Hybrid Sentiment Analysis

The sentiment analysis of news articles employs a hybrid ensemble model combining rule-based lexicon methods and machine learning–based contextual models. At the rule-based level, sentiment classification relies on an economic sentiment lexicon adapted from Loughran and McDonald (2011), which was further contextualized for Indonesian financial and policy terminology. Each term in the lexicon is associated with a polarity score $w_i \in [-1,1]$, indicating negative, neutral, or positive sentiment strength. The lexicon-based sentiment score for each article is calculated as:

$$S_{lex} = \frac{\sum_{i=1}^{n} w_i f_i}{N}$$

where f_i is the frequency of occurrence of each sentiment-bearing word, and N is the total count of identified sentiment words in the article.

At the machine learning level, sentiment is analyzed using a contextual transformer model trained on the Indonesian Sentiment Analysis corpus (Indo-SMSA). The model predicts sentiment polarity based on contextual embeddings rather than individual words. The output probability distribution $P = [p_{pos}, p_{neu}, p_{neg}]$ is converted into a continuous sentiment score:

$$S_{ml} = p_{pos} \times 1 + p_{neu} \times 0.5 + p_{neg} \times 0$$

The final article sentiment score combines both approaches using a weighted consensus inspired by Russo et al. (2021), which emphasizes the contextual accuracy of machine learning models while retaining lexicon interpretability:

$$S_{final} = 0.3 \times S_{lex} + 0.7 \times S_{ml}$$

The 70:30 weighting is supported by comparative evaluations in Bholat et al. (2015) and Ardia et al. (2019), which demonstrate that hybrid ensemble approaches outperform single-method sentiment models in policy communication analysis. This combination also reflects the methodological philosophy of balancing interpretability (rule-based) and contextual sensitivity (machine learning).

Stakeholder Identification Using Named Entity Recognition (NER)

To associate sentiment with its corresponding source, a custom Named Entity Recognition (NER) model was applied to the news corpus to identify stakeholder entities explicitly mentioned in the text. The model utilizes a semi-supervised annotation approach that combines expert-curated dictionaries and statistical learning. A domain-specific stakeholder lexicon, referred to as the "Expert Version Dictionary," was developed to guide entity tagging during the training process. This dictionary reflects the institutional ecosystem surrounding Bank Indonesia's communication and ensures precise classification of stakeholders into their respective policy relevance categories. Each entity recognized by the NER model is categorized into one of six stakeholder groups representing the main actors in Indonesia's monetary policy communication landscape. The categories and representative examples are presented in Table 3.1, which also includes direct references to the entity lists used within the expert dictionary that guided model annotation and training.

The methodological foundation of this approach draws upon established literature in domain-adaptive entity recognition and central bank communication studies, including Li et al. (2020) and Hansen and McMahon (2016). These studies emphasize the importance of aligning linguistic models with economic domain knowledge to ensure both precision and contextual relevance. By explicitly linking each stakeholder category with representative entities, the taxonomy provides analytical transparency and reproducibility, which are essential for the consistent interpretation of stakeholder sentiment in subsequent stages of the analysis.

Sentiment-Stakeholder Integration

Following stakeholder identification, the final step integrates the sentiment polarity of articles with the entities identified by NER to construct a stakeholder–sentiment mapping matrix. This integration links each stakeholder category with the corresponding sentiment expressed in the text, enabling a nuanced view of how different actors perceive or respond to central bank communication. For each stakeholder group g, the average sentiment score is computed as:

$$S_g = \frac{1}{n_g} \sum_{i=1}^{n_g} S_{final,i}$$

where n_g is the total number of articles mentioning stakeholder group g, and $S_{final,i}$ is the final sentiment score of the i^{th} article containing such mention.

Robustness Testing Methodology

This study applies a series of robustness tests to ensure that each communication feature is statistically reliable, conceptually valid, and free from structural bias. The tests were designed to assess the stability, internal consistency, and empirical reliability of all indicators that measure clarity, comprehensiveness, economic appropriateness, sentiment, and consistency. Four complementary approaches were used, namely stability testing, proxy validation, inter-model comparison, and inter-annotator reliability assessment.

a. Stability Testing

The stability test examines whether the feature scores are influenced by document structure, particularly by text length. The analysis employs Spearman's Rank Correlation Coefficient (ρ), which measures the strength and direction of a monotonic relationship between two ranked variables (Spearman, 1904; Corder & Foreman, 2014). The coefficient is expressed as follows:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

where $d_i$is the difference between the ranks of paired observations, and $n$is the total number of observations. A correlation close to zero indicates that the score is not structurally dependent on document length, confirming that the feature measures linguistic rather than mechanical variation. The use of this nonparametric correlation method follows the approach of Hansen and McMahon (2016), who emphasize the importance of isolating linguistic signals from structural noise in the study of central bank communications.

b. Proxy Validation

The proxy validation test evaluates whether each composite indicator accurately reflects its conceptual foundation. Each aggregated feature is compared with its principal subcomponent using Spearman's correlation. For example, overall clarity is tested against the Flesch readability score, while comprehensiveness is tested against topic coverage. Strong or moderate correlations confirm that the composite feature faithfully represents its underlying linguistic construct. This approach follows the recommendations of Bholat et al. (2015) and Gorodnichenko, Pham, and Talavera (2023), who emphasize the need for internal coherence in text-based indicators used in monetary communication analysis.

c. Inter-Model Comparison for Consistency

The analysis also evaluates the internal consistency of textual coherence using two metrics: lexical similarity and semantic similarity. Lexical similarity is computed through

the Term Frequency–Inverse Document Frequency (TF–IDF) vector model combined with cosine similarity, as described by Manning, Raghavan, and Schütze (2008):

$$S_{lex}(A,B) = \frac{A \cdot B}{\| A \| \| B \|}$$

where A and B represent TF–IDF-weighted document vectors, $A \cdot B$is their dot product, and $\| A \|$and $\| B \|$ are the magnitudes of each vector.

Semantic similarity, on the other hand, is calculated using the Soft-Cosine Similarity formulation introduced by Sidorov et al. (2014) and adapted for word-embedding representations as demonstrated in Charlet et al. (2017). Where the formula is:

$$S_{sem}(A,B) = \frac{\sum_{i,j} s_{ij}\, a_i b_j}{\sqrt{\sum_{i,j} s_{ij}\, a_i a_j}\sqrt{\sum_{i,j} s_{ij}\, b_i b_j}}$$

where $a_i$and $b_j$denote the weighted term frequencies for words $i$and j, and s_{ij}represents the semantic similarity between these words, derived from pre-trained embedding vectors. A low correlation between $S_{lex}$and S_{sem}indicates that the two measures capture different but complementary aspects of consistency. This interpretation is consistent with the theoretical framework of Hansen and McMahon (2016), which highlights that linguistic uniformity and conceptual coherence together strengthen central bank communication credibility.

d. Inter-Annotator Reliability

To assess the reliability of expert-annotated sentiment data, the study uses Fleiss' Kappa (κ), a statistical measure of inter-rater agreement (Fleiss, 1971). The coefficient is defined as

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

where $\bar{P}$ denotes the observed agreement among annotators and $\bar{P}_e$ represents the agreement expected by chance. According to Landis and Koch (1977), κ values between 0.41 and 0.60 indicate moderate agreement. This threshold is sufficient for establishing credible ground truth data, as supported by Gorodnichenko et al. (2023), who validated sentiment classification in monetary communication using expert-labeled datasets.

Through the integration of these methods, the robustness testing framework ensures that each linguistic feature in this study is empirically stable, internally consistent, and replicable. The framework builds upon established practices in computational linguistics and monetary policy communication, while adapting these approaches to the Bank Indonesia context.

Vector Autoregression with Exogenous Variables (VARX)

This approach offers several advantages. First, it captures the dynamic interrelationships among multiple exogenous and endogenous variables, allowing for a comprehensive understanding of the system. Second, by accounting for the impact of external shocks on endogenous variables, the model enhances realism and relevance in policy analysis. Finally, VARX models are capable of capturing both immediate and delayed responses to communication events, making them particularly suitable for evaluating the effectiveness and transmission of central bank messages over time. In general, the VARX (p,s) model can be formulated as follows:

$$y_t = c + \sum_{i=1}^{p} \Phi_i y_{t-i} + \Psi x_t + \epsilon_t$$

where y_t is a $k \times 1$ vector of the endogenous (dependent) variables at time t. The constant term c is a $k \times 1$ vector, representing the intercept of the model. The summation term

captures the relationship between the current values of the endogenous variables and their lagged values, with Φ_i being a $k \times k$ matrix of coefficients for the i-th lagged endogenous variables and $\boldsymbol{y}_{t-i}$ representing a $k \times 1$ vector of the endogenous variables at time $t-i$. The term Ψ represents a $k \times m$ matrix of coefficients for contemporaneous exogenous variables and x_t: An $m \times 1$ vector of the exogenous variables at time t. Finally, $\boldsymbol{\epsilon}_t$: A $k \times 1$ vector of the error terms at time t and p: The number of lagged periods included for the endogenous variables.

Linear Regression

Another model used in this research to examine relationship between central bank communication features and intermediary channels is Linear Regression. Linear regression allows us to assess the contemporaneous relationship between the communication features (such as sentiment) and the NER sentiment analysis result. By using linear regression, we can quantify how variations in these communication features directly influence the intermediary channels, providing a clear understanding of the immediate effect of communication characteristics on stakeholder category group. The linear regression model can be formulated as follows:

$$y_t = \alpha + \sum_{i=1}^{k} \beta_i x_{i,t} + \epsilon_t$$

where y_t represents dependent variable at time t, which is the variable of interest being explained. The independent variable, denoted as $x_{i,t}$, has an associated coefficient β_i, which quantifies the relationship between the independent variable and the dependent variable. The intercept α represents the baseline value of the dependent variable when all independent variables are equal to zero. The error term, ϵ_t, accounts for the unobserved factors that influence the dependent variable but are not captured by the independent variables in the model. This equation thus provides a framework for understanding the relationship between the dependent variable and the various independent variables, with coefficients estimating the strength and direction of thses relationships.

Anomaly Detection

In the evaluation stage, the trained Isolation Forest model is used to compute the anomaly score for each observation by tracing its path length across all isolation trees and converting this value into an anomaly score using the formula proposed by Liu et al. (2008) as follows:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

where $h(x)$ represents the path length, and $E(h(x))$ denotes the average path length across all isolation trees. The path length represents the number of steps required to isolate a data point from the root node within the decision tree structure. Meanwhile, $c(n)$ refers to the average value of $h(x)$ given the sample size n.

However, when compared with the rolling threshold approach, the latter captured a higher proportion of true signals, leading this study to adopt the rolling threshold method as the criterion for determining anomaly values. The average exchange rate over the past 12 periods is calculated as follows:

$$\mu_t = \frac{1}{n} \sum_{i=t-n+1}^{t} x_i$$

this moving average serves as a local mean that represents the normal pattern of exchange rate movements within a short-term horizon. Furthermore, to measure the degree of dispersion in exchange rates over the same horizon, a rolling standard deviation is employed:

$$\sigma_t = \sqrt{\frac{1}{n-1}\sum_{i=t-n+1}^{t}(x_i - \mu_t)^2}$$

this standard deviation captures the actual volatility during the given period, allowing the resulting threshold to be adaptive to prevailing market conditions. The anomaly threshold is defined as follows:

$$\tau_t = \mu_t + k \cdot \sigma_t$$

where $\tau_t$represents the upper control limit that distinguishes exchange rate movements remaining within the range of normal variation from those considered abnormal. The choice of the upper (rather than lower) limit is based on the analytical focus on depreciative pressures on the exchange rate (rupiah depreciation), which, in the context of external stability, is more relevant as an early warning signal. More specifically, in applying this formula, the variable x_t denotes the rupiah exchange rate at period t, representing the primary time series variable under analysis. The parameter n refers to the length of the observation window, also known as the rolling window, which is set to n=12 in this study to capture one-year dynamics of exchange rate movements. Meanwhile, k represents the standard deviation multiplier used to determine the anomaly threshold, with the initial implementation assigned a value of k=1. These definitions together establish the parameters necessary for constructing a consistent and statistically grounded analytical framework for anomaly detection within the exchange rate series. Each observation at time *t*is classified as an anomaly if the actual exchange rate exceeds the defined threshold:

$$A_t = \begin{cases} 1, & \text{if } x_t > \tau_t \\ 0, & \text{if } x_t \leq \tau_t \end{cases}$$

where (A_t) is the binary anomaly indicator. A value of 1 indicates the presence of an anomaly signal, whereas a value of 0 denotes a normal condition.

The Pearson correlation coefficient focuses on the linear prediction of values and the relationship between attributes X and Y (Dudáš, 2024). This relationship can be expressed through the following equation:

$$r_{XY} = \frac{\sum_{t=1}^{T}(X_t - \bar{X})(Y_t - \bar{Y})}{\sqrt{\sum_{t=1}^{T}(X_t - \bar{X})^2}\sqrt{\sum_{t=1}^{T}(Y_t - \bar{Y})^2}}$$

In this analysis, X_t represents the value of the rupiah exchange rate at time t, while Y_t denotes the value of the indicator with a lag of i at the same time t, where $i = 1$ refers to the indicator observed at $t-1$. The symbols $\bar{X}$ and $\bar{Y}$ Indicate the mean values of each respective time series, providing the baseline for assessing deviations from their averages. Meanwhile, T designates the total number of valid observations retained after excluding missing data entries. These components collectively form the basis for computing the Pearson correlation coefficient, which quantifies the linear relationship between the exchange rate variable and its corresponding lagged indicator, thereby revealing the direction and strength of their temporal association. Furthermore, the conceptual formulation of the correlation calculated is as follows:

$$r_{Exchange\ rate,\ indicator_{t-1}},\ r_{Exchange\ rate,\ indicator_{t-2}},\ \cdots,\ r_{Exchange\ rate,\ indicator_{t-6}}$$

The formula allows for the identification of how changes in public search indicators in previous periods correlate with exchange rate movements in the current period (t), thereby indicating whether the indicator possesses a leading characteristic. After constructing the lagged feature variables, the correlation matrix is then computed using the Pearson correlation coefficient (r_{XY}), which mathematically measures the ratio of the covariance between two variables to the product of their standard deviations.

In this study, the mathematical model is constructed under the following specification, where the dependent variable Yi represents the anomaly, such that:

$$Y_i = \begin{cases} 1, & \textit{when the rupiah exchange rate exceeds the threshold (anomaly)} \\ 0, & \textit{when the exchange rate is in a normal condition} \end{cases}$$

The independent variable X_t represents the lags of the leading indicator, specifically the indicator values from $t-1\ to\ t-6$. The probit model is based on the assumption that there exists an unobserved (latent) variable that represents the probability of the event $Y_t = 1$, with the following linear relationship:

$$Y_i^* = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i$$

where ε_i follows a standard normal distribution ($\varepsilon_i \sim N(0,1)$). The relationship between the actual probability of occurrence and the latent variable is expressed as follows:

$$P(Y_i = 1 \mid X_i) = \Phi(\beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik})$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution function. Accordingly, the probit model estimates the probability of an event's occurrence based on the cumulative normal distribution. In this study, the probit model used to estimate each indicator is specified as follows:

$$P(Y_t = 1) = \Phi(\beta_0 + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_6 X_{t-6})$$

The estimation results consist of the coefficients (β_i), z-statistics, and p-value used to assess the significance of each lagged variable. Furthermore, the coefficient (β_i) indicates the direction of the relationship between the indicator and the probability of an anomaly. A positive coefficient β_i>0 suggests that an increase in the leading indicator at lag i raises the likelihood of an exchange rate anomaly, whereas a negative coefficient β_i<0 indicates that an increase in the indicator reduces the likelihood of an anomaly. In the filtering test using the probit model, this approach aims to estimate the probability of an anomaly (exchange rate pressure) as a function of changes in the indicators over several preceding periods. Moreover, the inclusion of indicator lags ($t-1\ to\ t-6$) is intended to capture the leading effect of shifts in public sentiment on exchange rate movements.

Confusion Matrix

The effectiveness of these signals can be evaluated by calculating several categories of confusion matrix:

$$Type\ I\ error = \frac{FN}{(TP + FN)}$$

$$Type\ II\ error = \frac{FP}{(TN + FP)}$$

$$Predicted = \frac{TP}{(TP + FN)}$$

$$Noise\ to\ Signal\ Ratio = \frac{Type\ I\ error}{\%Predicted}$$

The formula is used to evaluate the performance of the anomaly signal detection system. Type I Error measures the proportion of false alarms that appear when there is no actual anomaly, while Type II Error shows how often the system fails to detect the anomaly that occurs. Predicted (or Recall/Sensitivity) reflects the system's ability to accurately capture anomalies. Finally, *the Noise to Signal Ratio (NSR)* compares the rate of Type I error to the accuracy of the prediction, so the lower the NSR, the more effective and reliable the indicator is in detecting the actual anomaly.

The confusion matrix represents a table that allows for a detailed analysis of the performance of a classification model. The confusion matrix is used to evaluate the performances of the models in the data experiments. In addition, it gives a breakdown of the predictions made by the model and the actual ground truth labels. The confusion matrix can be used to calculate various performance metrics, including recall, precision,

accuracy, and F1 score. Here are the definitions of true positive rate (TPR) and false positive rate (FPR):

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN}$$

The TPR, also known as sensitivity or recall, assesses the proportion of positive cases accurately identified by the model. The FPR calculates the proportion of negative cases mistakenly categorized as positive by the model. The confusion matrix provides a thorough perspective of the model's performance, allowing for more in-depth review of its strengths and weaknesses when distinguishing between classes.

Robustness Test For Early Warning System

Another method used for analyzing a model's performance is the receiver operating characteristic (ROC) curve. ROC curve represents true positive rates versus false positive rates to show the trade-off between correctly identifying and false predictions. The area under the curve (AUC) is used to compare the model's classification performances (Seiffert et al., 2010). Specifically, for each threshold, the ROC is calculated as follows:

$$ROC = \frac{True\ Positive\ Rate}{False\ Positive\ Rate} = \frac{\frac{TP}{TP + FN}}{\frac{FP}{FP + TN}} = \frac{\frac{TP}{TP + FN}}{1 - \frac{TN}{FP + TN}} = \frac{Sensitivity\ Rate}{1 - Specificity\ Rate}$$

Technically, the value of AUC (the integral of the area under the ROC curve) ranges from 0.5 to 1. While an AUC of 0.5 can be achieved by a random classifier (e.g., a coin toss), an AUC of 1 means that the model is a perfect classifier. In developing a robust Early Warning System (EWS), the model should ideally exhibit an AUC value closer to 1 than to 0.5; the higher the AUC, the better the model's performance. Furthermore, a robustness check in this EWS framework is conducted by employing usefulness and loss function analyses for each signal indicator. In general, policymakers tend to prefer issuing false alarms rather than missing an actual crisis; therefore, the value of μ can be set higher (for instance, μ=0.8). The accuracy of signal detection can then be assessed through the usefulness measure, following the approach proposed in previous studies. Sarlin (2013):

$$U_r(\mu) = \frac{U_a(\mu)}{\min(\mu P_1, (1 - \mu) P_2)}$$

The *usefulness* (U_r) measure represents the difference between the loss incurred by policymakers when no model is employed ($\min(\mu P_1, (1 - \mu) P_2)$) and the loss incurred when an Early Warning System (EWS) is utilized ($L(\mu)$).

Appendix 2: Figure

Figure 3. 1 Research Framework

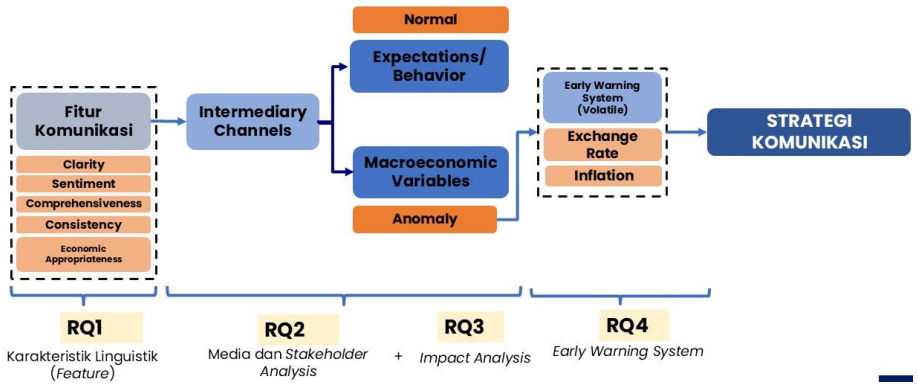


Figure 4. 31 Integrated Result Framework Linking Communication Features, Media Sentiment, and Econometric Impact Analysis

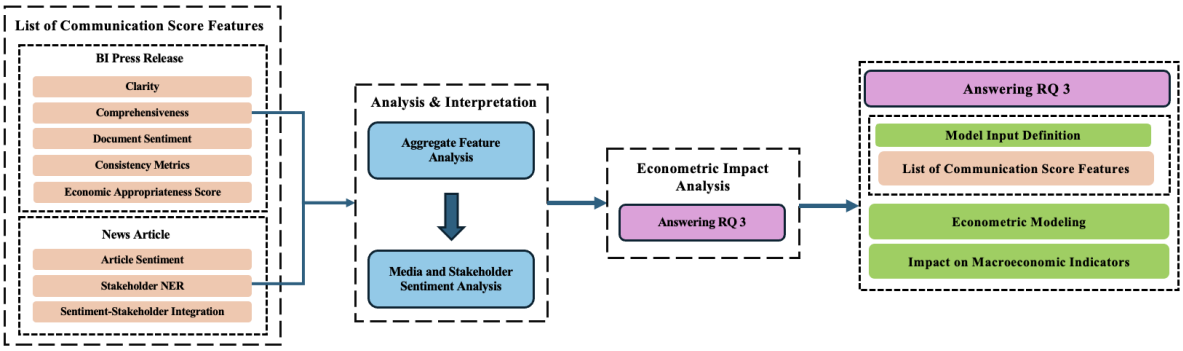


Figure 4. 32 Time-Series Trend of Overall Clarity Scores (2019–2024)

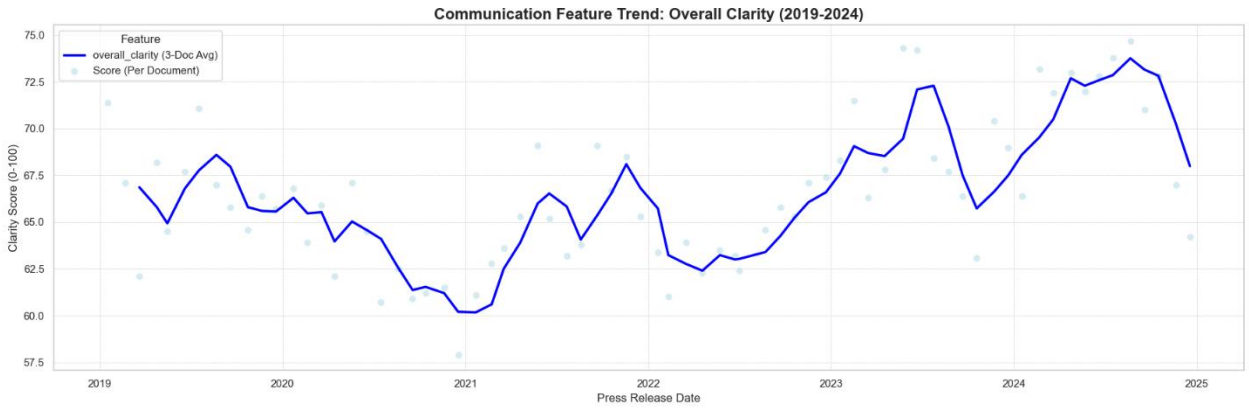


Figure 4. 33 Time-Series Trend of Economic Appropriateness Scores (2019–2024)

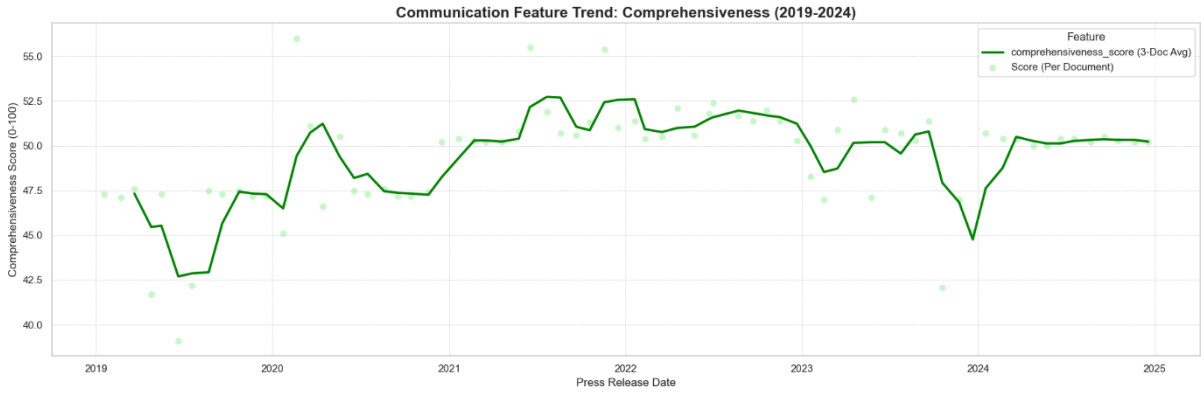


Figure 4. 34 Time-Series Trend of Economic Appropriateness Scores (2019–2024)

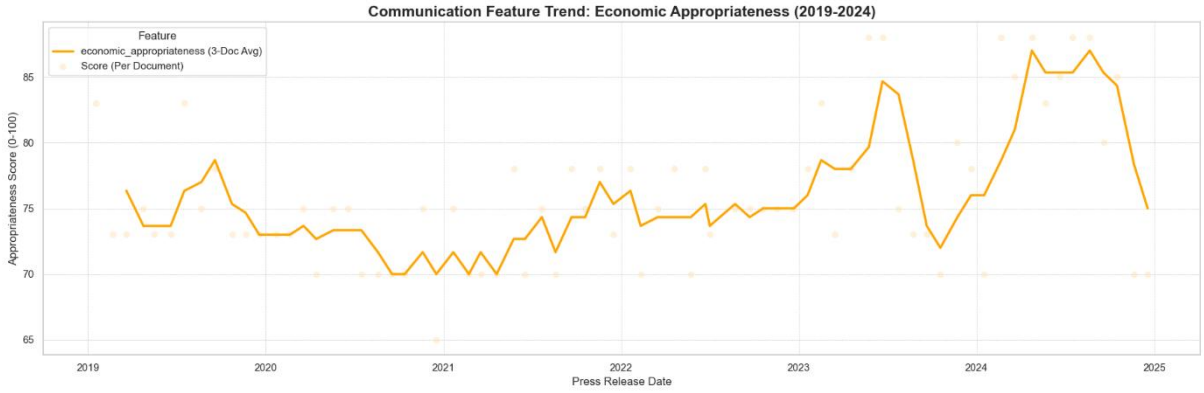


Figure 4. 35 Time-Series Trend of Sentiment Consensus Scores (2019–2024)

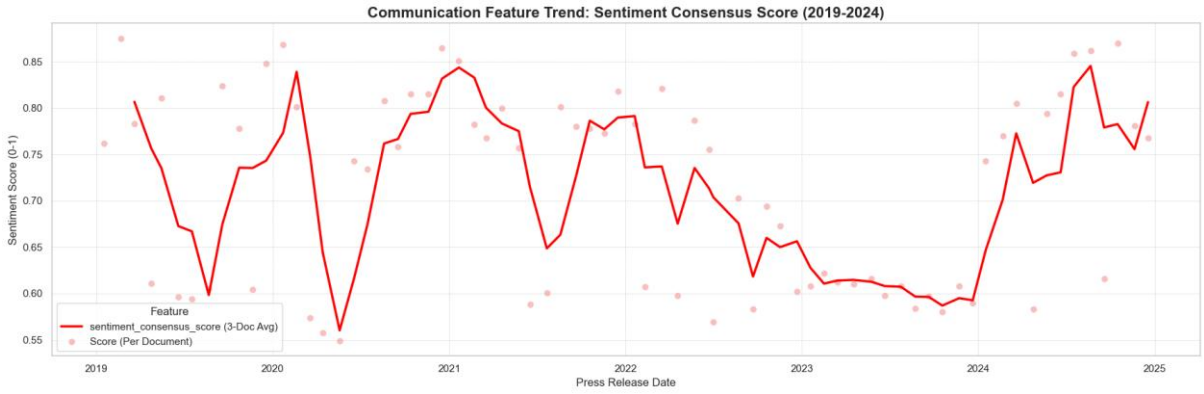


Figure 4. 36 Time-Series Trend of Lexical Consistency Scores (2019–2024)

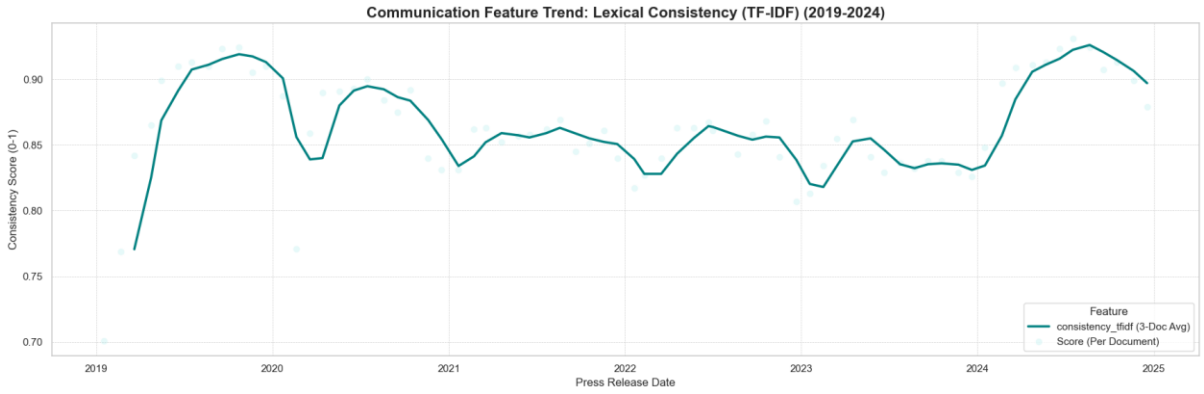


Figure 4. 37 Time-Series Trend of Semantic Consistency Scores (2019–2024)

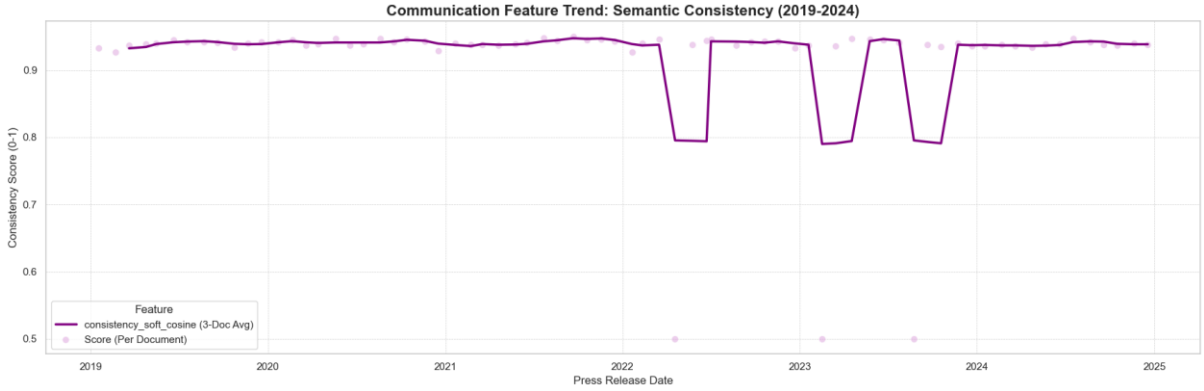


Figure 4. 38 Temporal Dynamics of Stakeholder Sentiment in Economic News (2019-2024)

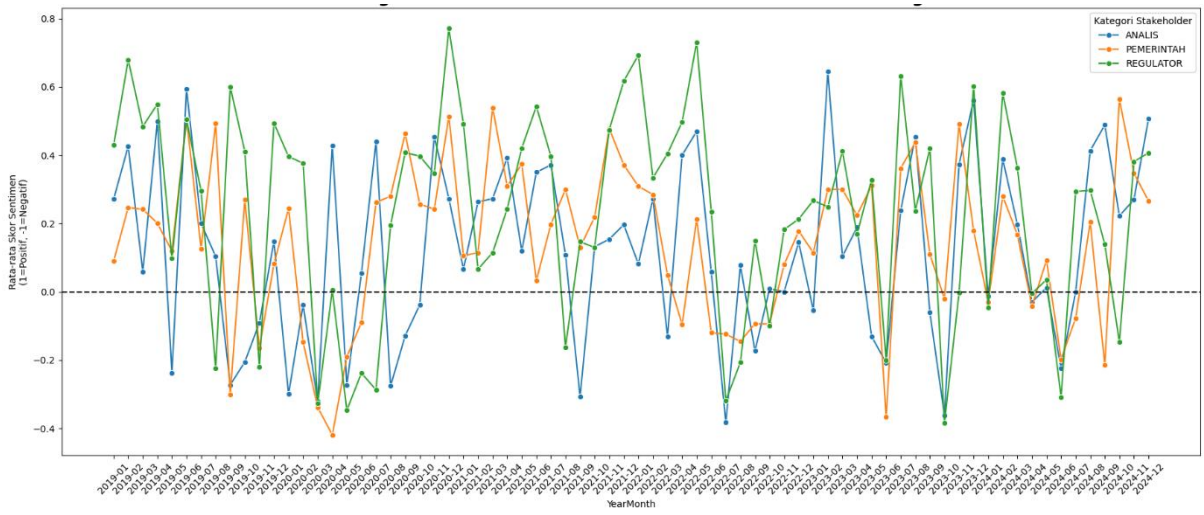


Figure 4. 39 Sentiment Distribution among Domestic Analysts and Research Institutions

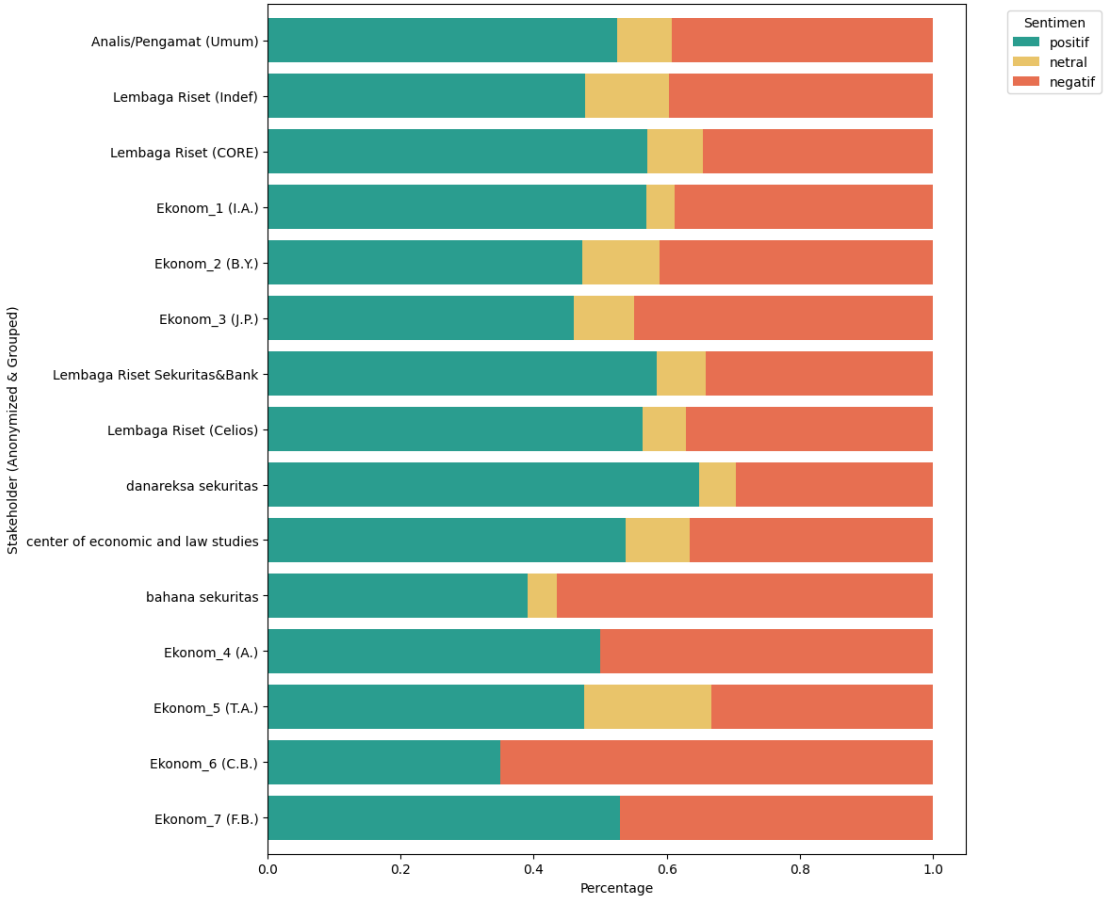


Figure 4. 40 Top Media Platforms with High Frequency of Critical or Negative Coverage

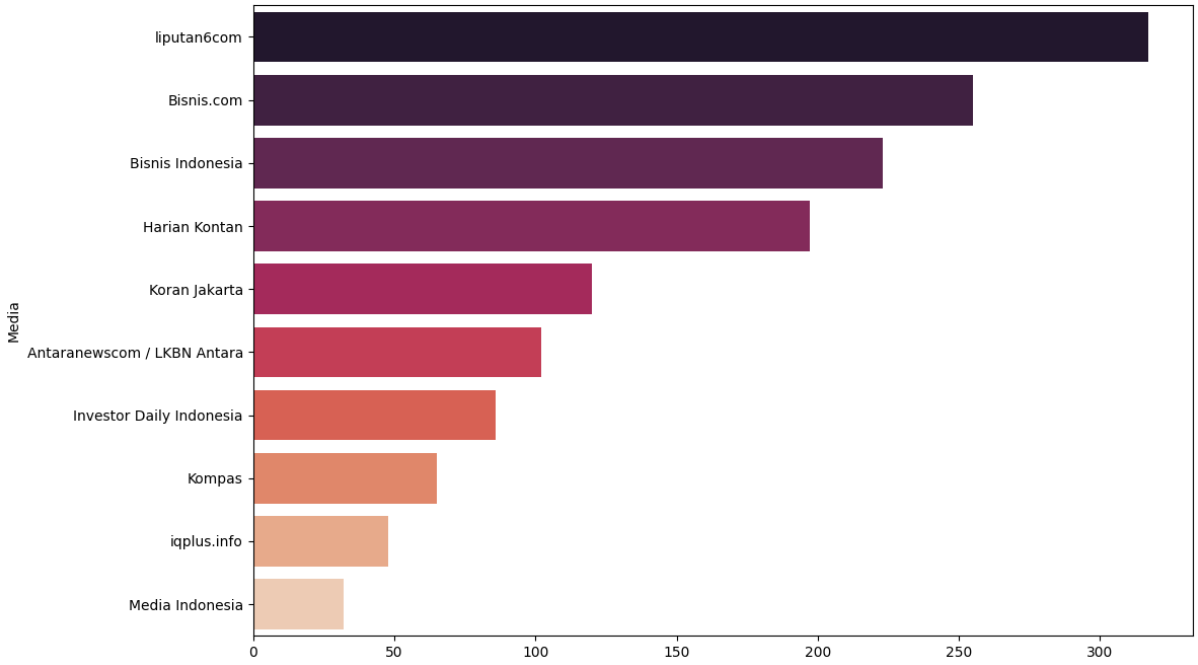


Figure 4. 41 Framework Communication Impact Analysis

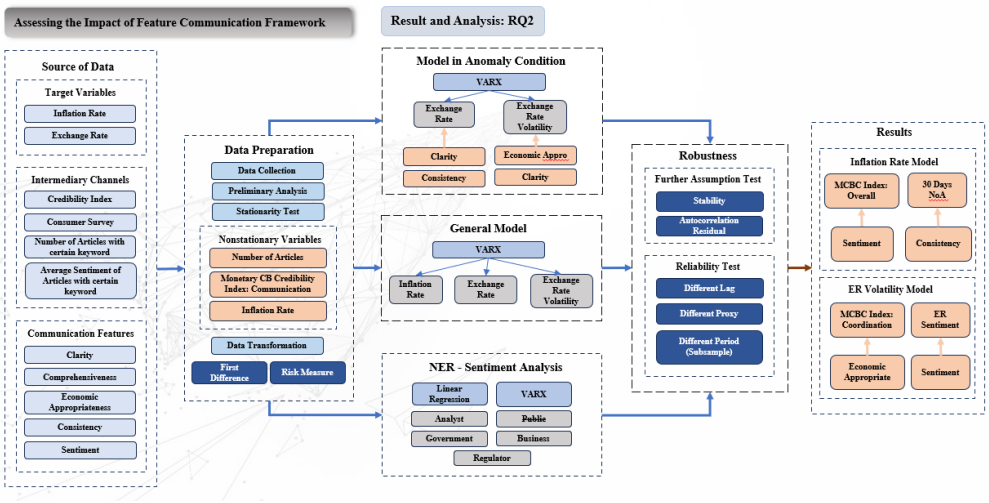
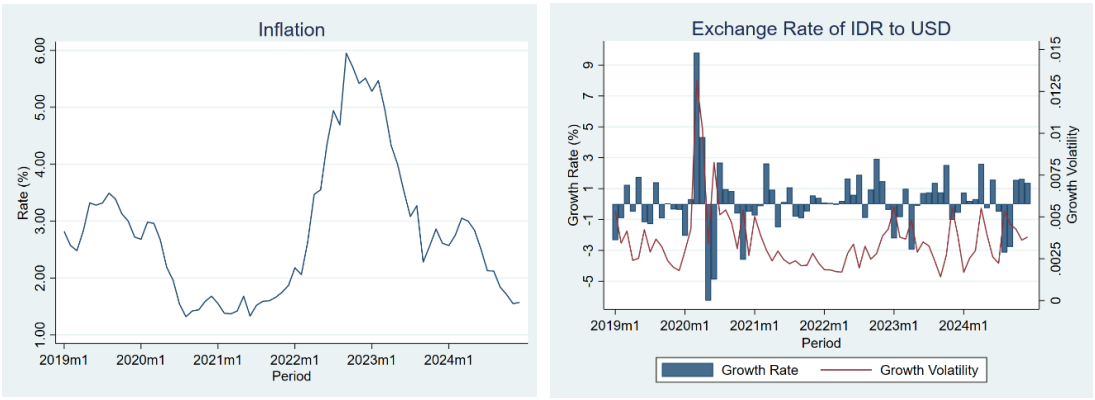


Figure 4. 42 Periodic Macroeconomic Data



Inflation Exchange Rate

Figure 4. 43 Periodic Number and Sentiment of News Articles

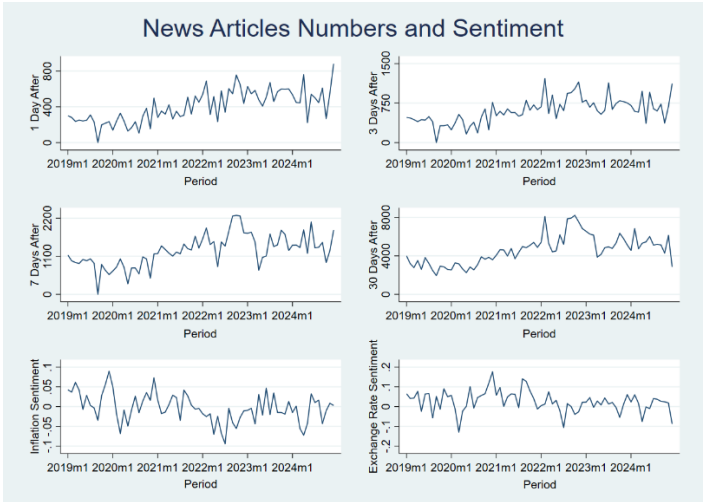


Figure 4. 44 Monetary Central Bank Credibility Index Dynamics

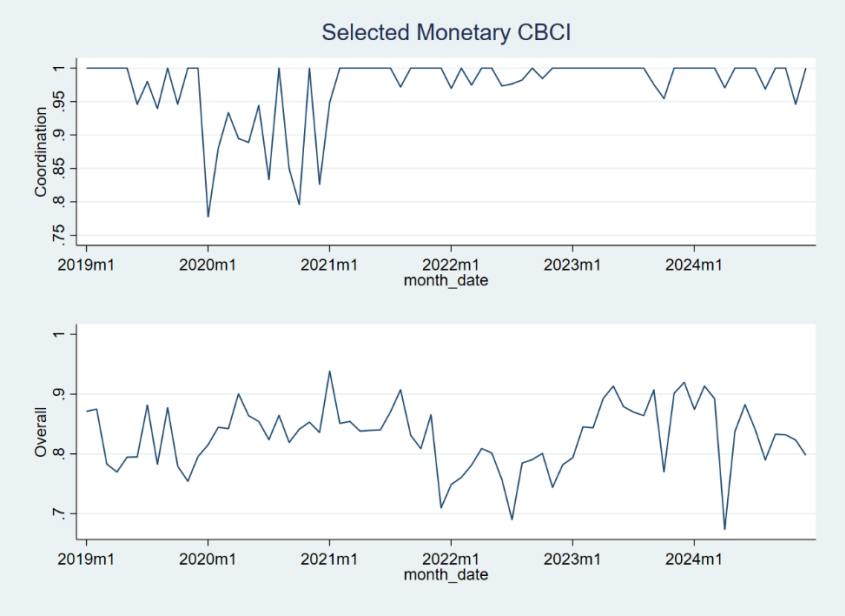


Figure 4. 45 IRF Function for Inflation Modelling

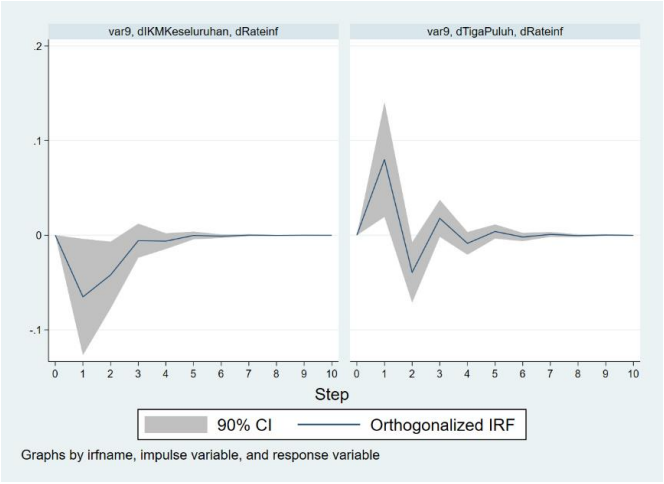


Figure 4. 46 Stability Test in Inflation Model

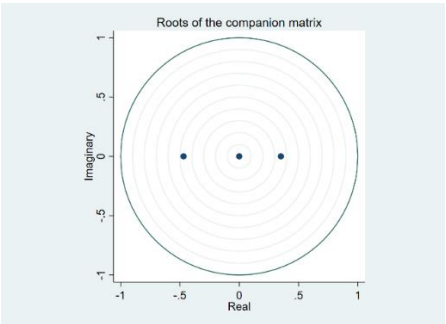


Figure 4. 47 Robustness Tests in Inflation Modelling

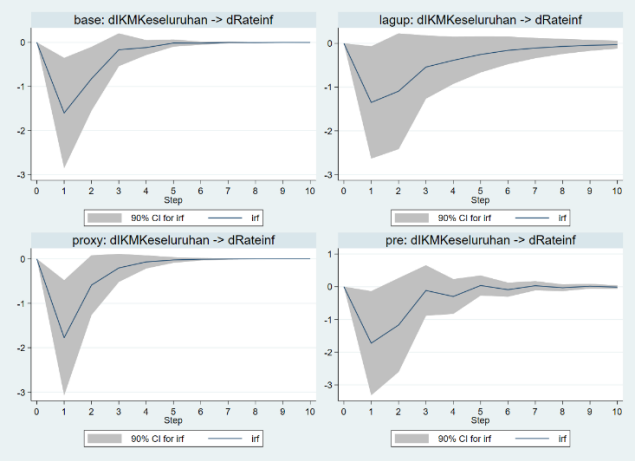


Figure 4. 48 IRF of Exchange Rate Modelling

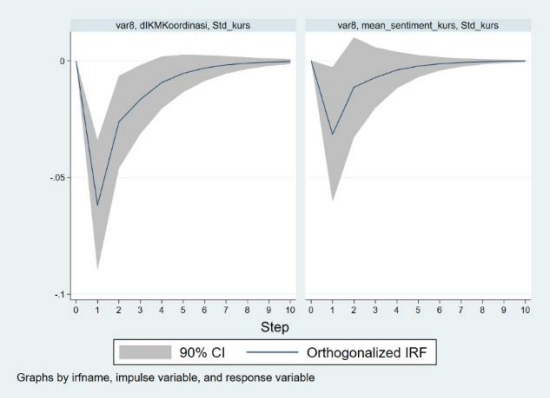


Figure 4. 49 Stability Test for Exchange Rate Model

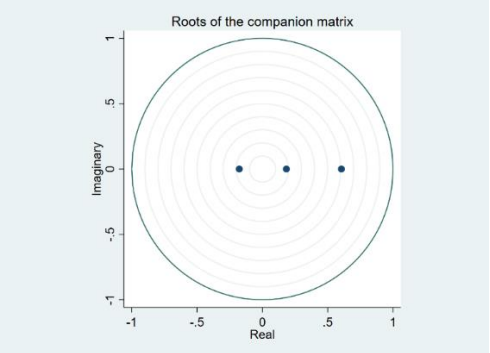


Figure 4. 50 Robustness Test in Exchange Rate Volatility Modelling

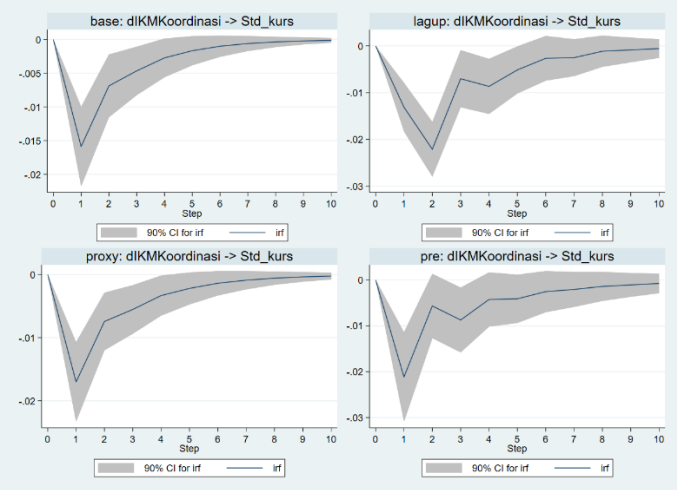


Figure 4. 51 The Architecture of the Early Warning System Development

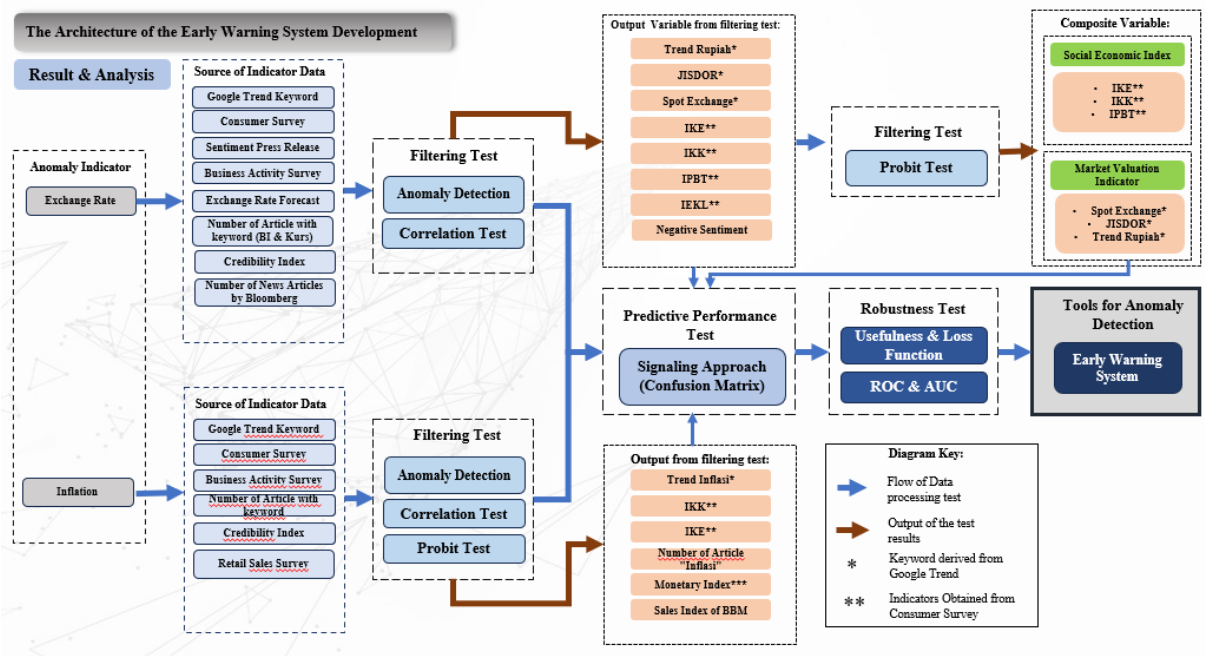


Figure 4. 52 Comparison of Inflation Anomaly Periods

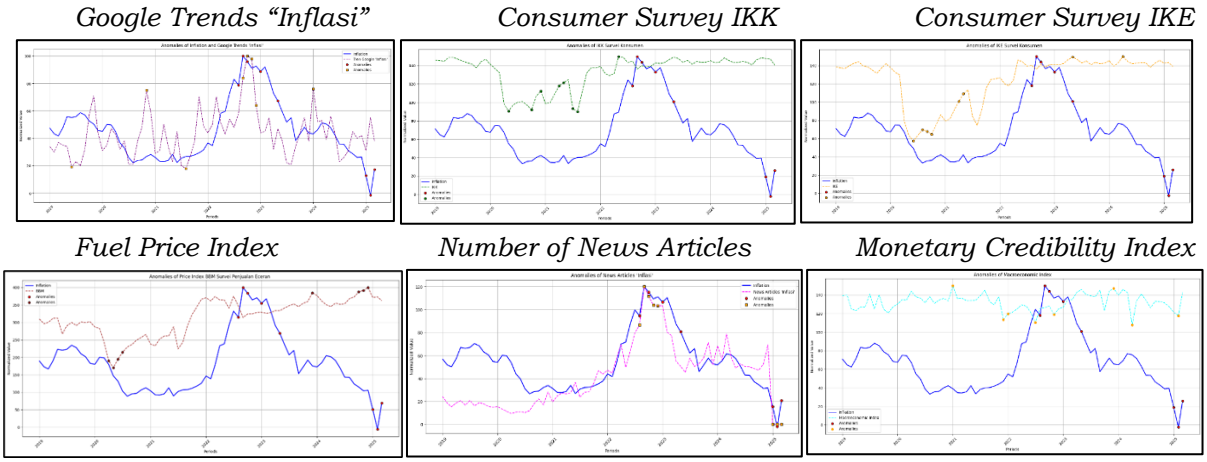
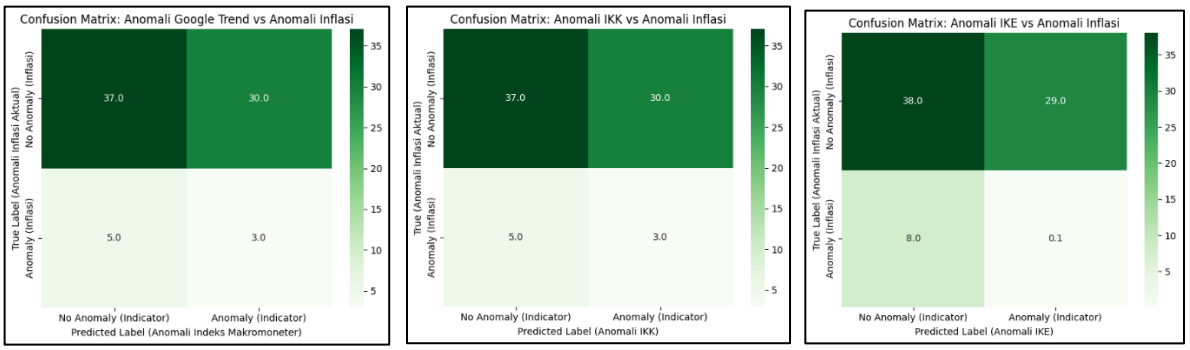


Figure 4. 53 Confusion Matrix of the Signaling Approach Results

The Google Trends Indicator *The IKK Indicator* *The IKE Indicator*



The Fuel Sales Index *the Number of News Articles* *the Monetary Credibility Index*

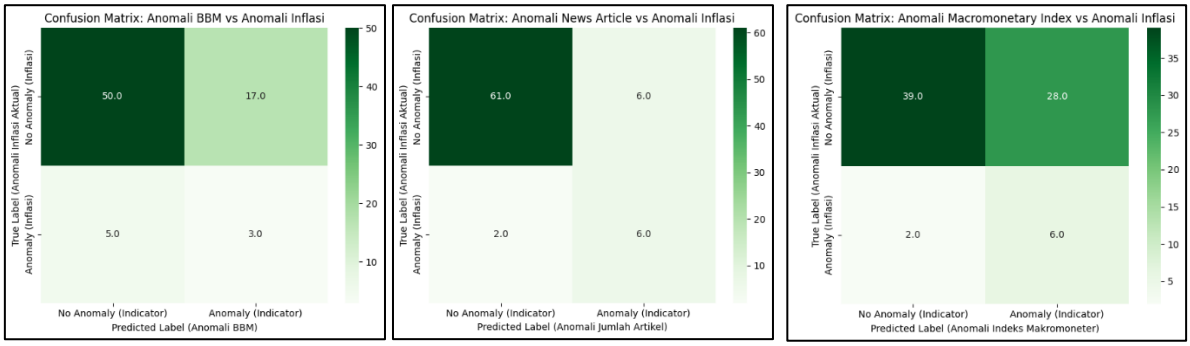
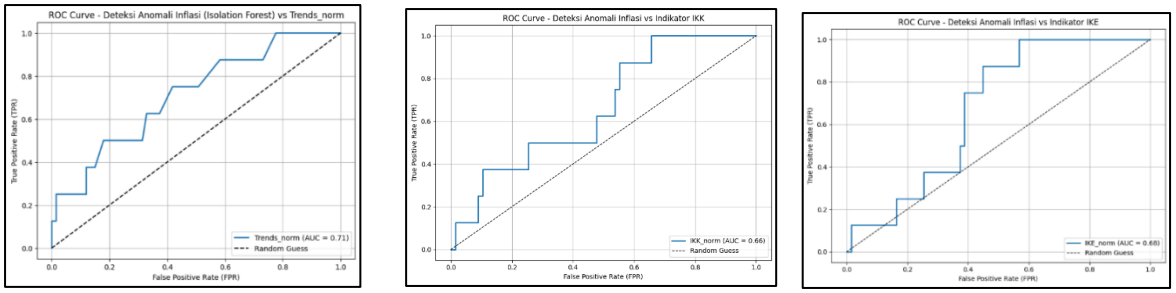


Figure 4. 54 ROC Curve and AUC Value

the Google Trends “Inflasi” the IKK Indicator the IKE Indicator



the Fuel Sales Index *the Number of News Articles* *the Monetary Credibility Index*

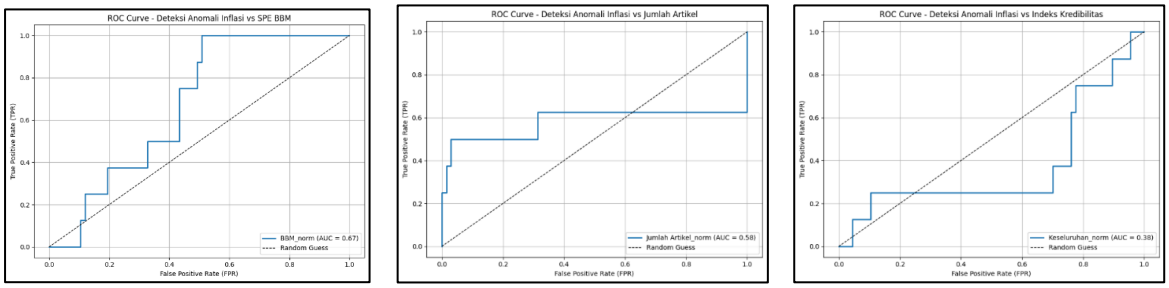


Figure 4. 55 Indicators of EWS and Exchange Rate Curve

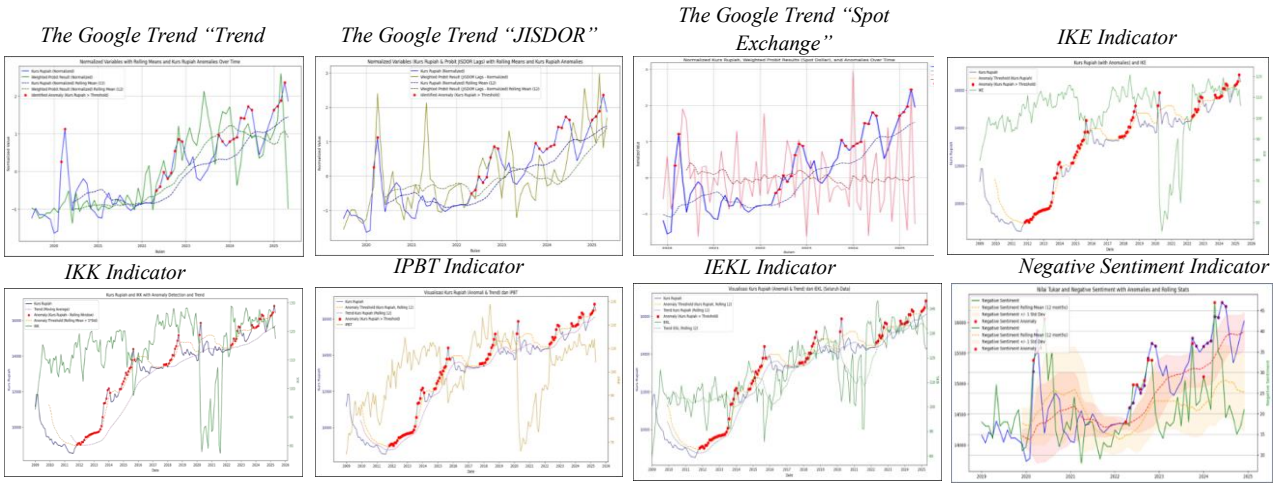


Figure 4. 56 Weighted Probit Result

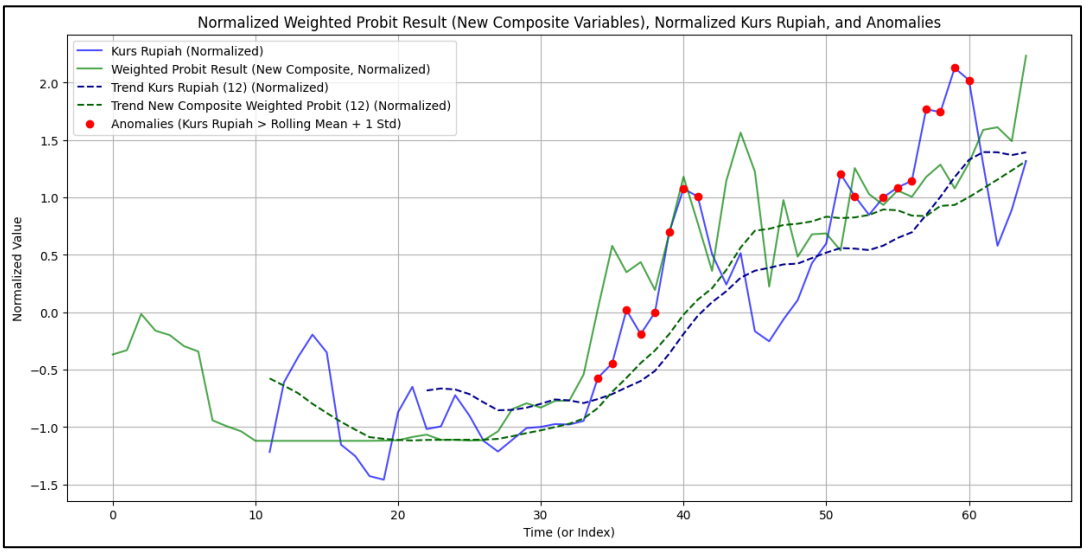


Figure 4. 57 Signal Analysis from each indicator

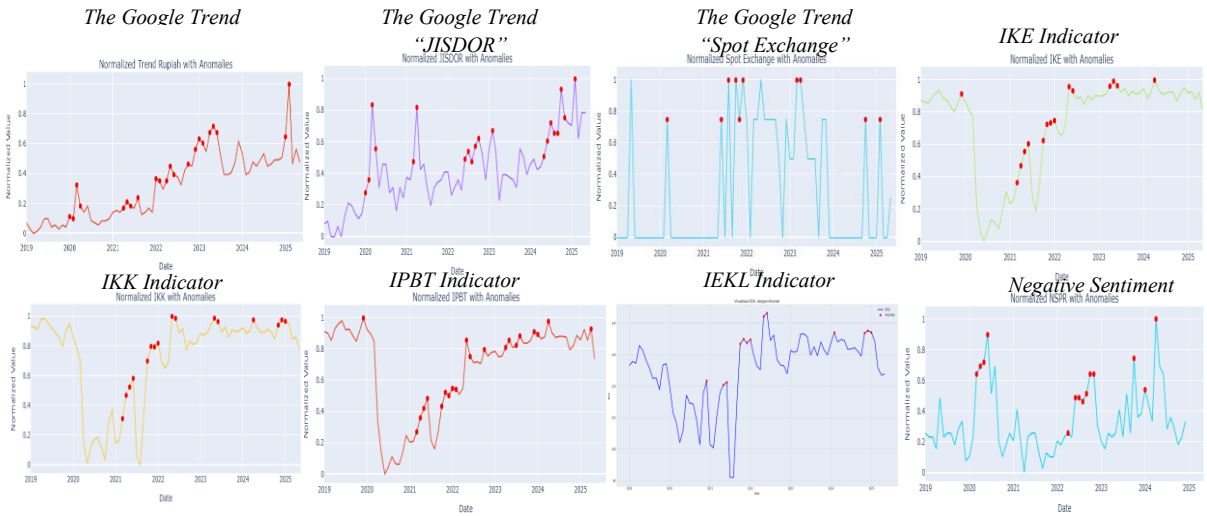
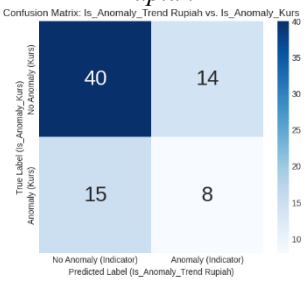
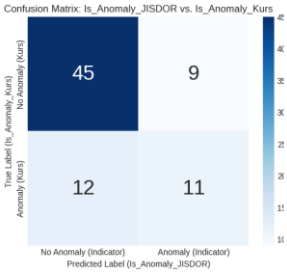


Figure 4. 58 Confusion Matrix of Exchange Rate

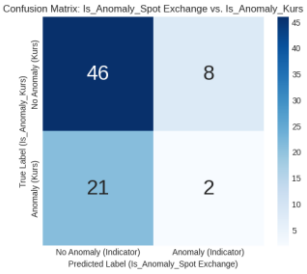
The Google Trend "Trend Rupiah"



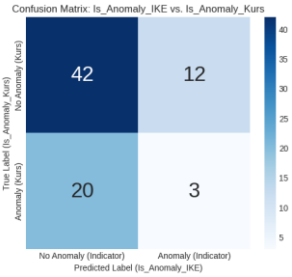
The Google Trend "JISDOR"



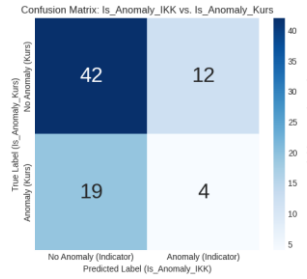
The Google Trend "Spot Exchange"



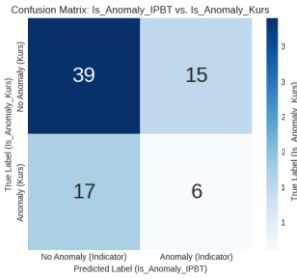
IKE Indicator



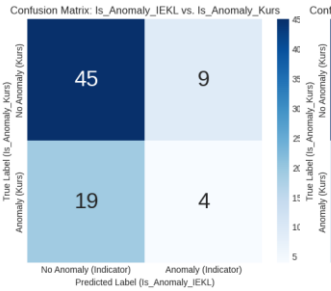
IKK Indicator



IPBT Indicator



IEKL Indicator



Negative Sentiment Indicator

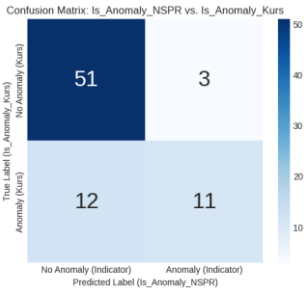


Figure 4. 59 The ROC and AUC Results of Each Indicator in EWS

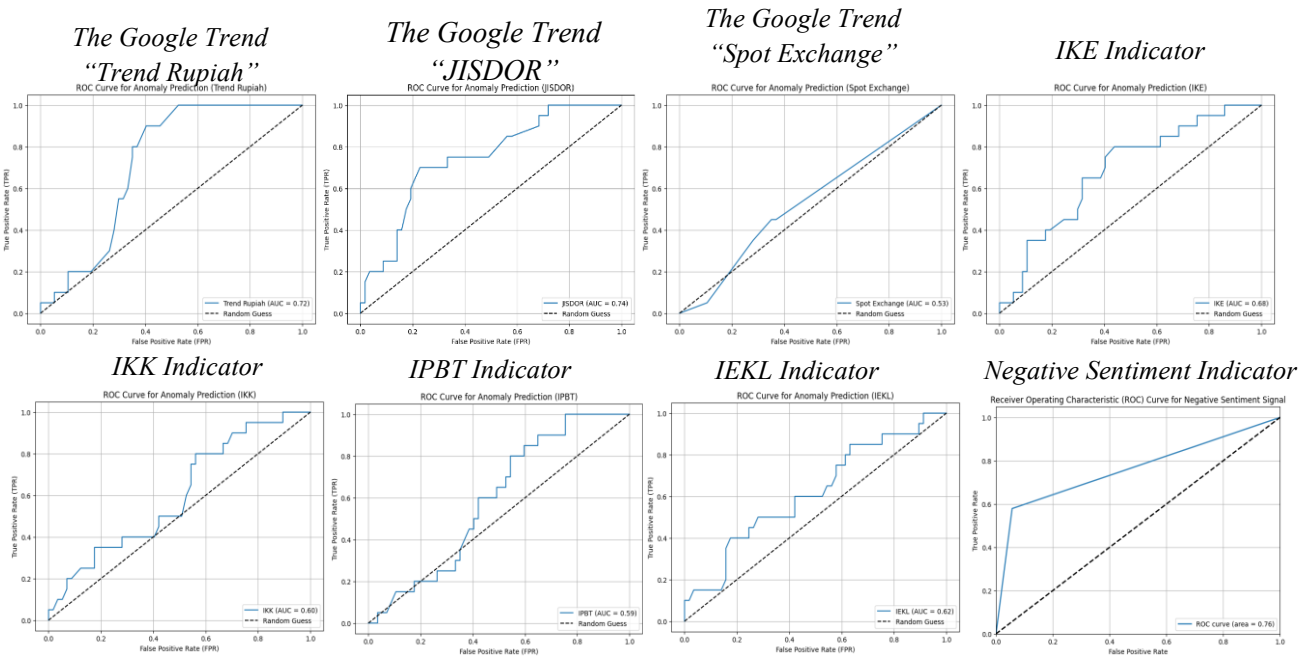
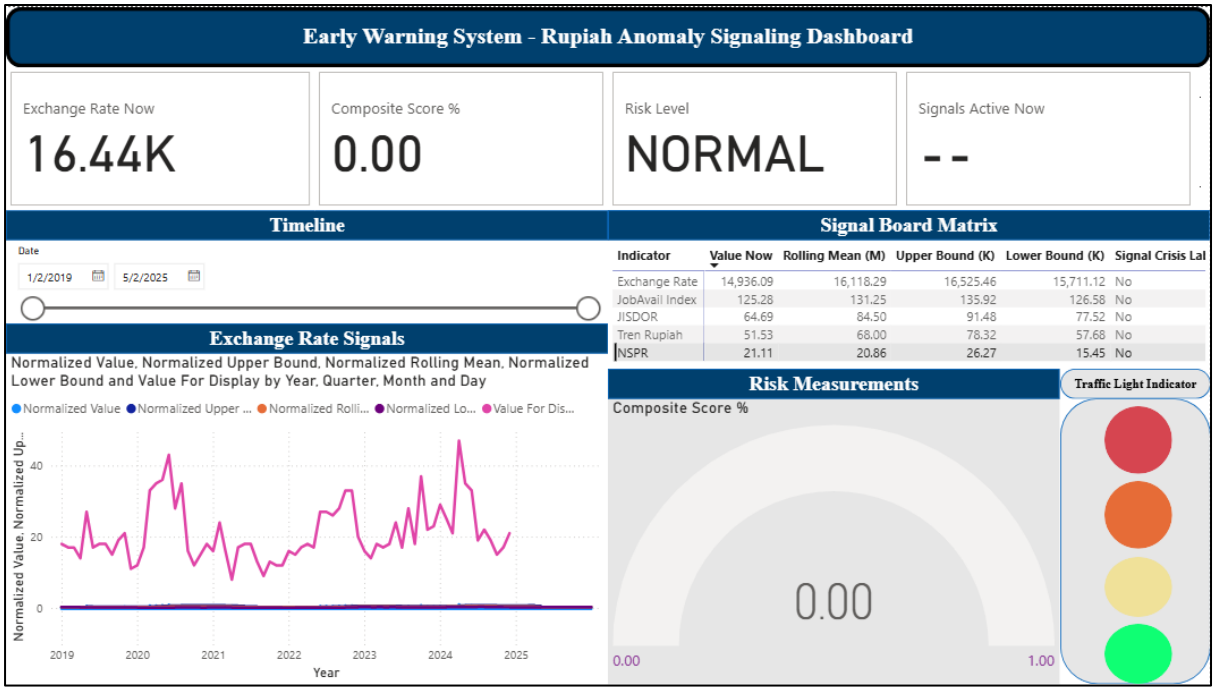


Figure 4. 60 Digital visualization of early warning system



Appendix 3: Table

Table 3.2: NER Stakeholder Category

Category Code	Stakeholder Group	Description
PEMERINTAH	Government	Ministries, state officials, and fiscal authorities contributing to macroeconomic policy coordination.
REGULATOR	Financial Regulators	Domestic regulatory and statistical agencies overseeing financial systems and macroprudential policy.
ANALIS	Analysts and Experts	Economists, academics, research institutions, and rating agencies providing economic analysis and commentary.
PELAKU_USAHA	Business Sector	Private sector actors, corporate executives, and trade associations representing market and industrial responses.
MASYARAKAT	Public and Civil Groups	General public, workers, consumers, and public organizations representing social sentiment and welfare perspectives.

Table 3.23: Explanation of the Confusion Matrix

Category	Logic	Explanation
True Positive	Signal = True and Anomaly = True	Indicator issues a signal, and the anomaly indeed occurs.
False Positive	Signal = True and Anomaly = False	The indicator gives a signal, but an anomaly does not occur (false alarm).
False Negative	Signal = False and Anomaly = True	The indicator does not give a signal, even though an anomaly occurs (a missed signal).
True Negative	Signal = False and Anomaly = False	The indicator does not provide a signal, and anomalies do not occur.

Table 3.33 Data

Data Type	Description and Function	Source
Bank Indonesia Press Release Corpus	A collection of official Bank Indonesia press releases issued between 2019 and 2024. This corpus serves as the main object of analysis for policy communication features, including clarity, comprehensiveness, sentiment, economic appropriateness, and semantic consistency.	Bank Indonesia
Online News Article Corpus	A compilation of economic and financial news articles published by major Indonesian media outlets during the same period. Articles are selected based on their relevance to monetary policy, inflation, and exchange rate topics. This corpus is used to examine media resonance and stakeholder sentiment toward Bank Indonesia's communications.	Cyberlibs BI
Macroeconomic Data	Data series, of inflation rate and exchange rate, taken from Bank Indonesia's dataset during the same period. This data is used to assess	Bank Indonesia, BPS

Data Type	Description and Function	Source
	the impact of central bank communication features to macroeconomic variables through intermediary channels.	
Central Bank Credibility Index	A compilation of scores developed by research team using structured scoring framework and administrative sources. The index comprises of sub-indices for monetary policy, macroprudential policy, and the payment system policy. These scores serve as intermediary channels in the communication impact analysis.	Bank Indonesia
Consumer Survey Index	Monthly measures sourced from Bank Indonesia's regular household survey, which track perceptions of current economic conditions and near-term expectations. These scores are employed as proxies for household perceptions and expectations within the intermediary channel framework.	Bank Indonesia
Google Trend	The monthly search volume for certain keywords obtained from Google Trends serves as a proxy for capturing the dynamics of public attention to fluctuations in the rupiah exchange rate. The keywords used are "Inflation", "Trend Rupiah", "JISDOR", and "Spot Exchange".	Google

Table 4.1 Robustness Test Result

Robustness Test	Metric	Feature Tested	Value (ρ / κ)
Stability vs Document Length	Spearman's ρ	Clarity	0.062
Stability vs Document Length	Spearman's ρ	Comprehensiveness	0.227
Stability vs Document Length	Spearman's ρ	Economic Appropriateness	0.061
Proxy Validation	Spearman's ρ	Clarity (Flesch vs Overall)	0.970
Proxy Validation	Spearman's ρ	Comprehensiveness (Topic vs Overall)	0.753
Inter-Model Comparison	Spearman's ρ	Consistency (Lexical vs Semantic)	0.214
Inter-Annotator Reliability	Fleiss' κ	Sentiment Annotation	0.4086

Table 4.2 VARX result for Inflation Rate Model

VARIABLES	Model 1 – Inflation Rate		
	Inflation Rate	Monetary CBCI (Overall)	30 Days After
L.Inflation_Rate	0.0753	-0.00557	-182.9
	(0.537)	(0.768)	(0.590)
L.Monetary_CBCI_Overall	-1.602**	0.320***	-2,072
	(0.0349)	(0.00636)	(0.326)
L.30_Days_After	9.21e-05**	3.03e-06	-0.514***
	(0.0268)	(0.637)	(8.67e-06)
Economic_Appropriateness	0.00433	0.00153	-11.99
	(0.556)	(0.177)	(0.557)
Clarity	-1.782	0.376	-4,558
	(0.315)	(0.170)	(0.355)
Comprehensiveness	0.00804	0.000383	7.907
	(0.262)	(0.729)	(0.691)
Consistency	-0.611	-0.0692	3,489***
	(0.130)	(0.267)	(0.00188)
Sentiment	0.247	0.0767*	-1,003
	(0.407)	(0.0948)	(0.224)
Constant	2.072	0.283	1,617
	(0.140)	(0.192)	(0.679)
Observations	70	70	70

pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.3 VARX results for Exchange Rate Volatility Model

VARIABLES	Model 2 – Exchange Rate Volatility		
	ER Volatility	Monetary CBCI Coordination	News Sentiment
L.Exchange_Rate_Volatility	0.279***	-9.080***	-0.896
	(0.00647)	(0.00302)	(0.794)
L.MCBCI_Coordination	-0.0159***	0.147	0.0183
	(9.51e-06)	(0.170)	(0.879)
L.News_Sentiment	-0.00665*	-0.0826	0.186
	(0.0684)	(0.448)	(0.128)
Economic_Appropriateness	4.00e-05	0.00226**	-0.000462
	(0.209)	(0.0174)	(0.665)
Clarity	0.000131	0.00254	-0.00205
	(0.108)	(0.298)	(0.994)
Consistency	0.00000610	-0.000166	-0.000595
	(0.731)	(0.754)	(0.316)
Comprehensiveness	-0.0000486	0.00123	-0.00325
	(0.459)	(0.531)	(0.139)
Sentiment	-0.00000250	-0.000334	0.000858**
	(0.820)	(0.307)	(0.0197)
Constant	0.0114	0.510**	0.272
	(0.159)	(0.0340)	(0.314)

pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.4 Linear Regression Result for Further NER Analysis

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Analyst**	**Public**	**Government**	**Regulator**	**Business**
Sentiment	0.352**	0.190	0.382***	0.512***	0.300*
	(0.0451)	(0.261)	(0.00743)	(0.00674)	(0.0680)
Constant	0.0606	0.168***	0.0800**	0.142***	0.177***
	(0.148)	(0.000161)	(0.0197)	(0.00167)	(2.06e-05)
Observations	72	72	72	72	72

Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.5 Correlation Value Analysis (Inflation)

Indicators	**Correlation**
Google Trend ”Inflasi” lag 2	0.48
IKK Consumer Survey lag 2	0.46
IKE Consumer Survey lag 2	0.48
Fuel Sales Index lag 6	0.48
Number of News Articles lag 4	0.6
Monetary Credibility Index lag 4	-0.47

Table 4.6 Probit Analysis (Inflation)

Probit Analysis						
Indicators	**coef**	**Std error**	z	$p > \|z\|$	**[0.025**	**0.975]**
Google Trends “Inflasi” lag 2	0.0594	0.018	3.25	0.001	0.024	0.095
IKK Consumer Survey lag 2	-0.3215	0.145	-2.21	0.027	-0.607	-0.036
IKE Consumer Survey lag 2	0.2964	0.134	2.216	0.027	0.034	0.559
Fuel Sales Index lag 6	0.1057	0.05	2.134	0.033	0.009	0.203
Number of News Articles lag 4	0.029	0.009	3.093	0.002	0.011	0.047
Monetary Credibility Index lag 4	-12.031	4.778	-2.52	0.012	-21.4	-2.666

Table 4.7 Predicted Value and NSR (Inflation)

Indicator	**Type 1 error**	**Type 2 error**	**Predicted**	**NSR**
Google Trend 'Inflasi'	0.448	0.152	0.375	0.404
IKK	0.448	0.152	0.375	0.404
IKE	0.448	0.166	0.125	1.329
Fuel Sales Index	0.254	0.250	0.375	0.667
Number of News Articles	0.090	0.167	0.750	0.222
Monetary Credibility Index	0.418	0.059	0.750	0.078

Table 4.8 Usefulness and Loss Score (Inflation)

Indicator	Usefullness	Loss Score
Google Trend "Inflasi"	0.2927	0.3488
IKK Consumer Survey	0.2927	0.3488
IKE Consumer Survey	0.0677	0.3768
Fuel Sales Index	0.3121	0.5254
Number of News Article	0.666	0.343
Monetary Credibility Index	0.6332	0.1598

Table 4.9 Summary of Robustness Tests (Inflation)

Indicator	Robustness Test	
	AUC	Usefullness
Google Trend "Inflasi"	0.71	0.29
IKK Consumer Survey	0.66	0.29
IKE Consumer Survey	0.68	0.07
Fuel Sales Index	0.67	0.31
News Article	0.58	0.67
Monetary Credibility Index	0.38	0.63

Table 4.10 Correlation Analysis Result (Exchange Rate)

Indicator	Source	Lagged	Correlation
Trend Rupiah	Keyword Google Trend	L6	0.77
JISDOR		L1	0.72
Spot Exchange		L6	0.5
Consumer Confidence Index	Consumer Survey	L4	0.5
Current Economic Confidence Index (IKE)		L6	0.54
Durable Goods Purchasing *Index*)		L1	0.5
Job Availability Expectations Index (IEKL)		L4	0.72
Negative Sentiment	Press Release Data	L1	0.51

Table 4.11 Probit Regression Result with Composite Variables

	coef	std error	z	P|z|	[0.025	0.975]
const	-12.1918	4.367	-2.792	0.005	-20.751	-3.633
Social economic index	0.0288	0.012	2.32	0.02	0.004	0.053
Market valuation indicator	0.0304	0.018	1.687	0.092	-0.005	0.066

Table 4.12 Signal Result Based on the Signaling Approach Analysis

Indicator	Type 1 error	Type 2 error	Predicted	NSR
Trend Rupiah	0.6522	0.2593	34.78%	0.7454
JISDOR	0.5217	0.1667	47.83%	0.3485
Spot Exchange	0.913	0.1481	8.70%	1.7037
IKE (Current Economic Confidence Index)	0.8696	0.2222	13.04%	1.7037
IKK (Consumer Confidence Index)	0.8261	0.2222	17.39%	1.2778
IPBT (Durable Goods Purchasing Index)	0.7391	0.2778	26.09%	1.0648
IEKL (Job Availability Expectations Index)	0.8261	0.1667	17.39%	0.9583
Sentiment Negative Press Release	0.6857	0.08108	31.42%	0.0218

Table 4.13 Comparison of the usefulness and loss scores per Indicators

Indicators	Usefulness	Loss Function
Trend Rupiah	0.436	1.564
JISDOR	0.790	1.210
Spot Exchange	0.026	1.974
Consumer Confidence Index	0.039	1.961
Current Economic Confidence Index (IKE)	0.126	1.874
Durable Goods Purchasing *Index*)	0.244	1.756
Job Availability Expectations Index (IEKL)	0.181	1.819
Negative Sentiment	0.901	1.099