Australian National University
Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

When Words Are Not Enough: Emotional Spillovers From Fed Chair Communication to Global Financial Markets

CAMA Working Paper 86/2026
October 2026

Yiru Chen
University of Leeds
Centre for Applied Macroeconomic Analysis, ANU

Muhammad Ali Nasir
University of Leeds
Peking University
University of Cambridge
Centre for Applied Macroeconomic Analysis, ANU

Xingjie Wei
University of Leeds

Abstract

Given the important role of the Federal Reserve Chair, their communications may generate substantial cross-border effects on international financial markets. Building on this perspective, this article investigates whether emotional signals conveyed by the Fed Chair during FOMC post-meeting press conferences generate emotional spillovers to global financial markets. Using press conferences held between 2011 and 2025, we construct multimodal emotional indices from textual, vocal, and facial cues using machine learning techniques, and examine their high-frequency effects on exchange rates and equity markets across six major economies. Using local projection methods with minute-level event windows, the results show that Fed Chair emotional signals have immediate and economically significant impacts on global financial markets. Among different modalities, facial and textual emotions exhibit stronger and more robust effects than vocal signals. In particular, positive emotional cues are generally associated with increases in equity ETF returns, while exchange rate responses display substantial cross-country heterogeneity, with emerging market currencies tending to appreciate. We also provide evidence that the effects of emotional cues are non-additive and depend on cross-modal interactions, with vocal signals conditioning the impact of textual and facial emotions. Moreover, the findings show that the transmission of emotional spillovers is shaped by macro-financial vulnerability and global risk conditions. On the one hand, economies with higher vulnerability exhibit stronger responses to emotional shocks, indicating that macro-financial vulnerability acts as a moderating mechanism. However, the effects of emotional signals, especially facial and textual emotion, are amplified during periods of higher global uncertainty. Our main findings remain robust across several robustness checks. In particular, for vocal and facial emotions, coefficient equality is not rejected in 85.8% and 78.3% of the market horizon comparisons, respectively, indicating that the main findings are not systematically driven by the choice of emotion extraction algorithms.

Keywords

central bank communication, multimodal signals, emotional spillover, high-frequency identification, machine learning, facial emotion analysis

JEL Classification

E44, E52, E58

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

When Words Are Not Enough: Emotional Spillovers from Fed Chair Communication to Global Financial Markets

Yiru Chen *

Leeds University Business School, University of Leeds, United Kingdom
Centre for Applied Macroeconomic Analysis, Australian National University, Australia
bdzh0534@leeds.ac.uk

Muhammad Ali Nasir

Leeds University Business School, University of Leeds, United Kingdom
National School of Development, Peking University, China
Department of Land Economy, University of Cambridge, United Kingdom
Centre for Applied Macroeconomic Analysis, Australian National University, Australia
M.A.Nasir@leeds.ac.uk

Xingjie Wei

Leeds University Business School, University of Leeds, United Kingdom
X.Wei1@leeds.ac.uk

*Corresponding author. Email: bdzh0534@leeds.ac.uk

Acknowledgements: We thank the participants at the seminars hosted by Tsinghua University and the University of Leeds, as well as participants at the 57th Annual Conference of the Money, Macro and Finance Society, for their valuable feedback on earlier versions of this paper. We are particularly grateful to Stephen Millard for his insightful comments and suggestions. We also gratefully acknowledge the financial support provided for participation in the 2026 Conference on Macroeconomic Policy, organised by the Swiss National Bank and the Study Center Gerzensee in partnership with the *Journal of Monetary Economics*. Any remaining errors are ours.

Abstract

Given the important role of the Federal Reserve Chair, their communications may generate substantial cross-border effects on international financial markets. Building on this perspective, this article investigates whether emotional signals conveyed by the Fed Chair during FOMC post-meeting press conferences generate emotional spillovers to global financial markets. Using press conferences held between 2011 and 2025, we construct multimodal emotional indices from textual, vocal, and facial cues using machine learning techniques, and examine their high-frequency effects on exchange rates and equity markets across six major economies. Using local projection methods with minute-level event windows, the results show that Fed Chair emotional signals have immediate and economically significant impacts on global financial markets. Among different modalities, facial and textual emotions exhibit stronger and more robust effects than vocal signals. In particular, positive emotional cues are generally associated with increases in equity ETF returns, while exchange rate responses display substantial cross-country heterogeneity, with emerging market currencies tending to appreciate. We also provide evidence that the effects of emotional cues are non-additive and depend on cross-modal interactions, with vocal signals conditioning the impact of textual and facial emotions. Moreover, the findings show that the transmission of emotional spillovers is shaped by macro-financial vulnerability and global risk conditions. On the one hand, economies with higher vulnerability exhibit stronger responses to emotional shocks, indicating that macro-financial vulnerability acts as a moderating mechanism. However, the effects of emotional signals, especially facial and textual emotion, are amplified during periods of higher global uncertainty. Our main findings remain robust across several robustness checks. In particular, for vocal and facial emotions, coefficient equality is not rejected in 85.8% and 78.3% of the market horizon comparisons, respectively, indicating that the main findings are not systematically driven by the choice of emotion extraction algorithms.

Keywords: Central bank communication, Multimodal signals, Emotional spillover, High-frequency identification, Machine learning, Facial emotion analysis

JEL: E44, E52, E58

Introduction

The Federal Reserve occupies a central position in the global financial system, and its monetary policy decisions and communications are closely monitored by investors, policymakers, and financial institutions worldwide. Beyond formal interest rate adjustments and balance sheet policies, the communication strategies of the Federal Reserve, particularly Federal Open Market Committee (FOMC) meeting and press conference by the Fed Chair, have gradually become critical policy instruments for shaping investor expectations, guiding market beliefs, and stabilizing financial conditions (Blinder et al., 2008; Gürkaynak et al., 2005). This role is particularly important because monetary policy affects households and firms indirectly through broader financial conditions, including long-term interest rates, equity prices, house prices, and exchange rates. Caballero and Simsek (2026) theoretically show that uncertainty about a central bank's desired financial conditions can generate market misunderstandings and amplify non-fundamental asset-price fluctuations, whereas clearer communication can reduce these distortions and improve monetary policy transmission.

In periods when conventional monetary policy instruments, such as adjustments to short-term policy interest rates and conventional open market operations, are limited, central bank communication becomes an especially important policy tool. Over much of the past decade and a half, nominal interest rates in major advanced economies have remained close to the zero lower bound, limiting the scope for further conventional easing (Krishnamurthy et al., 2014; Wu and Xia, 2016; Nasir, 2021). Against this background, policymakers have increasingly relied on communication to guide investor expectations, signal future policy intentions, and stabilize financial markets, which has substantially heightened the role of central bank communication in monetary policy transmission (Campbell et al., 2012; Coibion et al., 2025).

Building on this enhanced role of communication, the Federal Reserve further strengthened its communication framework by introducing post-meeting press conferences in 2011. According to Hansen and McMahon (2016), these press conferences represented a major institutional innovation in monetary policy communication, as they provided a real-time forum for the Fed Chair to explain the rationale behind policy decisions, elaborate on the economic outlook, and respond directly to journalists' questions. Conceptually, this development altered the way markets process policy news, because a press conference does not merely repeat the policy statement but helps clarify the underlying motivations, uncertainties, and forward-looking intentions of policymakers, thereby conveying additional emotional information to investors (Cieslak and Schrimpf, 2019). Since 2019, the Federal Reserve has held a press conference after every FOMC meeting, further institutionalizing this communication channel. As a result, the press conference has become a systematic and anticipated component of the policy announcement process, rather than an occasional supplement. This institutional change implies that a substantial share of market-relevant information may now be transmitted through the Chair's explanations, tone, and non-verbal cues, such as facial and vocal expressions, during the press conference, making it an ideal setting for studying how central bank communication shapes financial market expectations at high frequency (Nakamura and Steinsson, 2018).

While post-FOMC press conferences have been widely recognized as a vital channel for transmitting policy-relevant information, the information conveyed in this setting is not limited to the semantic content of prepared statements (Gürkaynak et al., 2005; Hansen and McMahon, 2016). The interactive nature of press conferences creates a richer communication environment in which policymakers' tone, emphasis, and affective expressions become observable to market participants.

In this context, financial markets do not directly observe policymakers' internal assessment of the economy but instead infer it from observable communication signals. While the textual statement provides structured policy guidance, delivery-based cues may reveal policymakers' confidence, uncertainty, or perceived risks surrounding the outlook (Cieslak and Schrimpf, 2019). Because these elements are difficult to fully encode in carefully drafted wording, emotional expressions can contain incremental information about beliefs and future policy paths (DePaulo et al., 2003; Belin et al., 2017). Consistent with this mechanism, evidence shows that vocal tone or facial expressions in policy communication affect asset prices independently of textual content (Mayew and Venkatachalam, 2012; Gorodnichenko et al., 2023; Curti and Kazinnik, 2023). Therefore, we hypothesize that emotional content constitutes an additional informational layer of monetary policy communication rather than a departure from conventional analysis, and rational investors may update expectations in response to such signals.

To test this hypothesis above, this study exploits the institutional setting of post-FOMC press conferences to investigate how the emotional content of Fed Chair communication may influence financial markets in six important economies, including Eurozone, Japan, UK, Brazil, South Africa, and China at high frequency. Existing research has contributed to understanding the domestic impact of Federal Reserve communication on the U.S. economy (Curti and Kazinnik, 2023). However, relatively little is known about how the emotional content of central bank communication may transmit internationally and influence foreign financial markets. The countries mentioned above exhibit strong financial exposure to U.S. monetary policy and the global dollar cycle, while spanning both advanced markets (Eurozone, Japan, UK) and major emerging markets (Brazil, South Africa, China), which provides the cross-country heterogeneity needed to test whether macro-financial vulnerability influences the transmission of emotional signals. Accordingly, emotional content is defined as multimodal affective signals including textual emotions, vocal acoustic features, and facial expressions. The press conference format, combining prepared statements with unscripted responses during the Q&A session, offers a unique environment in which both verbal and non-verbal emotional cues are likely to emerge. Therefore, this setting is particularly well suited for examining whether emotions conveyed by policymakers carry information beyond formal policy announcements. To achieve this objective, the study extracts textual, vocal, and facial emotional signals from transcripts, audio recordings, and videos of FOMC post-meeting press conferences using machine learning methods. These multimodal measures are aligned at the block level, defined as one paragraph in the introduction statement and one answer by the Fed Chair in Q&A session, and combined with high frequency financial data for several assets. Using narrow event windows, the analysis identifies the immediate market responses to Fed Chair emotional cues, while controlling for policy stance and other macro-financial factors. This design provides evidence on whether emotional communication constitutes a distinct transmission channel of U.S. monetary policy.

The results in this paper underscore the salience of the emotional signals contained in the Fed Chair's textual, vocal, and facial expressions in shaping global financial markets. Both positive facial and textual emotional signals have negative influences on USD/BRL, USD/CNH, USD/JPY, and USD/ZAR exchange rates, while have positive influences on USD/GBP and USD/EUR. In other words, positive facial and textual signals are associated with a broad weakening of the U.S. dollar against the Brazilian real, Chinese yuan, Japanese yen, and South African rand. Positive emotional signals conveyed by the Fed Chair likely reduce perceived uncertainty surrounding the economic outlook and future monetary policy path, thereby improving global risk sentiment. As global risk appetite increases, international investors tend to rebalance portfolios toward riskier and higher-yielding emerging market assets, contributing to appreciation in emerging market currencies. In contrast, equity ETF returns generally increase in response to positive changes in facial and textual emotional signals. To validate the estimated responses to textual, vocal, and facial emotional indices during FOMC press conferences, this paper also conducts a robustness analysis by decomposing the overall voice index into frequency-related, amplitude-related, time-related, and quality-related components. The results show that facial and textual emotions still negatively influence the exchange rates in emerging markets and positively influence the exchange rates in more advanced markets. However, how do emotional signals generate such differential effects across different countries?

To address this question, this study explores two possible mechanisms that may help explain why the cross-country heterogeneity exists, including a moderating mechanism and a state-dependent mechanism. First, our results show that the effects of emotional signals differ systematically between high-VIX and low-VIX environments, where regimes are defined cross-sectionally based on the pre-announcement VIX level relative to its sample median (Cloyne et al., 2023). In such environments, conventional policy signals may become less precise, leading investors to place greater weight on emotional cues, especially facial emotional signals. As a result, emotional signals carry greater informational value when uncertainty is high, reinforcing their role as a complement to traditional policy communication. Second, the transmission of emotional signals is shaped by country-specific conditions. Economies with higher macro-financial vulnerability (Ahmed et al., 2026) exhibit stronger and more persistent responses, while less vulnerable economies display more muted reactions. This suggests that macro-financial vulnerability acts as a moderating mechanism, increasing the sensitivity of domestic markets to external emotional shocks.

To the best of our knowledge, this study is the first to study the cross-border spillover of multimodal emotions by Fed Chair, including textual emotion, vocal emotion, and facial emotion, on six important economies,

including Eurozone, Japan, the United Kingdom, Brazil, South Africa, and China. Specifically, by shifting the focus from domestic markets to cross-border financial reactions, the study sheds light on how emotional signals conveyed by U.S. policymakers may transmit internationally and shape investor behavior in global financial markets. Moreover, the analysis provides new empirical evidence on whether combining different communication channels offers a more informative representation of policy communication than any single modality alone. In addition, the study examines whether the cross-border impact of these emotional signals varies systematically with countries' macro-financial vulnerability. Specifically, by interacting multimodal emotion measures with a country-level vulnerability index, the analysis tests whether markets in more vulnerable economies respond more strongly to emotional cues conveyed by the Fed Chair. In doing so, this study enhances our understanding of how central bank communication operates in practice, and highlights the potential importance of emotional spillover as a novel transmission channel of U.S. monetary policy.

The remainder of the paper is organized as follows. Section reviews the related literature on U.S. monetary policy spillovers and central bank communication. Section describes the data sources and the construction of textual, vocal, and facial emotion indices. Section presents the empirical analysis and main hypotheses. Section discusses the results and possible policy implications, and Section 4 concludes.

Literature review

Spillovers of U.S. Monetary Policy

Numerous studies have focused on the impact of U.S. monetary policy shocks on other countries (Hausman and Wongswan, 2011; Dedola et al., 2017; Miranda-Agrippino and Rey, 2020). For example, Kim (2001) shows that in developed countries with flexible exchange rate regimes, U.S. monetary policy shocks are transmitted internationally, with the global real interest rate serving as a central mechanism shaping trade balance outcomes. Similarly, Bauer and Neely (2014) investigate the international spillover effects of the Federal Reserve's unconventional monetary policy. Their findings show that U.S. large-scale asset purchase announcements significantly affect long-term government bond yields in several developed economies. For emerging markets, both Zehri et al. (2024) and Renzhi and Beirne (2024) examined how U.S. monetary policy spills over into emerging market economies (EMEs), though they emphasized different transmission mechanisms and conditioning factors. Moreover, Zehri et al. (2024) document asymmetric contagion effects from U.S. short-term and long-term interest rate shocks, arguing that EMEs must closely track U.S. policy shifts to ensure financial stability. Similarly, focusing on cross-border spillovers, Renzhi and Beirne (2024) show that a tighter U.S. monetary stance significantly shapes monetary policy transmission in EMEs. However, they further show that the adverse effects can be mitigated by stronger financial development, whereas higher capital account and trade openness tend to intensify exposure to global shocks. Overall, U.S. monetary policy has significant impacts on both developed countries and emerging markets.

A substantial strand of the literature investigates the cross-border spillover effects of U.S. monetary policy and provides strong evidence that such policy actions significantly influence financial conditions (Bruno and Shin, 2015; Georgiadis, 2016). For example, Hoek et al. (2022) use a high-frequency event study approach to decompose U.S. interest rate changes into "growth shocks", reflecting improved growth expectations, and "monetary shocks", reflecting hawkish policy shifts or inflation concerns, and show that only the latter significantly tighten EMEs financial market conditions. Their findings make an important contribution by challenging the conventional view that "higher U.S. policy rates" are uniformly bad news for EMEs. Moreover, Cheng and You (2025) distinguish between U.S. cost-push shocks and natural rate shocks and show that these two types of shocks generate markedly different spillover effects on EMEs. Their results suggest that U.S. interest rate increases driven by cost-push shocks tend to weaken the U.S. dollar and exert relatively limited pressure on EME financial markets. Extending this evidence to a broader international setting, Degasperi et al., 2026 employ an information-robust high-frequency identification strategy and Bayesian VAR models covering 30 advanced and emerging economies. They show that U.S. monetary tightening produces widespread contractionary effects through financial conditions, exchange rates, and a distinct commodity-price channel, while the magnitude and composition of these spillovers vary across recipient economies. More recently, Degasperi et al. (2026) identify a commodity-price channel through which U.S. monetary tightening reduces global output and headline inflation, although its limited effect on core inflation suggests that the transmission

is concentrated in globally traded inputs. By contrast, Ahmed et al. (2026) show that U.S. tightening can raise inflation and interest rates in EMEs with weakly anchored inflation expectations, indicating that cross-border effects depend jointly on the source of the U.S. policy shift and recipient-country vulnerabilities.

Therefore, given the central position of the U.S. in global financial markets, a monetary policy shock or change in policy stance leads to significant movements in other countries' exchange rates and financial markets.

Central Bank Communication and Textual Cues

We build on previous studies analyzing the important influences of central bank communications on the economy and financial market. Central bank communications increasingly embed critical information about inflation expectations, employment rates, economic outlook, policy rate changes, and monetary policy decisions. As financial markets have become more sensitive to such signals, communications have evolved from a supplementary tool into a policy instrument to shape economic expectations and influence financial market behavior.

In this sense, as attention to the importance of central bank communication has increased, many studies document a strong relationship between the content of communications and financial markets (Ahrens et al., 2025; Ho and Karagedikli, 2025; Gómez-Cram and Grotteria, 2022; Hüpper and Kempa, 2023). These studies conduct high-frequency event studies and leverage the Large Language Models (LLMs) or Natural Language Processing (NLP) to explore the influence of the verbal cues of communications on asset price movement, volatility, and trading volume. For example, Chen et al. (2026) examine how much additional linguistic information the subsequently released Minutes provide relative to the Statement for the same FOMC meeting, and how this additional information affects financial markets by using the GPT-4. Moreover, Bertsch et al. (2025) apply LLMs to capture various textual attributes to shed light on how communication reflects monetary policy objectives and financial stability concerns. Cho and Jung (2025) explore the effect of the tone of U.S. central bankers' speeches on inflation expectations. Their study suggests that the tone of central bank communications, including speeches by central bank officials beyond those of the Federal Reserve Bank chair alone, can serve as a potent policy instrument, particularly in shaping expectations during expansions. While most studies analyzing the function and impacts of central bank communication concentrate on FOMC statements, Cho and Rho (2026) highlight the importance of other forms of central bank communications, such as press conferences, congressional testimony, social media, and the Fed Chair's speeches. For example, Gorodnichenko et al. (2025) emphasize the significance of other central bank communications, such as tweets on Twitter. This research employs NLP to explore the use of Fed official accounts on Twitter as a communication tool. They find that when talking about central banking topics, the Fed's tweets attract greater attention from Twitter users and the information extracted from the tweets can provide a real-time, qualitative diagnostic for inflation expectations. Moreover, Manjhi and Shah (2025) reveal that there are more positive signals than negative signals between 1991 and 2024 through the analysis of budget speech sentiments. Their research also finds that polarity always aligns with the lower stock returns, making it an effective tool to influence the financial market.

Although the studies above shed light on the importance of verbal cues in central bank communication, relying solely on text-based analysis has important limitations. In public communication, central bankers tend to employ highly cautious and strategic language, particularly in prepared remarks, which leaves little room for emotional expression (Hansen and McMahon, 2016). At the same time, text-based sentiment measures have difficulty separating emotional signals from policy stances, as positive or negative wording often reflects assessment of economic conditions or policy direction rather than the speakers' emotional state, which may blur the distinction between emotion and information effect (Cieslak and Schrimpf, 2019). Moreover, textual analysis focuses on what is said but provides little insight into how it is delivered and thus cannot capture uncertainty or pressure that may arise in unscripted responses during press conference Q&A sessions. As financial markets increasingly attend to the manner of communication, a text-only approach may provide an incomplete picture of the information conveyed by central bank communication.

Non-verbal Emotional Cues in Central Bank Communication

Non-verbal emotional cues have received growing attention in business and finance research because they can add information that is not fully contained in the transcripts. A key idea is that observable behavior in video or audio can "leak" when the speaker is motivated and under pressure, making delivery-based cues harder to perfectly manage than carefully chosen wording (DePaulo et al., 2003). Consistent with this view, experimental evidence in accounting shows that investors react differently when the same forward-looking message is delivered via video, and they penalize managers who display non-verbal cues of uncertainty relative to identical disclosures in audio or written form (Cade, 2020). Related work in strategy and organizational research further demonstrates that combining facial signals with textual content helps identify distinct communication styles that are informative for real economic outcomes, suggesting that how a message is delivered matters alongside what is said (Qin and Yang, 2019). Therefore, within the broader category of non-verbal communication, vocal and facial cues constitute two primary and observable channels for conveying emotional and cognitive states.

Prior studies on speech communication have demonstrated that speakers' vocal features, such as voice pitch, amplitude, and intensity, possess rich information about speakers' underlying emotional and cognitive conditions, including anxiety, confidence, doubt, and perceived trustworthiness (Belin et al., 2017). Moreover, psychology research shows that individuals are less likely to control their non-verbal expressions than their words, such as tone of voice, body language, and facial expressions (Kahneman, 2011). Thus, when Fed chairs communicate their policy stance in public meetings, both the explicit content and manner of delivery in their emotional tone and non-verbal expressions may convey meaningful information to market investors (Kanelis and Siklos, 2023; Ewertz et al., 2026). For example, the emotions shown by central bankers can offer a form of unscripted information. And markets pay attention to these non-verbal signals because of information gaps. These emotional signals can also shape how markets react and may help explain movements in asset prices. In fact, research in finance shows that investors notice emotions in a speaker's voice and tone. For example, Gorodnichenko et al. (2023) studied the Fed chair's voice during FOMC press conferences and show how vocal emotions influence stock prices. They analyze FOMC press conference audio with a customized deep learning emotion-recognition model, showing that voice tone represents a distinct channel of policy communication that operates independently of textual content. Their findings align with the research of Mayew and Venkatachalam (2012), who argue that when policymakers hold negative information that contradicts public beliefs, the inner pressure often shows up in policymakers' non-verbal behaviors. However, focusing only on vocal cues, such as voice pitch and tone, also has important limitations.

Vocal signals capture only one narrow channel of communication and may miss other meaningful dimensions of policymakers' behavior. At the same time, vocal features can be influenced by many factors unrelated to policy information, such as speaking style, personal temperament, health conditions, or the technical setting of the press conference. Therefore, while vocal analysis provides valuable insights, it should be complemented with other modalities, such as facial expressions, to obtain a more complete understanding of how central bank communication influences financial markets and expectations.

While most of the studies have concentrated on exploring the vocal emotions of Fed chairs, the role and importance of facial expressions warrant research. The existing literature in monetary policy analysis applies the theory of human behavior to analyze facial expression and its influence on price movements and volatility. For example, Mathur et al. (2022) introduce the MONOPOLY dataset and a cross-modal transformer (MPCNet) for predicting volatility and price movements from MPC videos, demonstrating sizeable predictive gains when visual, vocal and verbal cues are jointly modeled. Their finding shows that central bank communication can be a vital mechanism through which the non-verbal emotions of Fed Chair's speeches may influence the financial market and economy. Recent work by Curti and Kazinnik (2023) and Alexopoulos et al. (2024) also finds that the Fed chair's facial expressions have immediate influences on the movement of asset prices, showing the importance of non-verbal cues in understanding how monetary policy is communicated.

Although facial expressions have received limited attention in the analysis of central bank communication, firm-level studies in finance indicate that facial expressions convey valuable information and emphasize the role of facial expressions. For example, Akansu et al. (2017) find that there are strong relationships between

facial expressions and profit margins. They show that when CEOs show signs of negative emotions in media interviews, such as disgust and anger, firms tend to see higher profit margins, stronger sales growth, and better returns on assets. However, when CEOs show signs of positive emotions, such as happiness, profit margins, return on equity, and return on assets usually fall. Moreover, Flam et al. (2020) show that investors respond negatively to CEO's facial expressions of anger during televised interviews, even after controlling for the verbal content of the communication, suggesting that facial emotions can independently influence market reactions. Similarly, Banker et al. (2024) find that CEO's dynamic hemifacial asymmetry, a subtle non-verbal facial cue observable in video disclosures, affects investor trust and is associated with adverse stock market responses. These studies all highlight the critical impact of nonverbal communication by executives, especially facial expression, on firm performance.

Multimodal Emotional Signals and Information Integration

While the previous sections discuss textual, vocal, and facial emotional cues separately, the real-world communication rarely occurs through a single channel. In practice, policymakers convey information simultaneously through multiple modalities, including language, voice, and facial expressions. A growing body of evidence suggests that relying on a single communication modality provides an incomplete representation of how information is conveyed and processed. Consistent with this perspective, Binder et al. (2026) show through randomized survey experiments that the vocal and visual components of central bank communication exert additional effects on consumers' expectations beyond those conveyed by textual content alone, with body language generating the largest incremental effect. Moreover, Martinec and Salway (2005) show that multimodal communication, including verbal via text and non-verbal via voice and video, can be semantically integrated to deliver overall information (Yazdani et al., 2024). Their study shows that textual measures alone fail to capture meaningful variation embedded in non-verbal delivery, while vocal or visual cues in isolation may overlook the informational context provided by language (Grebelsky-Lichtman, 2017). In fact, elaboration and extension as alignment and misalignment can be distinguished by how verbal and non-verbal resources interact semantically (Goldin-Meadow, 2003).

Previous research suggests that the semantic relationship between verbal and non-verbal modes plays a crucial role in how messages are processed. When verbal language and non-verbal cues express similar meanings, the message is easier to understand and can be processed with little mental effort (Kjeldsen and Hess, 2021; Lehmann et al., 2016). In contrast, when the meanings conveyed by the two modes differ, understanding becomes more difficult, as the receiver needs to combine information from both sources to make sense of the message (Lehmann et al., 2016; Kjeldsen and Hess, 2021). Thus, for recipients of multimodal communication, understanding a message generally requires more mental effort when verbal and non-verbal cues convey different meanings than when they are consistent with each other (Mehrabian, 1972; Sweller, 1988). In this context, policy messages are often delivered through multiple channels, including official meetings, press conferences and visual materials. When these channels are not fully aligned, market participants may face greater difficulty in interpreting the intended policy stance and show different expectations to the financial market. These arguments indicate that beyond the mere use of multiple communication channels, misalignment in emotional cues across modalities may substantially affect how policy messages are interpreted by financial market participants.

While central bank communication has been studied extensively from text-level and voice-level using LLMs or NLP and speech emotion analysis, there remain only a handful of studies that explore the impact of facial expression by Fed Chairs and how emotional signals interact across different modalities on the financial markets. Given the importance of policymakers' public communications, overlooking multimodal signal interaction may limit a comprehensive understanding of how monetary policy information is conveyed and interpreted by market participants. More importantly, existing studies on central bank communication typically treats communication channels in isolation, implicitly assuming that their effects are additive rather than interactive. However, emotional misalignment across modalities may amplify cognitive load, increase interpretative uncertainty, and thereby exert a distinct influence on market expectations beyond the effect of any single channel alone.

This gap is especially important in the context of FOMC press conference, where investors not only analyze the transcripts, but also watch the livestream. Failing to account for such cross-modal emotional interactions

may lead to an incomplete understanding of how monetary policy signals are transmitted and interpreted by financial markets. Moreover, the literature on central bank communication remains heavily concentrated on single-country settings, predominantly the U.S., leaving high-frequency cross-border evidence relatively scarce. Consequently, no study has systematically examined how textual, vocal, and facial cues jointly shape high-frequency cross-border financial spillovers, particularly in six countries. This paper studies how three kinds of emotional signals interact with each other and fills this gap in the literature. The current study contributes to the existing literature on Fed Chair communication theoretically and methodologically.

From the theoretical perspective, first, this study advances the theory of monetary policy communication by conceptualizing Fed Chair communication as a multimodal emotional signaling system rather than a purely textual or vocal process. Unlike existing studies that focus on a single channel, by integrating textual, vocal, and facial emotional cues into a unified framework, the analysis shows that non-verbal emotions may constitute an independent and economically meaningful communication channel, rather than mere noise. Second, the study introduces a novel emotional spillover channel into the international transmission of U.S. monetary policy. While existing studies primarily emphasize interest rates, risk-taking, and financial conditions, this research examines whether policymakers' emotional expressions can shape global investors' expectations and risk perceptions, thereby generating cross-border financial spillovers, particularly in important economies, including Eurozone, Japan, the United Kingdom, Brazil, South Africa, and China. Third, this study provides two possible mechanisms for why and when the effects of emotional signals vary across countries, including a moderating mechanism linked to macro-financial vulnerability and a state-dependent mechanism driven by global risk conditions.

From the methodology perspective, first, this study introduces a novel two-step PCA-based (Principal Component Analysis; Custódio et al., 2013) vocal emotion extraction strategy. Instead of relying on individual acoustic measures, the research classifies twelve basic vocal features, including vocal pitch, jitter, pitch variability, intensity, shimmer, intensity variability, pulse, speed rate, voice break, Harmonics-to-noise ratio (HNR), vocal entropy, and autocorrelation, into four theoretical dimensions, including frequency, amplitude, time, and voice quality (Yang et al., 2023; Jiang and Pell, 2017; Pon-Barry and Shieber, 2010). PCA is first applied within each dimension to extract dimension-specific latent factors, and a second PCA step is then used to aggregate these factors into a unified vocal emotion index. This hierarchical structure reduces noise and multicollinearity while preserving economically meaningful heterogeneity across vocal channels. Second, this study considers emotions jointly. Compared with Mathur et al. (2022), who employ a non-linear multimodal fusion of audio, video, and textual signals that prevent attribution of effects to any single feature, the present study first analyzes textual, vocal, and facial emotions separately and then examines their interactions. By explicitly modeling cross-modal interaction terms, this study tests whether emotional cues conveyed through different channels reinforce or offset one another in shaping market reactions, rather than assuming a black-box aggregation. This design allows the analysis to go beyond identifying whether emotions matter, and instead to investigate how different communication channels work together. Moreover, the study further interacts these multimodal emotion measures with a country-level macro-financial vulnerability index to examine whether the strength and direction of such joint effects depend on underlying structural conditions.

Data and measurement

FOMC Post-meeting Press Conference

To enhance transparency and accountability in monetary policy, the Federal Reserve introduced post-meeting press conferences in 2011. Initially, press conferences were held after selected FOMC meetings, while since 2019 a press conference has followed every scheduled FOMC meeting. In these press conferences, the Fed Chair elaborates on the Committee's policy decisions, discusses the economic outlook, and explains the rationale underlying recent monetary policy actions. The press conferences typically take place around 30 minutes after the release of FOMC policy statement and serve as an important mechanism through which policymakers can communicate with financial markets and the public. The timing and structure of events during a typical FOMC post-meeting press conference are illustrated in Figure 1.

The press conference usually begins with the release of the FOMC statement, followed by prepared opening

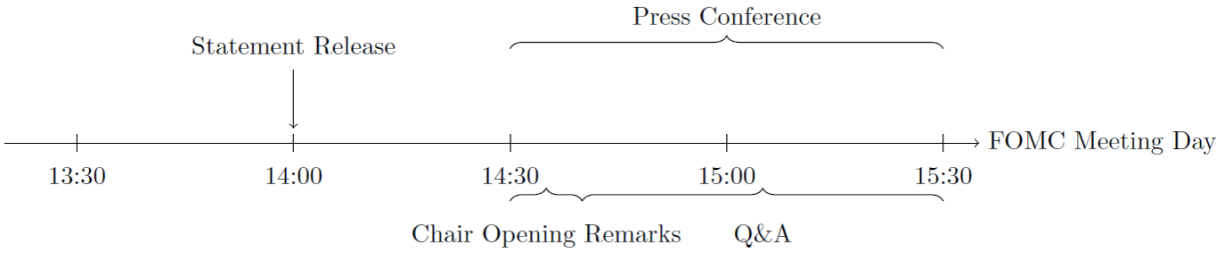


Figure 1. Timeline of the Press Conference.
Alt text: Timeline illustrating the sequence of events on an FOMC meeting day, from the statement release at 14:00 to the post-meeting press conference, including the Chair's opening remarks and the subsequent Q&A session.

remarks conveyed by the Fed Chair. These statements provide an overview of the Committee's assessment of current economic conditions and the policy stance and last about 10 minutes on average. After the Chair's opening statement, the Chair engages in a Q&A session with members of the press. During the Q&A session, journalists pose questions on monetary policy decisions, economic developments, and future direction of interest rates and Fed Chairs need to answer the questions. The Q&A session typically lasts about 44 minutes on average. The whole press conference generally lasts around 55 minutes and concludes with the discretion of the Fed chair.

The dataset of the current research include textual, vocal, and video inputs from FOMC post-meeting press conferences held between April 27, 2011, and December 10, 2025. The sample covers press conferences conducted by three Federal Reserve Chairs Ben Bernanke, Janet Yellen, and Jerome Powell. The transcripts of the press conferences are collected from official Federal Reserve Board website, while the corresponding video recordings are downloaded mainly from the official account of Federal Reserve on YouTube[1]. Although a total of 88 press conferences were originally held during this period, only 61 complete video recordings are available because some videos are missing from YouTube.

Finally, we remove journalists' questions and retain only the Chair's opening statement and answers during the Q&A session. Then, this study generates 2421 blocks totally, with each block consisting of one paragraph in introductory statement or one answer by Fed Chair during Q&A session.

Emotion Indices Using Machine Learning tools

According to the video and audio data from FOMC press conferences sourced from YouTube and textual transcripts collected from the Federal Reserve's official website, this research develops three distinct indices reflecting Fed Chairs' emotions conveyed through their vocal, facial, and textual emotions.

Before constructing the indices, this study starts with the segmentation and alignment of text, video, and audio data. Such accurate segmentation and data alignment (Baltrušaitis et al., 2019) are essential for examining the immediate effect of Fed Chairs' communication on financial markets. Figure 2 illustrates the conceptual framework of our study. Fed Chair communication conveys information through textual, vocal, and facial cues, which are used to construct corresponding emotion indices. These indices then form an emotional spillover channel that affects global financial markets.

Given that FOMC press conference typically lasts approximately one hour, identifying the corresponding time intervals in the audio recording for each sentence in the transcript is quite challenging. To address this problem, this study employs a classical forced alignment algorithm[2] (Sakoe and Chiba, 1978) and applies the alignment method based on the work of Alexopoulos et al. (2024). The study defines a block as a coherent

[1] Federal Reserve System. Official YouTube Channel, providing official speeches, policy briefings, and public communications. https://www.youtube.com/@federalreserve.
[2] We use aeneas Python library to align transcripts and audios. https://github.com/readbeyond/aeneas.

information unit. In the Q&A session, each uninterrupted answer by the Fed Chair to a single question is treated as one block. In the opening remarks, blocks are defined according to the natural paragraphs in the official transcript, with each paragraph representing one block.

The transcripts are first segmented into every sentence, with each block consisting of several sentences. After obtaining sentence-level timestamps based on transcripts, we combine the results across text segments and then use them to split the audio into sentence-level segments. Next, these audio segments are used as inputs to Praat,[3] the primary tool used to extract the vocal features (Boersma and Weenink, 2001). The main vocal features and the process of extraction are shown in Section . Because the audio is extracted directly from the video, there is no separate alignment procedure needed for the video files.

The detailed procedures and methods for constructing text-based, video-based, and audio-based indices are shown in Section , Section , Section respectively.

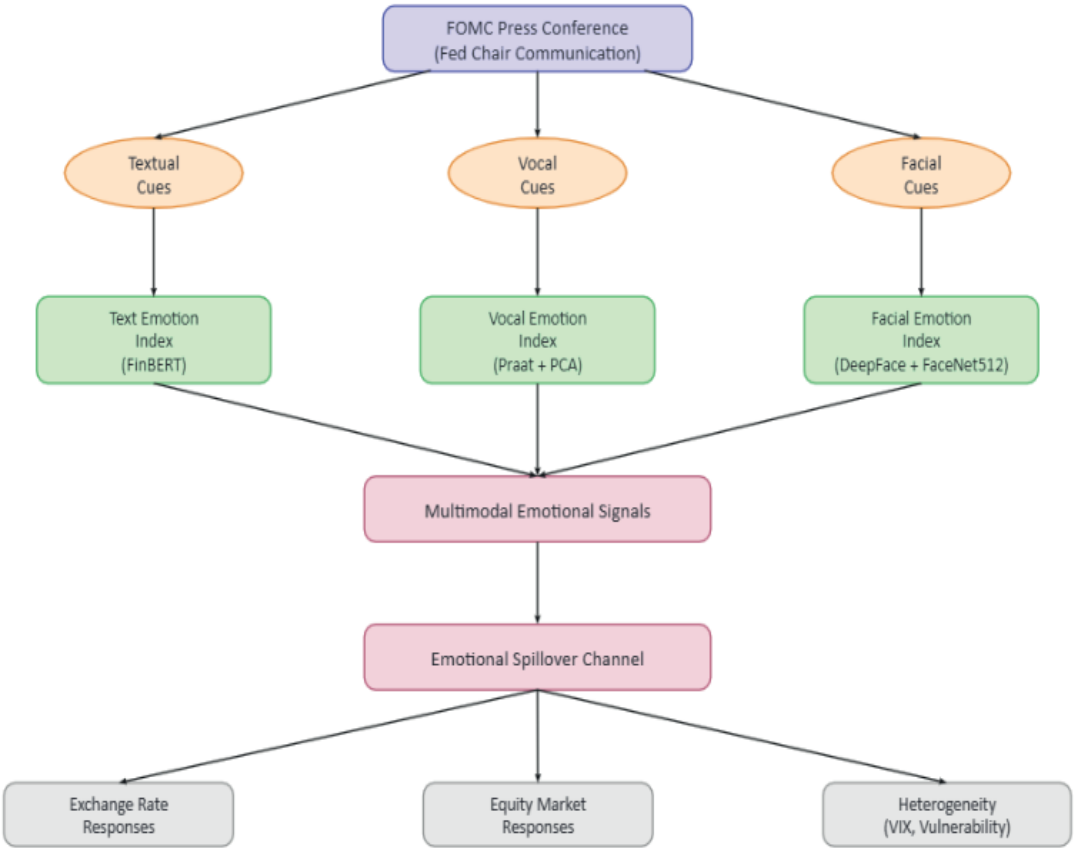


Figure 2. Conceptual Framework of Multimodal Emotional Spillovers from Fed Chair Communication. *Alt text:* Flow diagram illustrating the conceptual framework linking Fed Chair communication to global financial markets. Textual, vocal, and facial cues are used to construct three emotion indices, which form multimodal emotional signals and an emotional spillover channel affecting exchange rates, equity markets, and heterogeneous responses associated with VIX and country vulnerability.

Textual Emotion Index

To quantify the emotional content conveyed through the Fed Chairs' verbal communication, this study constructs a block-level textual emotion index using FinBERT, a domain-specific language model trained

[3] Praat is an open-source software package for phonetic and acoustic analysis. Its source code and documentation are available at https://github.com/praat/praat.github.io.

on financial text (Araci and Genc, 2020) to conduct classification tasks. FinBERT is particularly suitable for central bank communication because it is designed to capture sentiment in finance-related language and justify whether a sentence is positive or negative. Note that before analyzing the FOMC transcripts, this study performs some cleaning steps to remove page numbers, speaker names, timestamps, and other non-textual elements.

The construction of the textual emotion index follows the standard polarity-based approach in the literature, which measures emotional tone as the relative balance between the number of positive and negative sentences (Gorodnichenko et al., 2023). In the sentiment analysis literature, emotional polarity refers to the overall emotional orientation of a text along a negative–positive continuum (Pang and Lee, 2008). Building on this framework, we define emotional polarity at the block level as the difference between positive and negative sentiment scores aggregated across individual sentences. For each sentence spoken by the Fed Chair i in block m of press conference t, FinBERT assigns sentiment scores, taking values 1 (positive), -1 (negative), or 0 (neutral). Textual emotions are first extracted at the sentence level and then aggregated to the block level. The text emotion index for Fed Chair i, block m, in press conference t is shown as:

$$\text{Text}_{i,t,m} = \frac{\text{Positive}_{i,t,m} - \text{Negative}_{i,t,m}}{\text{Positive}_{i,t,m} + \text{Negative}_{i,t,m}}. \tag{1}$$

where $Positive_{i,t,m}$ and $Negative_{i,t,m}$ denote the numbers of sentences classified by FinBERT as positive and negative, respectively, within block m of meeting t by Fed Chair i. As a result, $Text_{i,t,m}$ ranges from -1 to 1, where higher values indicate a more positive emotional tone, and lower values indicate a more negative emotional tone.

Moreover, to separate emotional signals from the policy content itself, this study includes a policy stance index (Gorodnichenko et al., 2023) as a control variable. This index captures whether the communication in each block reflects a hawkish or dovish policy stance. Each sentence is classified into two categories: hawkish or dovish using a fine-tuned uncased BERT model. The policy stance index used in this research for Fed Chair i, block m, in press conference t is shown as:

$$\text{PolicyStance}_{i,t,m} = \frac{\text{Dovish}_{i,t,m} - \text{Hawkish}_{i,t,m}}{\text{Sentence}_{i,t,m}} \tag{2}$$

where $Dovish_{i,t,m}$ and $Hawkish_{i,t,m}$ represent the number of sentences with dovish and hawkish meaning, and $Sentence_{i,t,m}$ represent the number of sentences in the block m in press conference t by Fed Chair i.

Vocal Emotion Index

Because the responses in the Q&A session by Fed Chair are largely unscripted and cannot be fully prepared in advance, Fed Chairs are more likely to unintentionally reveal additional information through their vocal behaviors, reflecting underlying cognitive states and emotional conditions that are not explicitly stated in the verbal text (Gorodnichenko et al., 2023). Therefore, when Fed chairs convey the economic information to the public during the press conferences, especially in the Q&A session, implicit information may be captured by the vocal expressions.

Existing studies on the relationship between vocal expressions and emotions can be broadly grouped into two streams. One stream emphasizes the basic vocal cues, e.g., voice pitch, amplitude, fluency, and quality (Scherer, 2003; Laukka et al., 2008), whereas the other examines affective states that represent the high-level vocal constructs, e.g., happiness, sadness, anger, and fear (El Ayadi et al., 2011). According to the research of Yang et al. (2023), the studies using high-level vocal constructs may suffer from high measurement errors due to the vocal complexity challenge. Therefore, following the research of Alexopoulos et al. (2024), we extract basic vocal cues at 0.015s intervals to analyze the Fed chairs' vocal emotions. Building on the framework of Yang et al. (2023), which employs nine basic vocal features to predict financial risk, this study extends their approach by incorporating three additional cues, including pitch variability, intensity variability, and vocal

entropy (Hildebrand et al., 2020; Kamiloğlu et al., 2020), and organizing a total of twelve vocal features into four distinct dimensions. The full set of features is reported in Table 1.

To aggregate the information contained in these multiple and potentially correlated vocal features, this research novelly employs two-step PCA. Given differences in speaking habits across speakers, 12 vocal features shown in Table 1 are demeaned and standardized (Bitouk et al., 2010) at the speaker level before the implementation of PCA to mitigate differences in speaking styles across distinct Fed Chairs and keep changes in vocal features over time for the same Fed Chair.

The demeaned and standardized value of the 12 vocal features for speaker i in block m during FOMC press conference t is defined as follows:

$$\text{VoiceFeatures}_{i,t,m}^{(d,j)} = \frac{\text{VFeatures}_{i,t,m}^{(d,j)} - \overline{\text{VFeatures}}_{i}^{(d,j)}}{\text{SD}_{i}^{(d,j)}} \tag{3}$$

where $Vfeatures_{i,t,m}^{(d,j)}$ denotes the raw value of the j-th vocal feature belonging to one of the four vocal dimensions d for Fed Chair i in block m during the press conference t, and $\overline{Vfeatures}_{i}^{(d,j)}$ is the speaker-specific mean of 12 features computed across all blocks spoken by Fed Chair i. $Vfeatures_{i,t,m}^{(d,j)} - \overline{Vfeatures}_{i}^{(d,j)}$ are voice scores de-meaned by individual Fed Chair to remove differences in Fed Chairs' average vocal cues. The result $VoiceFeatures$ denotes the observed values of the vocal features for Fed Chair i in block m during press conference t.

Step 1: Principal component analysis within Vocal Dimensions

This research starts with the use of PCA separately within each vocal dimension (Frequency-related, Amplitude-related, Time-related, and Quality-related) and extract the first principal component as the dimension-specific voice index:

$$\text{PCA1}_{i,t,m}^{(d)} = w_1^{(d)}\text{VoiceFeatures}_{i,t,m}^{(d,1)} + w_2^{(d)}\text{VoiceFeatures}_{i,t,m}^{(d,2)} + w_3^{(d)}\text{VoiceFeatures}_{i,t,m}^{(d,3)} \tag{4}$$

where the weight vector $w_j^d = \left(w_1^d, w_2^d, w_3^d\right)$ is selected to maximize the variance of $PCA1_{i,t,m}^{d}$, which corresponds to the direction capturing the greatest common variation in the data.

The results from step 1 are used to conduct heterogeneity analysis in Section .

Step 2: Principal component analysis across Vocal Dimensions

Then, this study aggregates information across the four vocal dimensions by applying PCA to the dimension-level indices obtained from Step 1. Specifically, for each speaker i and block m, the research collects the first principal components from each vocal dimension into a vector:

$$\mathbf{PCA}_{i,t,m} = \begin{pmatrix} \text{PCA1}_{i,t,m}^{(1)} \\ \text{PCA1}_{i,t,m}^{(2)} \\ \text{PCA1}_{i,t,m}^{(3)} \\ \text{PCA1}_{i,t,m}^{(4)} \end{pmatrix} \tag{5}$$

Finally, this study applies PCA to this vector and retains the first principal component as the overall voice emotion index. The final voice index can be shown as follows:

$$\text{Voice}_{i,t,m} = \alpha_1\text{PCA1}_{i,t,m}^{(1)} + \alpha_2\text{PCA1}_{i,t,m}^{(2)} + \alpha_3\text{PCA1}_{i,t,m}^{(3)} + \alpha_4\text{PCA1}_{i,t,m}^{(4)} \tag{6}$$

where the weight vector $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ is selected to maximize the variance of $Voice_{i,t,m}$.

The use of PCA to construct vocal indices offers several advantages. First, vocal features can be highly correlated. For example, high arousal emotions, such as happiness and anger, are associated with higher vocal pitch, increased pitch variability, and higher jitter. Therefore, these features may move together and reflect similar underlying vocal emotion states. PCA can help to reduce dimensionality by extracting the most important common variation among correlated vocal features, while retaining the important information contained in the data. Second, the study uses two-step PCA approach, which provides a flexible framework for further analysis. Step 1 allows us to explore heterogeneity across four distinct vocal dimensions, which is important because different dimensions may convey different types of information.

Table 1. Vocal Feature Dimensions and Definitions

Vocal feature dimensions	Vocal features	Definition
Frequency-related	Vocal pitch	The frequency of a sound wave, measured in hertz (Hz).
	Jitter	The frequency perturbation or instability of a sound wave.
	Pitch variability	Variation in the frequency of a sound wave.
Amplitude-related	Intensity	The amount of power carried by a sound wave.
	Shimmer	The amplitude perturbation or instability of a sound wave.
	Intensity variability	Variation in the amplitude of a sound wave.
Time-related	Pulse	The periodic pulse pattern of a sound wave.
	Speech rate	The number of words spoken per minute.
	Voice breaks	The number of voice breaks or unvoiced frames in a sound wave.
Quality-related	Harmonics-to-noise ratio (HNR)	The ratio of periodic signal energy to noise.
	Vocal entropy	The degree of irregularity and uncertainty in a speech signal.
	Autocorrelation	The extent to which a sound wave resembles white noise or a periodic signal.

Notes: Feature classifications and definitions are based on prior studies of vocal emotion and speech analysis, including Scherer (2003) and Jiang and Pell (2017).
Alt text: Table listing twelve vocal features grouped into four dimensions—frequency, amplitude, time, and voice quality—with a brief definition of each acoustic feature used in the construction of the vocal emotion index.

Facial Emotion Index

Using the FOMC press conference videos described in Section , we construct a block-level facial emotion index based on the Fed Chair's facial expressions.

To ensure that the emotional signals extracted from different data sources correspond to the same communication block, this study first aligns the textual transcripts and the audios recordings by using the aforementioned forced-alignment algorithm (Pettarin, 2017) in Section . Each block corresponds to either one paragraph in the opening statement or a complete answer given by the Fed Chair during the Q&A session. For each block, the corresponding video segment is identified based on the block-level timestamps obtained through the aeneas Python library. This alignment procedure allows the video frames associated with each block to be accurately extracted. Then the videos are segmented by 5 fps and then processed frame by frame to identify the face of the Fed Chairs. Since press conference videos often include multiple participants such as journalists, Fed Chairs, and audience members, an additional filtering procedure is required to ensure that only the Chair's facial expressions are analyzed.

This study first applies a face detection algorithm DeepFace [4] developed by Serengil and Ozpinar (2024), to each video frame to locate facial bounding boxes. Frames without detectable faces are excluded from the analysis.

[4] https://github.com/serengil/deepface.

Then, to distinguish the Chair from other individuals appearing on screen, a face recognition procedure is implemented. For each Chair, including Bernanke, Yellen, and Powell, a small set of reference images is manually selected. These reference images are used to generate facial embeddings, which are numerical vector representations summarizing the key facial features of an individual's face, using the FaceNet512 model[5]. This study then calculates the cosine distance between its embedding and the embeddings of the reference images for each detected face in a frame. A frame is classified as belonging to the Fed Chair if the similarity distance is below a predefined threshold[6]. This procedure ensures that only the frames containing the Chair's face are retained for subsequent emotion analysis.

Next, DeepFace returns emotion scores for the seven facial emotions, including anger, disgust, fear, happiness, neutral, sadness, and surprise. Because disgust is less likely to appear in central bank communication, we exclude it from the index. The current research classifies anger, fear, and sadness as negative emotions, and classifies happiness and surprise as positive emotions. After the facial emotion scores are obtained for frames, they are subsequently averaged to form block-level indicators. Consistent with the construction of the textual emotion index above, the positive and negative facial emotions are then aggregated into a net polarity measure. The final facial emotion index is shown as follows:

$$\text{Face}_{i,t,m} = \frac{\text{Positive}_{i,t,m} - \text{Negative}_{i,t,m}}{\text{Positive}_{i,t,m} + \text{Negative}_{i,t,m}} \tag{7}$$

where $Positive_{i,t,m}$ denotes the average intensity of positive facial emotions (happiness and surprise) expressed during block m in press conference t by Fed Chair i, with $Negative_{i,t,m}$ defined vice versa. The index ranges from -1 to 1, with values closer to 1 indicating a more positive facial emotion and values closer to -1 indicating a more negative facial emotion.

As shown in Figure 3, the facial emotion index remains predominantly below zero, suggesting that Fed Chairs generally display a restrained facial emotional tone. The most pronounced negative observations occur before 2017, broadly coinciding with the period surrounding the normalization of U.S. monetary policy. In December 2015, the Federal Reserve raised the federal funds rate for the first time since 2006, supported by continued labor market improvement and a decline in the unemployment rate to approximately 5%. Although the subsequent tightening path was gradual and extensively communicated, this period was also characterized by uncertainty surrounding the timing of policy normalization, China's economic slowdown, falling oil prices, and global financial market turbulence.

Further negative spikes appear around 2019–2020, 2021–2022, and 2024, coinciding respectively with the COVID-19 shock, rising inflation and rapid monetary policy tightening, and uncertainty about the timing of policy easing. Nevertheless, the five-meeting moving average generally becomes less negative over the sample period, reaching its highest level around 2023–2024 before declining moderately toward 2025.

Market Responses in six economies

Nakamura and Steinsson (2018) construct an identification method that significantly mitigates endogeneity and omitted variable concerns by exploiting high-frequency asset prices and exchange rates changes in narrow windows. The use of narrow event windows limits the influence of macroeconomic news during FOMC press conferences. This approach ensures that observed movements in asset prices and exchange rates reflect responses to policy actions rather than endogenous policy reactions to financial market conditions and exchange rates in China, Eurozone, Japan, Brazil, South Africa, and the United Kingdom.

Considering that FOMC press conferences are typically held based on U.S. Eastern Time, while stock markets across countries and regions operate in different time zones with non-synchronous trading hours, the immediate responses of international equity markets to U.S. monetary policy announcements may be subject to timing

[5] FaceNet512 is a pre-trained deep neural network model for face recognition that maps facial images into high-dimensional embedding vectors, allowing similarity comparisons across faces based on vector distance.

[6] We use a cosine-distance threshold of 0.30 for face identity matching. Frames with embedding distances below 0.30 are classified as belonging to the Fed Chair.

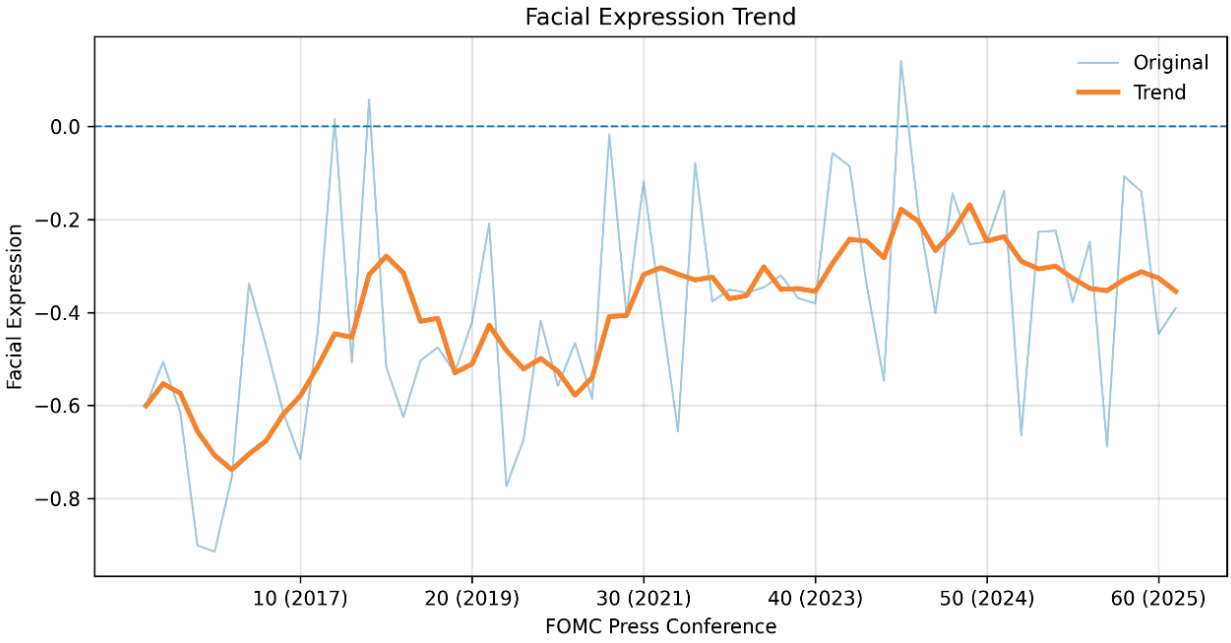


Figure 3. Meeting-Level Facial Expression Index and Its Five-Meeting Moving Average.

Notes: The figure shows the meeting-level facial expression index over the sample period. The horizontal axis reports the chronological sequence number of each FOMC press conference, with the corresponding calendar year shown in parentheses. The light-blue line represents the original meeting-level facial expression index, while the orange line shows its five-meeting moving average, which illustrates the underlying trend in the Fed Chair's facial emotional tone.

Alt text: Line chart showing the Fed Chair's meeting-level facial expression index and its five-meeting moving average across FOMC press conferences from 2011 to 2025. The index is predominantly negative, while the moving average generally becomes less negative over much of the sample period.

Table 2. Examples of Fed Chairs' Facial Emotion Indices

Ben Bernanke, April 25 2012	Janet Yellen, March 18 2015	Jerome Powell, Sept 18 2024
(a)	(b)	(c)
		
Facial index: -0.88497184	Facial index: -0.929025527	Facial index: -0.900516077
Main emotion: sad	Main emotion: fear	Main emotion: angry

Notes: The table presents examples of the facial emotion indices of three Fed Chairs during the press conferences held on 25 April 2012, 18 March 2015, and 18 September 2024. Facial emotion scores are generated using DeepFace.
Alt text: Table presenting selected examples of Fed Chair facial emotion measures from three FOMC press conferences in 2012, 2015, and 2024, with facial emotion scores generated using DeepFace.

mismatches. To mitigate potential measurement biases arising from non-synchronous trading and to enhance the comparability of cross-country market reactions, this study employs 6 U.S.-listed Exchange Traded Funds (ETFs) that are traded at the New York Stock Exchange (NYSE) and track the equity market performance of 6 different countries as proxy variables for national stock markets. These country ETFs primarily track benchmark indices composed of the largest listed firms in each country, typically covering approximately 60%–85% of total domestic market capitalization. Moreover, since ETFs, unlike indices, can be traded, such information has practical meanings for the market participants. As these ETFs are traded during U.S. market hours, their prices can promptly incorporate information released during FOMC press conferences, thereby providing a more accurate measure of short-term equity market responses to FOMC policy shocks. The details of these variables are listed in the following Table 3.

Table 3. Variable Definitions

Variable	**Definition**
Dependent variables	
ETFs (one-minute frequency)	Each ETF serves as a high-frequency proxy for the corresponding economy's equity market.
EZU (iShares MSCI Eurozone ETF)	Return on the iShares MSCI Eurozone ETF, used as a proxy for the euro-area equity market.
EWJ (iShares MSCI Japan ETF)	Return on the iShares MSCI Japan ETF, used as a proxy for the Japanese equity market.
EWU (iShares MSCI United Kingdom ETF)	Return on the iShares MSCI UK ETF, used as a proxy for the UK equity market.
EWZ (iShares MSCI Brazil ETF)	Return on the iShares MSCI Brazil ETF, used as a proxy for the Brazilian equity market.
EZA (iShares MSCI South Africa ETF)	Return on the iShares MSCI South Africa ETF, used as a proxy for the South African equity market.
FXI (iShares China Large-Cap ETF)	Return on the iShares China Large-Cap ETF, used as a proxy for the Chinese equity market.
Exchange rates	Each exchange rate is measured at one-minute frequency and expressed in local currency units per U.S. dollar.
USD/EUR	Change in the Euro per U.S. dollar exchange rate.
USD/JPY	Change in the Japanese yen per U.S. dollar exchange rate.
USD/GBP	Change in the British pound per U.S. dollar exchange rate.
USD/BRL	Change in the Brazilian real per U.S. dollar exchange rate.
USD/ZAR	Change in the South African rand per U.S. dollar exchange rate.
USD/CNH	Change in the offshore Chinese yuan per U.S. dollar exchange rate.
Independent variables	
$\text{Text}_{t,m}$	Standardized block-level textual emotion index for block m in FOMC press conference t, constructed by aggregating sentence-level FinBERT sentiment scores.
$\text{Voice}_{t,m}$	Standardized block-level vocal emotion index for block m in FOMC press conference t, constructed from twelve acoustic features extracted using Praat and aggregated through a two-step principal component analysis.
$\text{Face}_{t,m}$	Standardized block-level facial emotion index for block m in FOMC press conference t, constructed from frame-level facial emotion scores generated using DeepFace.
Control variables	
Policy stance	Block-level index measuring the hawkish or dovish direction of the communication, constructed using a fine-tuned BERT-based classification model.

Variable	Definition
30-day federal funds futures	Change in the current-month 30-day federal funds futures rate around the FOMC communication window, used to capture market-perceived monetary policy surprises.
U.S. 2-year Treasury yield	Change in the U.S. 2-year Treasury yield, used to capture changes in short-term interest-rate expectations.
U.S. 10-year Treasury yield	Change in the U.S. 10-year Treasury yield, used to capture changes in longer-term interest-rate expectations.
MPU	The Monetary Policy Uncertainty index, measuring uncertainty surrounding U.S. monetary policy.
S&P 500 E-mini futures	Return on S&P 500 E-mini futures, used as a real-time measure of U.S. equity market conditions during the FOMC communication window.

Notes: The table presents the definitions of the dependent, independent, and control variables. After their construction at the block level, the textual, vocal, and facial emotion indices are standardized separately to have a mean of zero and a standard deviation of one over the estimation sample. This common scaling facilitates comparisons of coefficient magnitudes across the three modalities. The final standardization of the vocal emotion index is distinct from the speaker-level normalization of the underlying acoustic features conducted before PCA. The latter removes stable differences in vocal characteristics across Fed Chairs.

Alt text: Table defining the dependent, independent, and control variables used in the analysis, including high-frequency ETF and exchange-rate returns, textual, vocal, and facial emotion indices, monetary policy controls, Treasury yields, uncertainty measures, and U.S. equity-market conditions.

Control variables

To separate the effects of emotional communication from conventional policy signals and the broader information environment, the baseline specification controls for the block-level policy stance and the high-frequency change in current-month 30-day federal funds futures, the latter capturing market-perceived monetary policy surprises. It also includes the Monetary Policy Uncertainty (MPU) index developed by Baker et al. (2016), which captures broader policy-related uncertainty that may influence financial market responses around FOMC press conferences.[7]

Moreover, to account for prevailing financial market conditions, the analysis includes several high-frequency financial indicators. Specifically, the U.S. 2-year Treasury yield is used to capture short-term interest rate expectation, while the U.S. 10-year Treasury yield reflects longer-term expectation about economic outlook and inflation, thereby capturing movements along the yield curve. To further control for real-time equity market expectations, the model incorporates the return on S&P 500 E-mini Futures, which provides a forward-looking and high-frequency measure of U.S. stock market sentiment during the FOMC communication window.

Empirical analysis

Market responses in six economies

This section examines the high-frequency relationship between the Fed Chair's textual, vocal, and facial emotional cues and financial market responses across six economies. Using a local projection framework, the analysis estimates the dynamic responses of equity and exchange rate returns following each communication block m.

The empirical specification is given by:

$$\begin{aligned}\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m} = \beta_T^h \text{Text}_{t,m} + \beta_V^h \text{Voice}_{t,m} + \beta_F^h \text{Face}_{t,m} \\ + \theta^{h\prime} \mathbf{X}_{t,m} + \text{FE}_t + \varepsilon_{t,m,h},\end{aligned} \tag{8}$$

[7] The MPU index is available at https://www.policyuncertainty.com/.

where $\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m}$ denotes the h-minute cumulative change in financial outcome for economy e following the end of communication block m during FOMC press conference t. The variables $\text{Text}_{t,m}$, $\text{Voice}_{t,m}$, and $\text{Face}_{t,m}$ denote the textual, vocal, and facial emotion indices extracted from the same communication block, respectively. The vector $\mathbf{X}_{t,m}$ contains the control variables, including the policy stance index, changes in 30-day federal funds futures, U.S. 2-year and 10-year Treasury yields, the Monetary Policy Uncertainty (MPU) index, and S&P 500 E-mini futures returns. The term FE_t denotes FOMC press conference fixed effects, while $\varepsilon_{t,m,h}$ is the error term. The coefficients β_T^h, β_V^h, and β_F^h capture the estimated responses associated with textual, vocal, and facial emotional cues, respectively, at horizon h.

H1: Fed Chair emotional cues, including textual, vocal, and facial cues, are associated with immediate cross-border financial market responses across six economies at high frequency.

Multimodal Interaction effects

This section examines whether the relationship between one emotional modality and financial market responses depends on the emotional signals conveyed through another modality. Specifically, the analysis considers whether relatively high vocal or facial emotion indices amplify or attenuate the market associations of textual and vocal emotional cues.

The empirical specification is given by:

$$\begin{aligned}\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m} = {} & \beta_T^h \text{Text}_{t,m} + \beta_V^h \text{Voice}_{t,m} + \beta_F^h \text{Face}_{t,m} \\ & + \gamma_{TV}^h \left(\text{Text}_{t,m} \times \text{HighVoice}_{t,m}\right) \\ & + \gamma_{TF}^h \left(\text{Text}_{t,m} \times \text{HighFace}_{t,m}\right) \\ & + \gamma_{VF}^h \left(\text{Voice}_{t,m} \times \text{HighFace}_{t,m}\right) \\ & + \theta^{h\prime} \mathbf{X}_{t,m} + \text{FE}_t + \varepsilon_{t,m,h}.\end{aligned} \tag{9}$$

where $\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m}$ denotes the h-minute cumulative change in financial outcome for economy e following the end of communication block m during FOMC press conference t. The variables $\text{Text}_{t,m}$, $\text{Voice}_{t,m}$, and $\text{Face}_{t,m}$ denote the standardized textual, vocal, and facial emotion indices extracted from the same communication block, respectively.

The indicator $\text{HighVoice}_{t,m}$ equals one when the vocal emotion index is above its sample median and zero otherwise. Correspondingly, $\text{HighFace}_{t,m}$ equals one when the facial emotion index is above its sample median and zero otherwise. Thus, γ_{TV}^h captures whether the association between textual emotion and the financial outcome differs between blocks with relatively high and low vocal emotion indices. Similarly, γ_{TF}^h captures how the association with textual emotion varies according to the level of the facial emotion index, while γ_{VF}^h captures how the association with vocal emotion varies according to the level of the facial emotion index.

The vector $\mathbf{X}_{t,m}$ contains the same control variables as in Equation 8, including the policy stance index, changes in 30-day federal funds futures, U.S. 2-year and 10-year Treasury yields, the Monetary Policy Uncertainty (MPU) index, and S&P 500 E-mini futures returns. The term FE_t denotes FOMC press conference fixed effects, while $\varepsilon_{t,m,h}$ is the error term.

A positive interaction coefficient indicates that the association between one emotional modality and the financial outcome is stronger when the conditioning modality takes a relatively high value. A negative interaction coefficient indicates that this association is attenuated when the conditioning modality takes a relatively high value. The interaction coefficients therefore test whether the estimated market associations of emotional communication vary across modalities rather than operating independently and additively.

H2: The association between an emotional cue and financial market responses depends on the level of emotional cues conveyed through other communication modalities.

Macro-financial Vulnerability as a Moderating Mechanism

Lastly, this section examines whether the cross-border financial spillovers triggered by Fed Chair emotional cues differ systematically across six economies. While the emotional content of U.S. monetary policy communication may represent a common external shock, the ability of different countries to absorb such shocks depends on their underlying macro-financial conditions. As a result, the emotional signals conveyed by the Fed Chair may generate stronger market reactions in more vulnerable economies.

In order to explore this structural heterogeneity, the study introduces a country-level macro-financial vulnerability ranking based on the methodology conducted by Ahmed et al. (2026). For each year in the sample, these six economies are first ordered according to a set of macro-financial indicators that capture external imbalances and exposure to foreign currency risk. Especially, countries are ranked each year based on the following indicators: current account balance as a percent of GDP, gross government debt as a percent of GDP, average annual inflation over the past three years, the ratio of foreign exchange reserves to GDP. For each indicator, higher values correspond to greater vulnerability. The individual rankings are then averaged across all indicators to obtain a composite country-level vulnerability rank for each year. A higher value of the final index indicates a more macro-financially vulnerable economy.

The baseline local projection model is extended as follows for each horizon h minute:

$$\begin{aligned}\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m} = \beta_T^h \text{Text}_{t,m} + \beta_V^h \text{Voice}_{t,m} + \beta_F^h \text{Face}_{t,m} \\ + \delta_T^h \left(\text{Text}_{t,m} \times \text{VUL}_{e,t-1}\right) \\ + \delta_V^h \left(\text{Voice}_{t,m} \times \text{VUL}_{e,t-1}\right) \\ + \delta_F^h \left(\text{Face}_{t,m} \times \text{VUL}_{e,t-1}\right) \\ + \theta^{h\prime} \mathbf{X}_{t,m} + \text{FE}_t + \varepsilon_{t,m,h}.\end{aligned} \tag{10}$$

where $\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m}$ denotes the h-minute cumulative change in financial outcome for economy e following the end of communication block m during FOMC press conference t. The variables $\text{Text}_{t,m}$, $\text{Voice}_{t,m}$, and $\text{Face}_{t,m}$ denote the standardized textual, vocal, and facial emotion indices extracted from the same communication block, respectively.

The variable $\text{VUL}_{e,t-1}$ denotes the lagged macro-financial vulnerability measure for economy e, with higher values indicating greater vulnerability. The interaction coefficients δ_T^h, δ_V^h, and δ_F^h indicate how the estimated associations between textual, vocal, and facial emotional cues and financial market responses vary with an economy's macro-financial vulnerability.

More specifically, δ_T^h captures how vulnerability modifies the association between textual emotion and the financial outcome at horizon h. Correspondingly, δ_V^h and δ_F^h capture how vulnerability modifies the associations with vocal and facial emotional cues. The total estimated association for each modality therefore depends on the level of vulnerability. For example, the association with textual emotion is given by

$$\beta_T^h + \delta_T^h \text{VUL}_{e,t-1}.$$

A positive interaction coefficient indicates that the estimated association becomes more positive as vulnerability increases, whereas a negative interaction coefficient indicates that it becomes more negative. Whether this represents amplification or attenuation depends on the sign of the corresponding baseline coefficient β_T^h, β_V^h, or β_F^h.

The vector $\mathbf{X}_{t,m}$ contains the same control variables as in Equation 8, including the policy stance index, changes in 30-day federal funds futures, U.S. 2-year and 10-year Treasury yields, the Monetary Policy Uncertainty (MPU) index, and S&P 500 E-mini futures returns. The terms FE_t denotes FOMC press conference fixed effects, while $\varepsilon_{t,m,h}$ is the error term.

H3: The associations between Fed Chair emotional cues, including textual, vocal, and facial cues, and financial market outcomes vary systematically with country-level macro-financial vulnerability.

Global Risk Conditions as a State-Dependent Mechanism

This section examines whether the transmission of emotional signals depends on global risk conditions. While the previous section focuses on cross-country heterogeneity driven by macro-financial vulnerability, this section investigates whether the impact of emotional communication varies over time with changes in market-wide uncertainty.

To capture the role of global risk conditions, the sample is divided into high-VIX and low-VIX regimes according to the pre-announcement level of the VIX index using state-dependent local projection (Cloyne et al., 2023):

$$\text{HighVIX}_t = \begin{cases} 1, & \text{if VIX}_t > \text{median(VIX)}, \\ 0, & \text{if VIX}_t \leq \text{median(VIX)}. \end{cases} \tag{11}$$

where VIX_t denotes the level of the VIX index prior to the FOMC press conference at time t.

The state-dependent local projection model is estimated separately for the low-VIX and high-VIX regimes at each horizon h. For the low-VIX regime, the specification is given by:

$$\begin{aligned} \text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m} = \beta^h_{T,L}\text{Text}_{t,m} + \beta^h_{V,L}\text{Voice}_{t,m} + \beta^h_{F,L}\text{Face}_{t,m} \\ + \theta^{h\prime}_L \mathbf{X}_{t,m} + \text{FE}_t + \varepsilon^L_{t,m,h}. \end{aligned} \tag{12}$$

For the high-VIX regime, the corresponding specification is:

$$\begin{aligned} \text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m} = \beta^h_{T,H}\text{Text}_{t,m} + \beta^h_{V,H}\text{Voice}_{t,m} + \beta^h_{F,H}\text{Face}_{t,m} \\ + \theta^{h\prime}_H \mathbf{X}_{t,m} + \text{FE}_t + \varepsilon^H_{t,m,h}. \end{aligned} \tag{13}$$

In Equations 12 and 13, $\text{Coun}_{e,t,m+h} - \text{Coun}_{e,t,m}$ denotes the h-minute cumulative change in the financial outcome for economy e, measured from the end of communication block m during FOMC press conference t. The variables $\text{Text}_{t,m}$, $\text{Voice}_{t,m}$, and $\text{Face}_{t,m}$ denote the standardized textual, vocal, and facial emotion indices extracted from the same communication block, respectively.

The coefficients $\beta^h_{T,L}$, $\beta^h_{V,L}$, and $\beta^h_{F,L}$ capture the estimated associations with textual, vocal, and facial emotional cues in the low-VIX regime. Correspondingly, $\beta^h_{T,H}$, $\beta^h_{V,H}$, and $\beta^h_{F,H}$ capture these associations in the high-VIX regime. Comparing the corresponding coefficients across the two specifications shows whether the estimated market associations vary with financial market volatility.

The vector $\mathbf{X}_{t,m}$ contains the same control variables as in Equation 8, including the policy stance index, changes in 30-day federal funds futures, U.S. 2-year and 10-year Treasury yields, the Monetary Policy Uncertainty (MPU) index, and S&P 500 E-mini futures returns. The terms FE_t denotes FOMC press conference fixed effects. Finally, $\varepsilon^L_{t,m,h}$ and $\varepsilon^H_{t,m,h}$ denote the error terms in the low-VIX and high-VIX regimes.

H4: The associations between Fed Chair emotional cues, including textual, vocal, and facial cues, and financial market outcomes are stronger in the high-VIX regime than in the low-VIX regime.

Findings and Policy implications

Market responses in six economies

Figure 4 provides the dynamic responses of exchange rates and equity ETFs in six countries to the Fed Chair's textual, vocal, and facial emotional indices during FOMC press conferences. Overall, all these three emotional channels generate high-frequency cross-border market reactions, but the effects differ substantially across different modalities, asset classes, and countries. Among these three channels, facial emotion appears to have the strongest and most stable effects, followed by textual emotion, while vocal emotion remains comparatively weak and less stable. Moreover, this initial evidence suggests that financial markets react not

only to what the Fed Chair says, but also to how the message is delivered and what non-verbal emotional signals accompany it. In particular, facial expressions are more direct and difficult to strategically control than vocal expressions, making them more credible indicators of policy confidence and uncertainty.

It is worth noting that both facial and textual expressions induce sizable movements, and the direction of the responses exhibits significant cross-country heterogeneity in the exchange rate market. For Japan, Brazil, South Africa, and China, positive changes in facial emotion lead to persistent negative responses by roughly 1.2 bps, 4 bps, 3 bps, and 0.8 bps, respectively, whereas the euro and the pound respond positively by about 2 bps at longer horizons. A similar situation can be observed for text emotion, where exchange rates in the Eurozone and the United Kingdom respond positively. While those in Brazil, South Africa, Japan, and China decline by 2 bps, 0.5 bps, 1 bps, and 1.5 bps respectively. In contrast to the exchange rate, where responses from facial emotion and textual emotion exhibit substantial cross-country heterogeneity in both magnitude and direction, the equity ETFs increase in response to positive changes in facial emotion. In particular, for Brazil and South Africa, ETF returns increase by approximately 8 bps and 10 bps at longer horizons, while China, the Euro area, Japan, and the United Kingdom also show consistent positive responses in the range of 3-5 bps. A broadly similar pattern can be observed for textual emotion, although the magnitude is significantly smaller and the dynamics are less smooth. Most economies exhibit positive responses of around 1-2 bps at medium to longer horizons, indicating that more positive textual sentiment is generally associated with higher equity valuations. By contrast, vocal emotion still produces the weakest and least consistent effects in the ETF markets. This is not very surprising. Vocal cues are significantly difficult to disentangle and identify in isolation, which limits their immediate interpretability and responses from market participants. In Section 4, this study provides a heterogeneity analysis across different vocal dimensions that yields statistically significant results.

These results suggest that investors interpret more positive facial and textual signals as indicating lower financial uncertainty or greater policy confidence, which tends to improve global risk appetite and support equity prices across both advanced and emerging markets. However, the exchange rate responses point to a more complicated situation, as the same emotional signals produce markedly different outcomes across countries. In emerging markets such as Brazil, China, and South Africa, more positive facial and textual emotional signals are consistently associated with declines in USD/BRL, USD/CNH, and USD/ZAR exchange rates, indicating an appreciation of local currencies. In contrast, positive emotional signals are followed by increases in USD/EUR and USD/GBP, which indicates an appreciation of the U.S. dollar rather than the local currency.

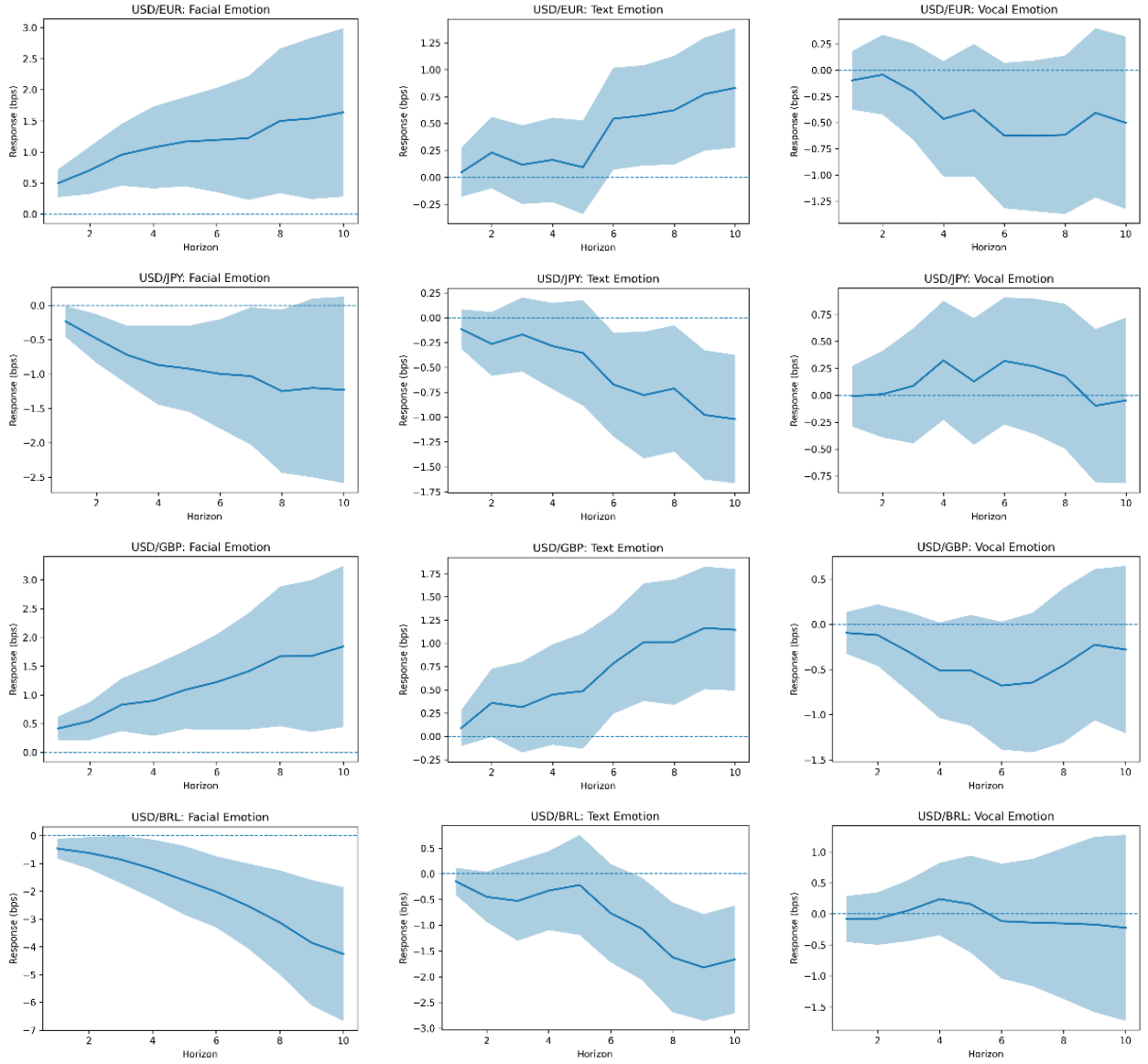
Two possible mechanisms may help explain these cross-country differences. One possible explanation builds on the literature on global capital flows (McCauley, 2012), which shows that international investors are more likely to allocate capital to riskier assets in emerging markets during calmer periods characterized by lower volatility. This suggests that investors' investment preferences may systematically vary across different global risk conditions. When uncertainty is high, market participants may place greater weight on supplementary information cues, including emotional signals, when forming expectations. In particular, when uncertainty is high, conventional policy signals may become less precise, reducing their informational content. Then market participants may be more likely to rely on additional information cues to guide their expectations and decisions, such as emotional signals. As a result, the relative informational value of emotional signals may increase as uncertainty increases. In other words, differences in VIX levels may lead to differences in how investors perceive emotional signals, thereby amplifying or attenuating their underlying investment preferences and then contributing to the observed cross-country heterogeneity in exchange rate and equity market responses. Using the VIX index as a proxy for global risk uncertainty, this paper examines whether the effects of emotional signals vary across different risk environments in Section 4.

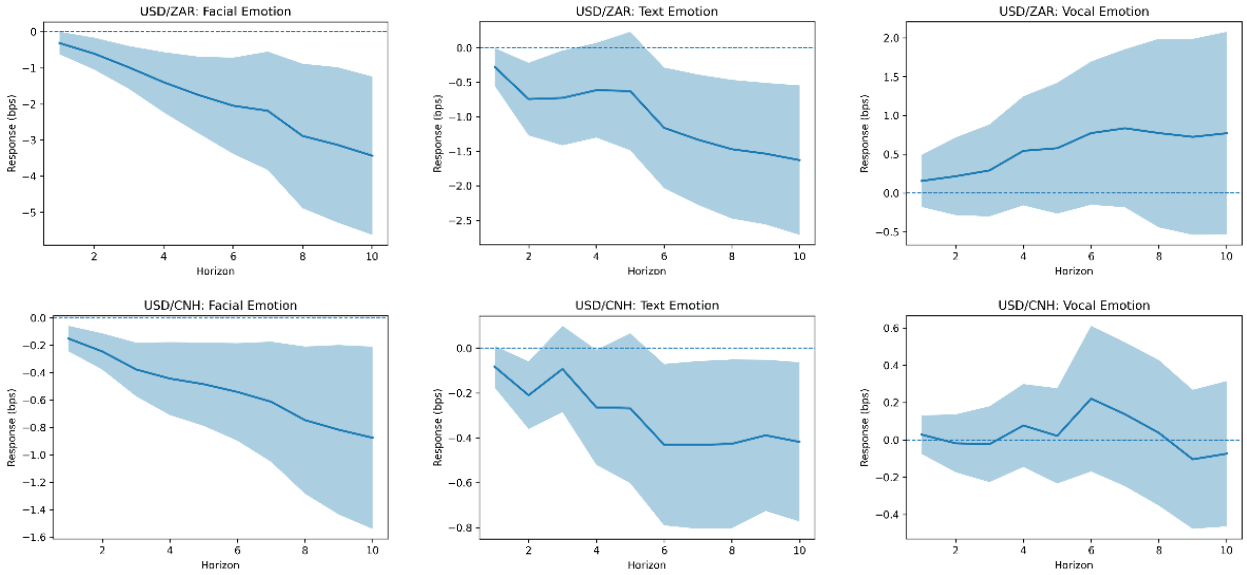
Another possible interpretation is that macro-financial vulnerability also shapes how emotional signals are transmitted across countries. Economies with higher vulnerability may be more sensitive to such external shocks, while less vulnerable economies tend to be less sensitive to external shocks. This mechanism is closely related to the broader literature on international spillovers, which shows that increases in an emerging market economy's macroeconomic and financial vulnerabilities influence the sensitivity of its asset prices to U.S. monetary policy shocks, particularly to interest rate surprises (Hoek et al., 2022). In this sense, macro-financial vulnerability may act as a moderating mechanism of emotional spillovers, which is tested in

Section 4.

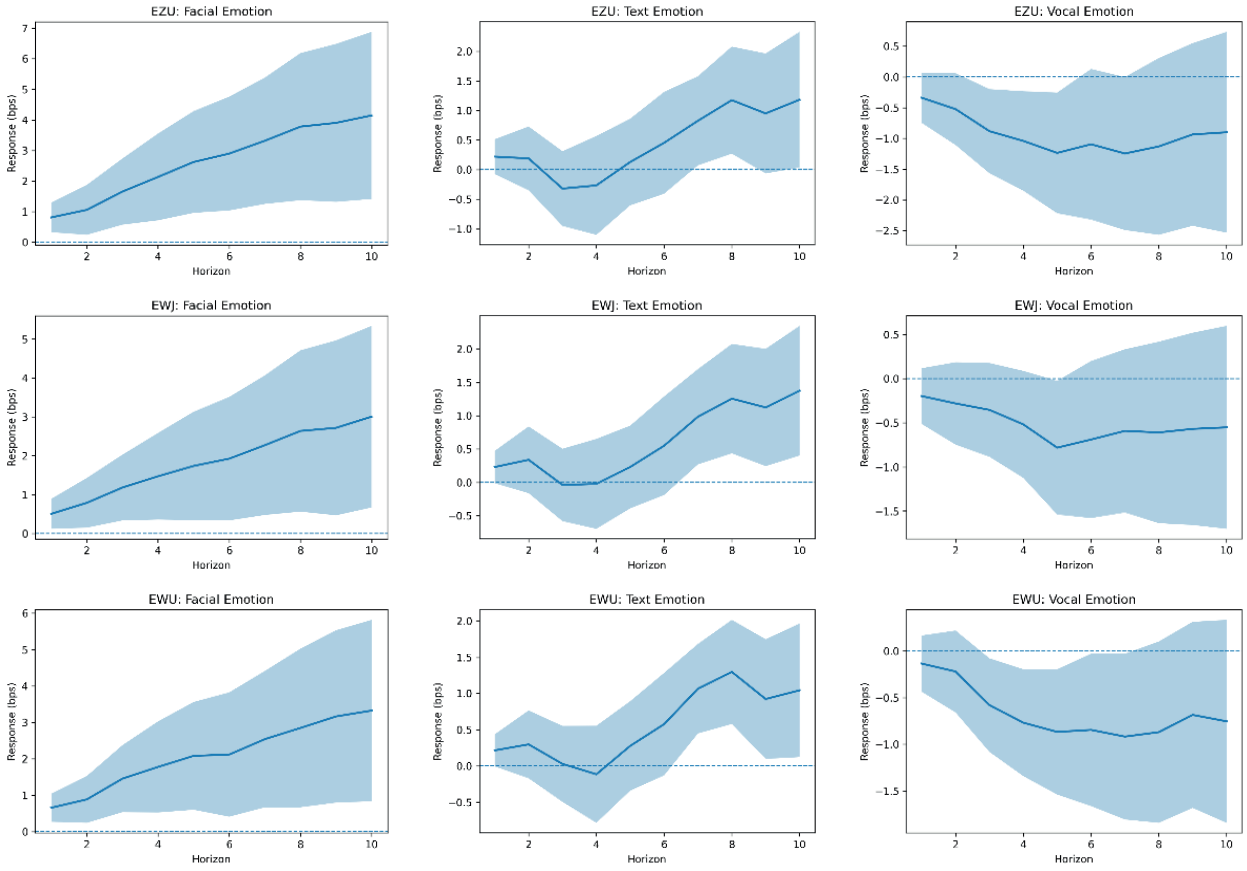
Overall, these findings show that emotional communication by the Fed Chair constitutes a meaningful and valuable source of information in international financial markets. Its effects vary systematically across modalities and countries, reflecting differences in informational content and interpretation and giving rise to pronounced cross-market and cross-country heterogeneity. Importantly, the evidence provides strong support for H1, indicating that emotional shocks have economically significant high-frequency effects on global asset prices. Moreover, this paper further decomposes the vocal emotional index into four components to conduct a robustness analysis in Appendix B. Figures B.2 and B.3 in Appendix B show that the responses to facial and textual emotional signals remain broadly unchanged. Their magnitudes and dynamic movements are broadly consistent with the baseline results, suggesting the effects of textual and facial signals are robust to alternative specifications of vocal emotion. Beyond confirming the relevance of emotional cues, the results contribute to the literature by offering a multimodal perspective on central bank communication and by highlighting the role of the Fed Chair's emotional signals in global financial markets. In the next section this study investigates whether emotional communication is inherently non-additive.

Exchange Rates




USD/ZAR: Facial Emotion
USD/ZAR: Text Emotion
USD/ZAR: Vocal Emotion
USD/CNH: Facial Emotion
USD/CNH: Text Emotion
USD/CNH: Vocal Emotion
Response (bps)
Horizon

ETFs


EZU: Facial Emotion
EZU: Text Emotion
EZU: Vocal Emotion
EWJ: Facial Emotion
EWJ: Text Emotion
EWJ: Vocal Emotion
EWU: Facial Emotion
EWU: Text Emotion
EWU: Vocal Emotion
Response (bps)
Horizon

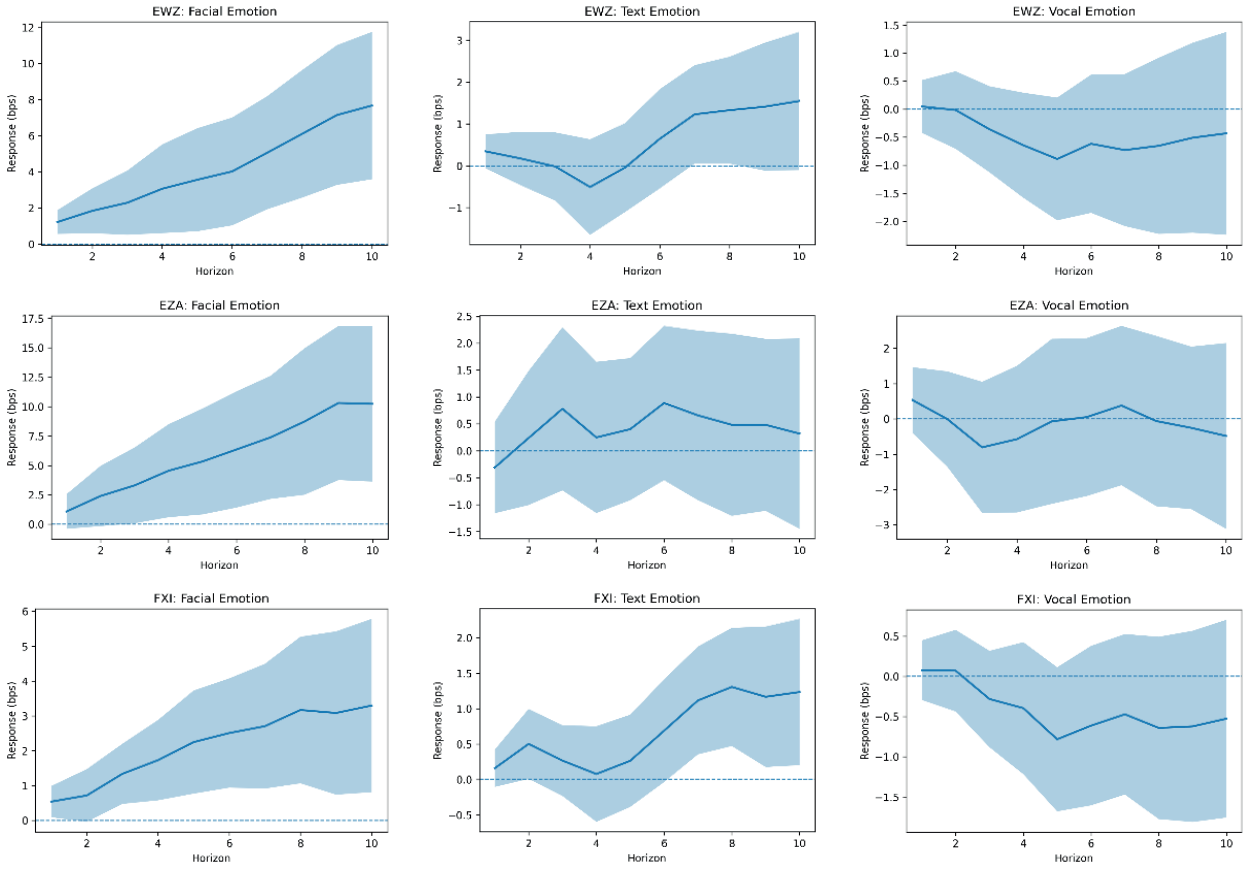


Figure 4. Responses of Exchange Rates and ETFs to Fed Chair Emotional Cues.

Notes: The figure presents the estimated responses of exchange rates and ETFs to the Fed Chair's facial, textual, and vocal emotional cues during FOMC press conferences. The three columns report the responses to facial, textual, and vocal emotional cues, respectively. Within each panel, the economies are ordered as the Eurozone, Japan, the United Kingdom, Brazil, South Africa, and China. The shaded areas represent the 90% confidence intervals.

Alt text: Impulse-response plots showing the estimated responses of exchange rates and equity ETFs in six economies to facial, textual, and vocal emotional cues from the Fed Chair over ten one-minute horizons. Shaded areas represent 90% confidence intervals.

Multimodal interaction effects

Table 4 examines the interaction effects across textual, vocal, and facial emotional cues. The results reveal that the impact of emotional communication is non-additive, where the marginal effect of one modality depends on the intensity of others.

It is worth noting that vocal emotion plays a conditioning role in the transmission of information. Although vocal emotion exhibits relatively weak and unstable effects on its own in Section , it plays a more systematic conditioning role once the interaction terms are introduced. However, once interaction terms are introduced, vocal cues exhibit a much more systematic role in shaping market responses. In particular, vocal emotion significantly influences the impact of textual signals in several exchange rate markets, as reflected in the positive and significant interaction terms for USD/BRL and USD/ZAR. This implies that when vocal emotion is strong, the marginal effect of textual sentiment increases. Moreover, this contrast also suggests that vocal emotion may not operate as an independent source of information, but rather as a complementary channel that modifies how other signals are interpreted by market participants.

In contrast, advanced economy currencies exhibit a different pattern. For USD/EUR and USD/GBP, the interaction between vocal emotion and HighFace is negative and statistically significant at medium to longer horizons. This indicates that when facial emotion is already stronger, the marginal effect of vocal emotion declines. In other words, the additional informational contribution of vocal cues becomes limited once facial signals are sufficiently pronounced. This pattern is consistent with a substitution effect across modalities, suggesting that overlapping signals may reduce the need for further information updating in these markets.
The results in this study also show that the interaction effects display a broadly consistent negative pattern in emerging market ETFs. For Brazil (EWZ), China (FXI), and South Africa (EZA), the interaction between textual emotion and HighVoice is predominantly negative and statistically significant at several horizons, indicating that the marginal contribution of textual emotion tends to weaken when vocal emotion is relatively high. These results indicate that when vocal signals are already strong, the marginal contribution of textual emotion becomes weaker, leading to attenuated price responses. This suggests that the joint presence of strong signals across modalities reduces the incremental informational value of each individual channel.
The findings in this study show that multimodal emotional signals do not uniformly reinforce each other. Instead, their joint effects depend on both the asset class and the market environment. Reinforcement effects are more evident in emerging market exchange rates, where vocal cues amplify textual signals, while attenuation effects dominate in advanced economy currencies and equity markets, where overlapping information reduces marginal impacts. This provides strong support for H2 and highlights that emotional communication functions as an interactive signaling system rather than a simple aggregation of independent channels.
these results suggest that the market relevance of an emotional signal depends not only on the signal itself but also on the emotional information conveyed through other modalities. The effectiveness of communication not only depends on the message itself, but also depends on the interaction and alignment across different channels, which jointly determine its credibility and informational impact.

Table 4. Interaction Effects Across Textual, Vocal, and Facial Emotional Cues

Market	Interaction term	Conditional IRF h=2	h=4	h=6	h=8	h=10	Weighted sum
Exchange rates							
	Text*High-	-0.47	0.01	-0.55	-0.49	-0.41	-3.05
USD/BRL	Face	(0.42)	(0.61)	(0.80)	(0.99)	(1.01)	
	Text*High-	0.90**	0.58	0.80	1.36*	1.73**	10.63
	Voice	(0.39)	(0.48)	(0.67)	(0.73)	(0.83)	
	Voice*High-	0.54	0.90	0.82	2.01	1.69	11.08
	Face	(0.51)	(0.78)	(0.94)	(1.27)	(1.51)	
	Text*High-	0.02	0.05	0.15	-0.06	-0.14	0.29
USD/CNH	Face	(0.14)	(0.13)	(0.17)	(0.24)	(0.26)	
	Text*High-	0.05	0.13	0.03	0.17	0.56*	1.63
	Voice	(0.16)	(0.17)	(0.21)	(0.26)	(0.29)	
	Voice*High-	-0.02	-0.01	-0.03	0.28	0.60**	1.49
	Face	(0.15)	(0.20)	(0.24)	(0.28)	(0.30)	
	Text*High-	0.04	-0.19	-0.31	-0.37	-0.50	-2.86
USD/EUR	Face	(0.28)	(0.33)	(0.46)	(0.55)	(0.60)	
	Text*High-	-0.21	0.01	0.10	-0.17	-0.56	-1.23
	Voice	(0.22)	(0.35)	(0.45)	(0.53)	(0.56)	
	Voice*High-	-0.46	-0.77	-1.10**	-1.39**	-1.57**	-9.96
	Face	(0.43)	(0.50)	(0.55)	(0.65)	(0.74)	
	Text*High-	0.11	0.46	0.58	0.33	0.34	3.73
USD/JPY	Face	(0.26)	(0.35)	(0.48)	(0.54)	(0.64)	
	Text*High-	0.21	0.10	-0.33	-0.30	0.41	-0.02
	Voice	(0.30)	(0.35)	(0.49)	(0.54)	(0.55)	
	Voice*High-	0.29	0.56	0.67	1.20*	1.51**	7.74
	Face	(0.31)	(0.40)	(0.52)	(0.69)	(0.64)	

	Text*High-	0.02	-0.04	-0.24	-0.15	-0.11	-1.32
USD/GBP	Face	(0.26)	(0.30)	(0.38)	(0.50)	(0.58)	
	Text*High-	0.09	0.24	0.30	-0.14	-0.66	-0.07
	Voice	(0.35)	(0.44)	(0.47)	(0.52)	(0.55)	
	Voice*High-	-0.20	-0.62	-1.34**	-1.46**	-1.53**	-8.95
	Face	(0.40)	(0.42)	(0.58)	(0.63)	(0.68)	
	Text*High-	-0.25	0.56	0.82	0.48	1.13	5.21
USD/ZAR	Face	(0.45)	(0.58)	(0.66)	(0.94)	(1.01)	
	Text*High-	0.92**	0.55	0.65	0.56	1.16	7.13
	Voice	(0.46)	(0.51)	(0.71)	(0.80)	(0.93)	
	Voice*High-	0.02	0.54	0.62	1.40	1.20	8.21
	Face	(0.44)	(0.73)	(0.98)	(1.11)	(1.23)	
	Text*High-	0.59	0.92	2.01	2.07	2.23	13.85
EWZ	Face	(0.77)	(0.98)	(1.44)	(1.95)	(2.04)	
	Text*High-	-1.00	-0.86	-2.03	-3.56**	-4.20**	-22.82
	Voice	(0.80)	(0.89)	(1.26)	(1.44)	(1.70)	
	Voice*High-	-0.04	-1.20	-1.17	-2.63	-2.22	-13.69
	Face	(0.92)	(1.35)	(1.50)	(1.93)	(2.05)	
	Text*High-	0.59	0.95	0.62	0.95	0.49	6.16
FXI	Face	(0.58)	(0.82)	(0.94)	(1.09)	(1.21)	
	Text*High-	-1.35***	-0.84	-1.23	-2.46**	-3.07**	-17.28
	Voice	(0.52)	(0.71)	(1.04)	(1.11)	(1.24)	
	Voice*High-	-0.43	-0.98	-0.76	-1.53	-1.14	-8.9
	Face	(0.49)	(0.73)	(0.83)	(1.10)	(1.29)	
	Text*High-	0.34	0.58	0.30	0.55	0.18	3.14
EZU	Face	(0.62)	(0.71)	(0.95)	(1.17)	(1.36)	
	Text*High-	-0.56	-0.51	-1.28	-1.75	-1.44	-12.13
	Voice	(0.67)	(0.78)	(0.96)	(1.07)	(0.98)	
	Voice*High-	0.37	0.09	-0.58	-0.80	-1.45	-13.15
	Face	(0.38)	(0.67)	(0.72)	(0.78)	(0.98)	
	Text*High-	0.34	0.44	0.33	0.47	0.47	2.6
EWJ	Face	(0.50)	(0.61)	(0.81)	(1.00)	(1.16)	
	Text*High-	-0.49	-0.30	-0.45	-0.71	-1.14	-6.27
	Voice	(0.55)	(0.66)	(0.90)	(0.92)	(1.07)	
	Voice*High-	0.51*	-0.32	-0.59	-1.41	-1.48	-6.14
	Face	(0.31)	(0.66)	(0.88)	(1.06)	(1.12)	
	Text*High-	0.13	0.35	0.13	0.37	0.43	2.13
EWU	Face	(0.48)	(0.62)	(0.80)	(0.97)	(1.12)	
	Text*High-	-0.31	-0.01	-0.43	-0.97	-1.59	-6.38
	Voice	(0.58)	(0.73)	(0.86)	(0.93)	(1.04)	
	Voice*High-	-0.13	-1.07	-1.99**	-2.44**	-1.96	-13.83
	Face	(0.57)	(0.71)	(0.99)	(1.05)	(1.22)	
	Text*High-	-2.26	-2.67	-0.07	-0.89	-0.15	-11.93
EZA	Face	(1.77)	(2.12)	(2.92)	(3.06)	(3.22)	
	Text*High-	-3.29***	-3.41**	5.23**	-4.77**	-5.14**	-41.47
	Voice	(1.25)	(1.60)	(2.11)	(2.33)	(2.38)	
	Voice*High-	-0.54	-2.32	-2.44	-2.95	-2.55	-22.04
	Face	(1.29)	(2.23)	(2.64)	(2.99)	(3.39)	

Notes: This table reports the estimated interaction effects across textual, vocal, and facial emotional cues. Standard errors are reported in parentheses. The weighted sum summarizes the interaction estimates across the reported horizons. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Alt text: Table reporting interaction effects among textual, vocal, and facial emotional cues across exchange-rate and equity ETF markets at selected local-projection horizons, together with standard errors, statistical significance, and weighted sums of the interaction estimates.

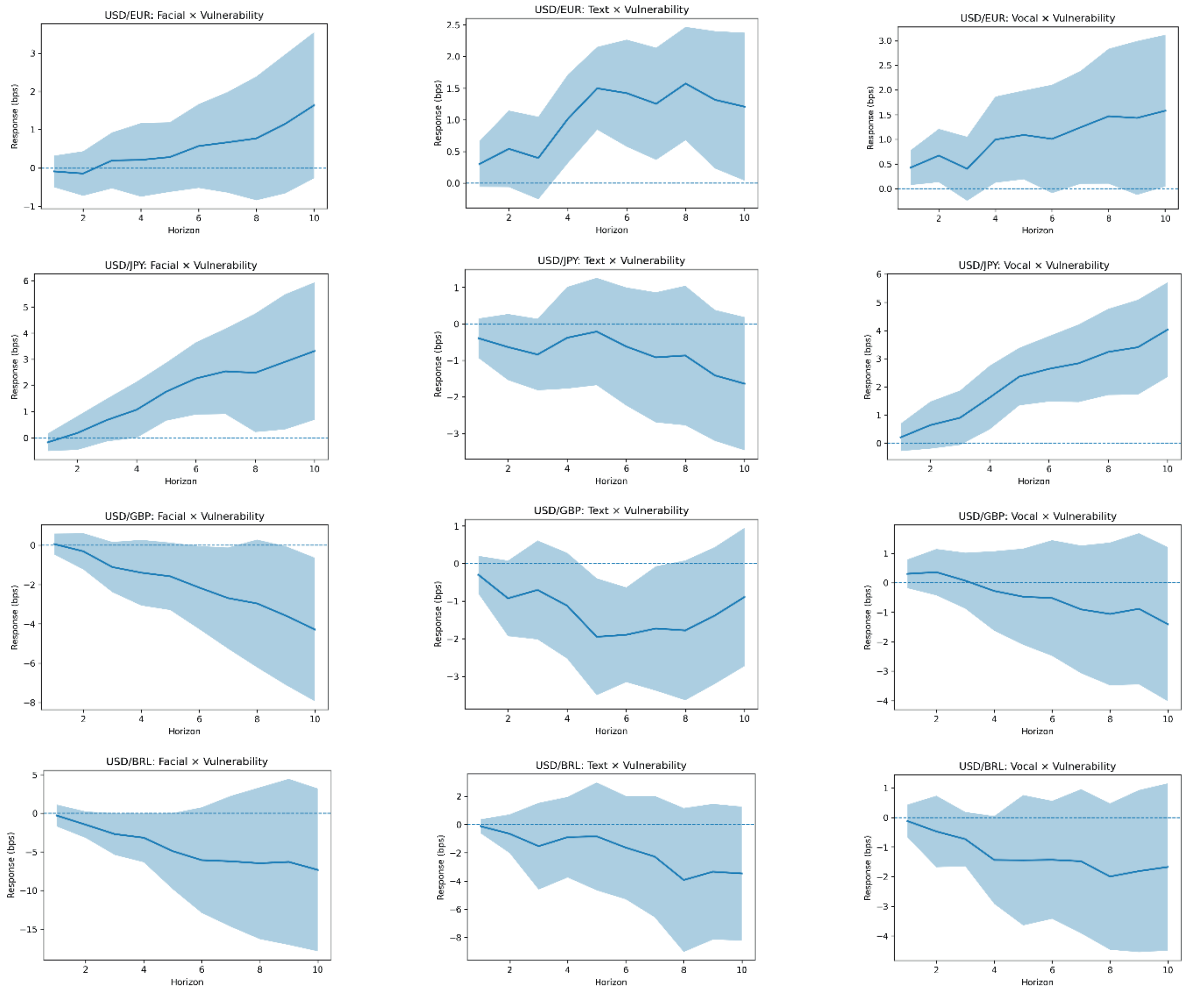
Moderating Effects of Macro-Financial Vulnerability

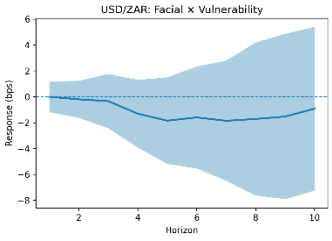
This section builds on the cross-country differences documented in the previous section and further investigates the moderating role of macro-financial vulnerability in different countries.

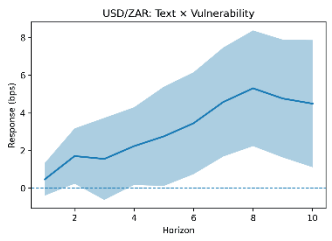
The results in Figure 5 indicate that emotional signals conveyed by the Fed Chair can generate significant cross-border spillover effects, with impacts that are not simply uniform across countries but instead depend critically on their macro-financial vulnerability. Economies with higher vulnerability, such as the United Kingdom and Brazil, tend to exhibit relatively stronger and more pronounced responses in their financial markets, including the exchange rate and equity markets. Positive emotional signals are associated with local-currency appreciation and higher equity returns in these economies. In this context, macro-financial vulnerability plays a role as a moderating mechanism, increasing the sensitivity of domestic financial markets to external emotional shocks and resulting in larger asset price adjustments. This result coincides with the country-level vulnerability rankings in Appendix A, where economies with higher vulnerability tend to respond more strongly to emotional signals, while less vulnerable economies, such as China and Eurozone, display more muted reactions.

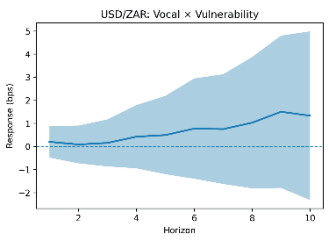
Overall, the results highlight that the effects of emotional signals are not uniform but depend on how such signals are interpreted across different economic environments. These findings provide supporting evidence that the transmission of emotional signals is not uniform across countries but instead depends critically on country-specific structural conditions.

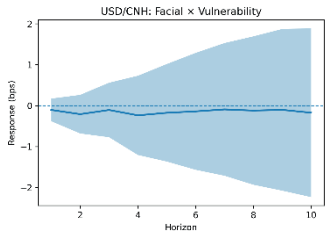
Exchange Rates

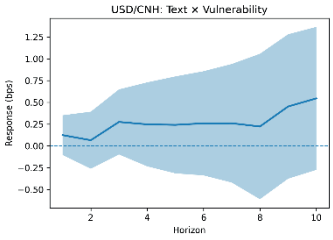


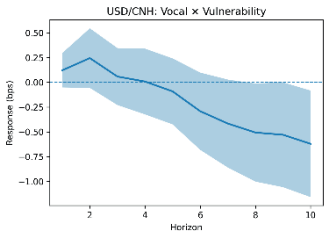

USD/ZAR: Facial × Vulnerability
Response (bps)
Horizon


USD/ZAR: Text × Vulnerability
Response (bps)
Horizon

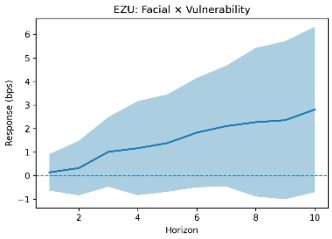

USD/ZAR: Vocal × Vulnerability
Response (bps)
Horizon

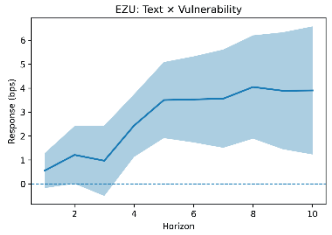

USD/CNH: Facial × Vulnerability
Response (bps)
Horizon

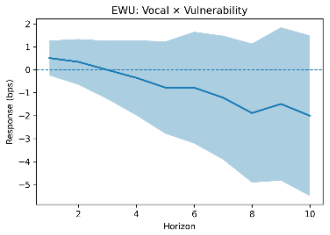

USD/CNH: Text × Vulnerability
Response (bps)
Horizon

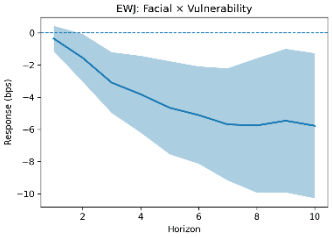

USD/CNH: Vocal × Vulnerability
Response (bps)
Horizon

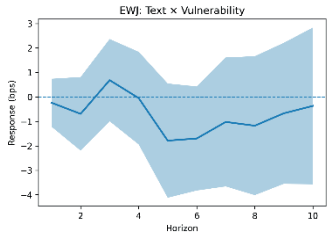
ETFs

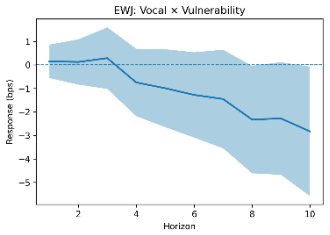

EZU: Facial × Vulnerability
Response (bps)
Horizon

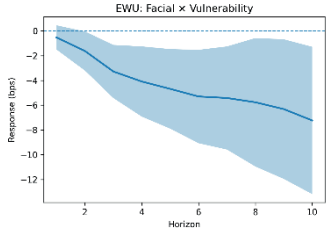

EZU: Text × Vulnerability
Response (bps)
Horizon

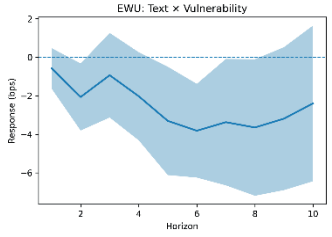

EWU: Vocal × Vulnerability
Response (bps)
Horizon

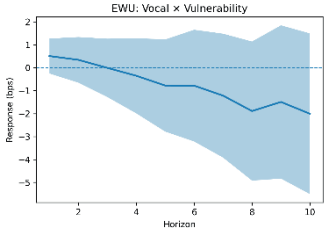

EWJ: Facial × Vulnerability
Response (bps)
Horizon

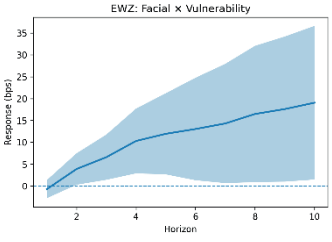

EWJ: Text × Vulnerability
Response (bps)
Horizon

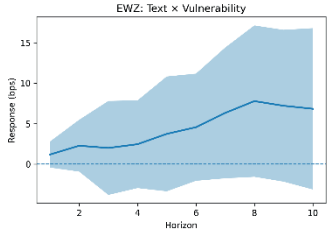

EWJ: Vocal × Vulnerability
Response (bps)
Horizon

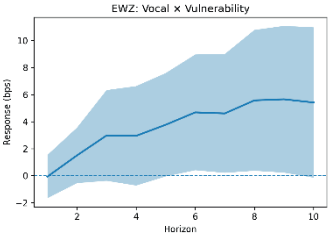

EWU: Facial × Vulnerability
Response (bps)
Horizon


EWU: Text × Vulnerability
Response (bps)
Horizon


EWU: Vocal × Vulnerability
Response (bps)
Horizon


EWZ: Facial × Vulnerability
Response (bps)
Horizon


EWZ: Text × Vulnerability
Response (bps)
Horizon


EWZ: Vocal × Vulnerability
Response (bps)
Horizon

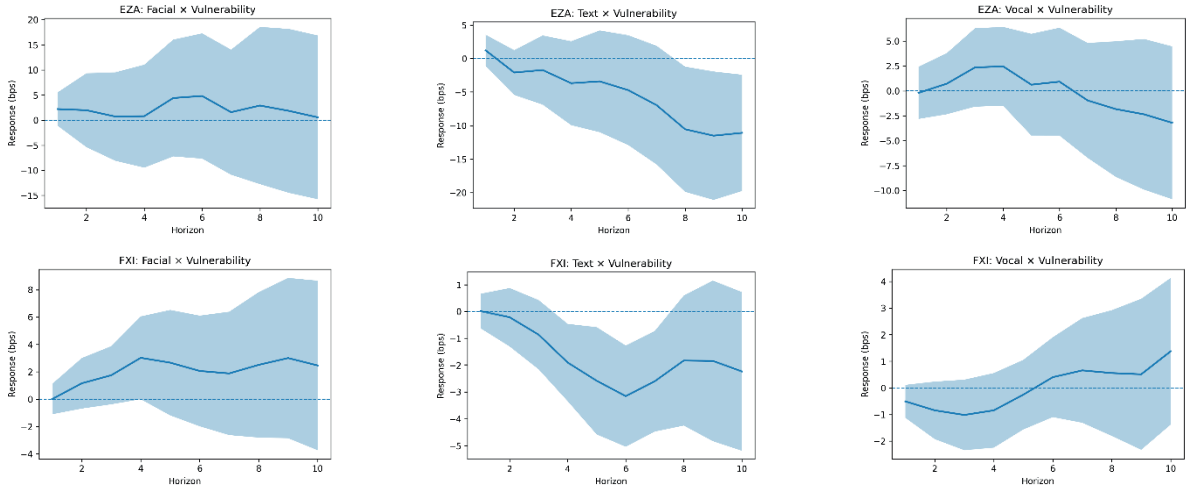


Figure 5. Responses of ETFs and Exchange Rates to Emotional Cues across Different Levels of Country Vulnerability.

Notes: The figure presents the estimated responses of exchange rates and ETFs to the Fed Chair's facial, textual, and vocal emotional cues during FOMC press conferences, conditional on country-level macro-financial vulnerability. The three columns report the interaction effects involving facial, textual, and vocal emotional cues, respectively. Within each panel, the economies are ordered as the Eurozone, Japan, the United Kingdom, Brazil, South Africa, and China. The shaded areas represent the 90% confidence intervals.

Alt text: Impulse-response plots showing how the estimated responses of exchange rates and equity ETFs to facial, textual, and vocal emotional cues vary with country-level macro-financial vulnerability across six economies. Shaded areas represent 90% confidence intervals.

State-Dependent Effects under Global Risk Conditions

The state-dependent impulse responses are presented in Figure 6, which compares the effects of emotional signals under low-VIX and high-VIX, which present low- and high-risk environments, respectively. The results show a significant amplification effect in high-VIX periods.

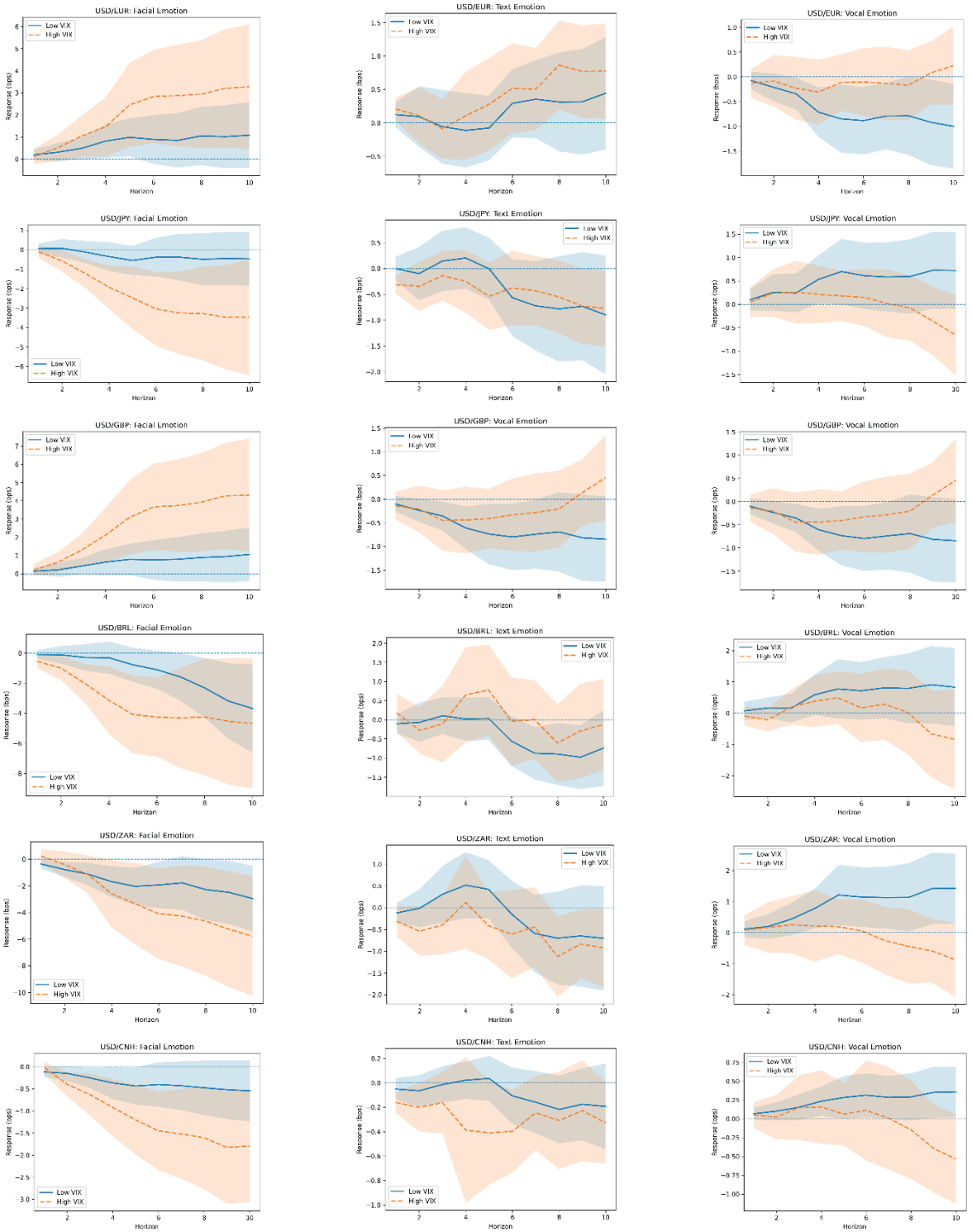
For exchange rates, positive facial emotional signals are associated with significantly larger declines in USD/BRL, USD/ZAR, and USD/CNH under high-VIX conditions relative to low-VIX periods. This indicates a stronger appreciation of local currencies in emerging markets during periods of elevated global uncertainty. In contrast, the responses in low-VIX environments are generally smaller in magnitude and less persistent. These findings suggest that emotional communication has a more pronounced effect on exchange rate dynamics when market-wide uncertainty is high.

A similar pattern emerges in equity markets. Emerging market ETFs, such as EWZ and EZA, exhibit substantially larger and more persistent positive responses under high-VIX regimes. In low-VIX periods, the responses remain positive but are notably weaker and less dynamic. This amplification effect is particularly evident at medium to longer horizons, indicating that emotional signals, especially facial emotion, not only trigger immediate reactions but also have more persistent effects when uncertainty is elevated. For Europe and the United Kingdom, the state dependence is less pronounced. While some differences between high-VIX and low-VIX regimes can still be observed, the magnitude of the responses is generally smaller. This suggests that the state dependence of emotional spillovers is generally less pronounced in the advanced economies considered in the sample.

Overall, the evidence highlights that the effects of emotional communication are highly heterogeneous across countries and market environments. Emotional signals from the Fed Chair generate stronger and more persistent responses in emerging markets and during periods of elevated global uncertainty, while their impact is comparatively weaker in advanced economies. The results suggest that the transmission of emotional signals depends critically on global risk conditions. Importantly, these findings support H4 and further indicate

that emotional signals are informative and that their influence varies systematically across countries and market states, contributing to the observed cross-country heterogeneity in exchange-rate and equity-market responses.

Exchange Rates



ETFs

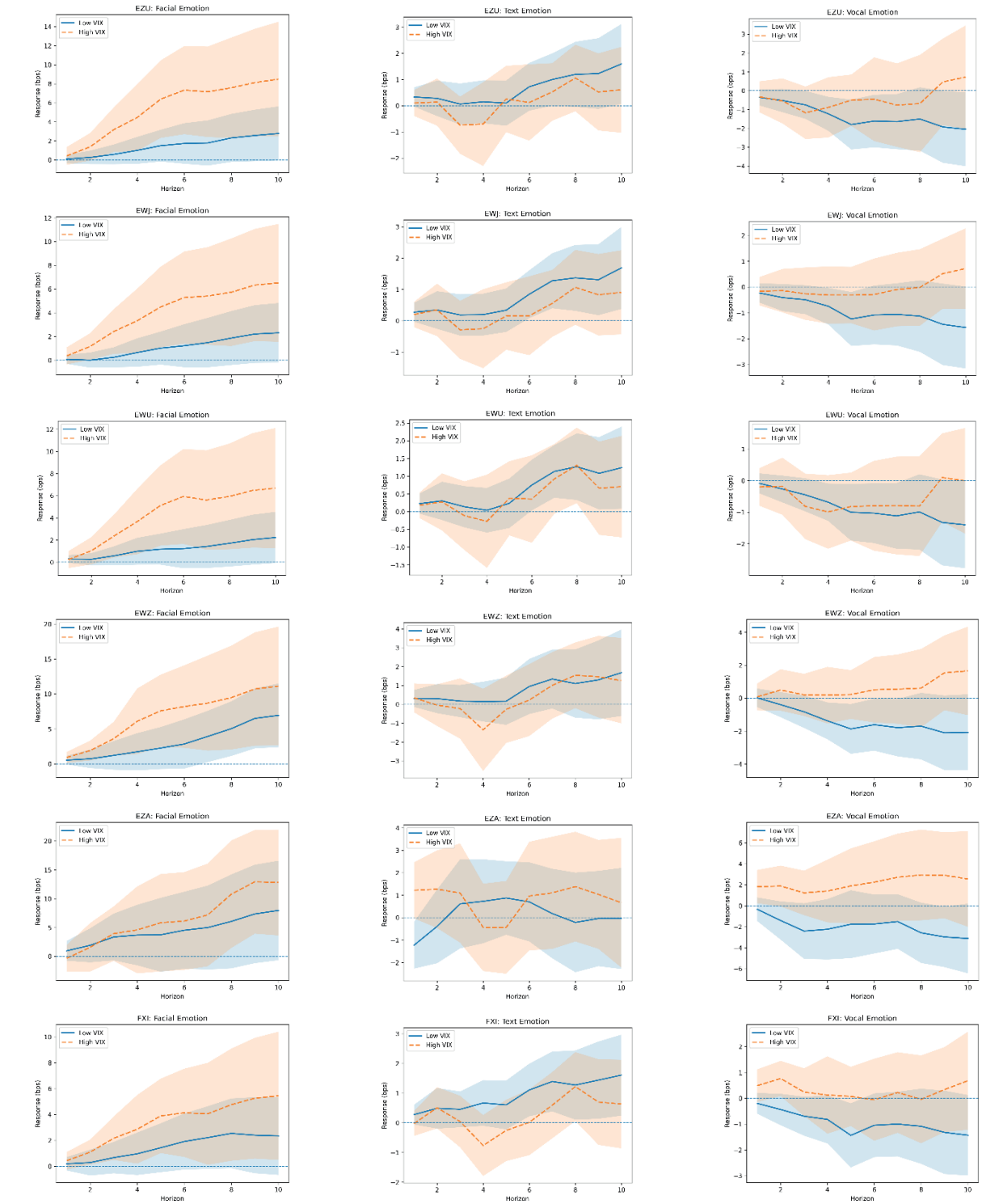


Figure 6. ETF and Exchange Rate Responses to Emotional Cues under High-VIX and Low-VIX Regimes.

Notes: The figure presents the estimated responses of exchange rates and ETFs to the Fed Chair's facial, textual, and vocal emotional cues during FOMC press conferences, estimated separately under high-VIX and low-VIX regimes. The three columns report the responses to facial, textual, and vocal emotional cues, respectively. Within each panel, the economies are ordered as the Eurozone, Japan, the United Kingdom, Brazil, South Africa, and China. The solid blue lines represent the estimates under the low-VIX regime, while the dashed orange lines represent the estimates under the high-VIX regime. The corresponding shaded areas represent 90% confidence intervals.

Alt text: Impulse-response plots comparing exchange-rate and equity ETF responses to facial, textual, and vocal emotional cues under low-VIX and high-VIX regimes across six economies. Solid lines show low-VIX estimates and dashed lines show high-VIX estimates, while shaded areas represent 90% confidence intervals.

Robustness Check

The ML-Generated Regressor Problem

Existing studies support the use of machine-learning methods to extract economically meaningful information from textual, vocal, and facial signals. FinBERT performs well in financial sentiment classification (Araci and Genc, 2020), while ML-generated vocal and facial measures have been shown to capture affective information relevant to financial market outcomes (Gorodnichenko et al., 2023; Curti and Kazinnik, 2023). These studies provide supporting evidence that ML-generated measures can serve as informative proxies for otherwise unobservable emotional signals.

Nevertheless, these measures may still contain prediction errors. Zhang et al. (2026) show that using ML-generated variables as regressors can bias subsequent econometric estimates, potentially affecting coefficient magnitudes and statistical significance. In this study, such errors may also distort the comparison across textual, vocal, and facial emotions. We therefore assess robustness at two levels. First, we examine cross-algorithm consistency using alternative emotion measures. Second, we re-estimate the baseline local projections to test whether the main response patterns and relative importance of the three modalities remain unchanged.

Alternative Emotion-Extraction Approaches

To assess whether the results are sensitive to the choice of emotion-extraction algorithm, we reconstruct each emotion index using an alternative approach. For textual emotion, the FinBERT-based measure is replaced with a FinRoBERTa-based sentiment classifier. FinRoBERTa improves the pre-training procedure of BERT and has demonstrated strong performance across a range of NLP tasks (Liu et al., 2019). Sentence-level sentiment predictions are aggregated to the block level using the same procedure as the baseline textual index.

For vocal emotion, OpenSMILE is used as an alternative to the Praat-based measure. OpenSMILE is a widely used toolkit for extracting acoustic features in speech and affective-computing research (Eyben et al., 2010). The extracted features are processed using the same standardization, principal-component analysis, and block-level aggregation procedures as in the baseline analysis. For facial emotion, Facial Expression Recognition (FER) is employed as an alternative to DeepFace. FER, a Python package that uses a pre-trained deep-learning model, detects facial expressions and estimates probabilities across seven emotion categories. To maintain comparability, video frames are sampled at the same frequency and the alternative facial index is constructed using the same emotion classification and block-aggregation rules as the baseline measure. All alternative indices are subsequently standardized and incorporated into the baseline local projection specifications without changing the dependent variables, controls, fixed effects, or estimation horizons.

Cross-Algorithm Measurement Consistency

Table 5 reports the consistency between the baseline and alternative emotion measures. Pearson and Spearman coefficients assess whether the alternative algorithms preserve the linear and rank-order variation in the baseline indices, while Lin's Concordance Correlation Coefficient (CCC) evaluates their numerical concordance (Lin, 1989). Directional agreement measures the proportion of matched blocks classified in the same positive,

neutral, or negative direction.

The results indicate substantial cross-algorithm consistency across all three modalities. For textual emotion, the FinBERT and FinRoBERTa measures exhibit strong linear and rank-order associations, with Pearson and Spearman coefficients of 0.832 and 0.858, respectively. The two measures also agree on the direction of textual emotion in 86.77% of the matched blocks. Vocal emotion displays the strongest cross-algorithm consistency. The Praat- and OpenSMILE-based measures yield Pearson and Spearman coefficients of 0.987 and 0.985, a CCC of 0.985, and directional agreement of 94.88%. These results indicate a high degree of consistency in the vocal emotion index across alternative acoustic-feature extraction approaches.

The DeepFace and FER facial measures are also strongly associated, with Pearson and Spearman coefficients of 0.786 and 0.798, respectively, and a CCC of 0.785. Their directional agreement reaches 79.31%, indicating that the two facial-expression algorithms assign the same emotional direction to approximately four-fifths of the matched blocks. Although the consistency of the facial measures is lower than that of the textual and vocal measures, the results still indicate that their broad variation and directional classification are preserved across algorithms. Overall, the evidence suggests that the constructed emotion measures are not entirely dependent on a particular extraction algorithm. Nevertheless, cross-algorithm consistency does not demonstrate that the measures are free from measurement error. The alternative indices are therefore subsequently incorporated into the baseline local projection models to examine whether the estimated financial-market responses remain stable.

Table 5. Cross-Algorithm Consistency of Emotion Measures

Modality	Algorithms compared	Pearson	Spearman	Lin's CCC	Directional agreement
Textual	FinBERT–FinRoBERTa	0.832***	0.858***	0.878	86.77%
Vocal	Praat–OpenSMILE	0.987***	0.985***	0.985	94.88%
Facial	DeepFace–FER	0.786***	0.798***	0.785	79.31%

Notes: Pearson and Spearman coefficients measure the linear and rank-order associations between the baseline and alternative emotion measures, respectively. Lin's concordance correlation coefficient (CCC) assesses numerical concordance by accounting for both correlation and differences in means and scales (Lin, 1989). Directional agreement represents the percentage of matched blocks assigned to the same positive, neutral, or negative category by both algorithms. *, **, and *** denote statistical significance of the Pearson and Spearman coefficients at the 10%, 5%, and 1% levels, respectively.

Alt text: Table comparing baseline and alternative emotion-extraction algorithms for textual, vocal, and facial measures. Consistency is evaluated using Pearson correlation, Spearman correlation, Lin's concordance correlation coefficient, and directional agreement.

Stability of Local Projection Estimates

To examine whether the estimated market responses are sensitive to the choice of emotion-extraction algorithm, the baseline local projection models are re-estimated using the alternative textual, vocal, and facial measures. For each modality, the baseline and alternative coefficients are compared across 120 market–horizon combinations, comprising 12 financial-market outcomes and 10 projection horizons. Coefficient stability is assessed using a paired bootstrap coefficient-difference test, where the null hypothesis is that the baseline and alternative coefficients are equal. The procedure resamples FOMC meetings and re-estimates both specifications using the same bootstrap sample, thereby preserving the dependence between the two estimates. Table 6 provides evidence that the main response patterns are generally stable across alternative measurement approaches. The sign-agreement rates are 97.5% for textual emotion, 96.7% for vocal emotion, and 95.8% for facial emotion, indicating that the estimated response directions are rarely altered by the choice of algorithm. Moreover, coefficient equality cannot be rejected at the 5% level in 75.8% of the textual comparisons, 85.8% of the vocal comparisons, and 78.3% of the facial comparisons. Although some coefficient-level differences remain, the high directional agreement and the predominance of non-rejections suggest that the main empirical findings are not systematically driven by a particular emotion-extraction algorithm.

Table 6. Stability of Local Projection Estimates across Alternative Emotion Measures

Modality	Algorithms compared	Sign agreement	Equality rejected share	Equality not rejected share
Textual	FinBERT–FinRoBERTa	0.975	0.242	0.758
Vocal	Praat–OpenSMILE	0.967	0.142	0.858
Facial	DeepFace–FER	0.958	0.217	0.783

Notes: The table compares the baseline and alternative local projection coefficients across 120 market–horizon combinations for each modality, comprising 12 financial market variables and 10 projection horizons. Sign agreement is the proportion of comparisons in which the two estimates have the same sign. The equality rejected share denotes the proportion of comparisons for which the null hypothesis of coefficient equality is rejected at the 5% significance level based on the bootstrap distribution of the coefficient difference. The equality not rejected share denotes the corresponding proportion for which the null hypothesis is not rejected.

Alt text: Table comparing the stability of baseline and alternative local-projection estimates for textual, vocal, and facial emotion measures across 120 market–horizon combinations. The table reports sign agreement and the shares of coefficient-equality tests that are rejected or not rejected at the 5% significance level.

Conclusion

The expectations transmission channel of U.S. monetary policy to global financial markets has attracted considerable attention in recent years. This article advances the literature on central bank communication by introducing an emotional spillover mechanism from U.S. Fed Chair to global financial market, explaining how textual, vocal, and facial emotional signals jointly affect international financial markets beyond conventional monetary policy surprises.

Our theoretical framework and empirical findings suggest that central banks may influence financial markets not only through what they communicate, but also through how they communicate. Moreover, Fed Chair communication is associated not only with U.S. financial markets but also with international financial markets through cross-border emotional spillovers. This study tests these propositions through an analysis of the Fed Chairs' communication from 2011 to 2025 using machine learning methods. The findings suggest that emotional spillovers represent an additional channel through which Federal Reserve communication is associated with global asset prices.

This paper makes several contributions to the research agenda on central bank communication. In particular, the paper highlights that the effects of emotional signals are not uniform across countries, but are driven by different mechanisms. On the one hand, economies with greater macro-financial vulnerability tend to exhibit stronger and more persistent responses, whereas less vulnerable economies generally display more moderate reactions. This cross-country heterogeneity suggests that the same emotional signal can lead to distinct market outcomes depending on country-specific characteristics. On the other hand, the impact of emotional signals is state-dependent and varies systematically with global risk conditions. In particular, emotional spillovers are significantly stronger during periods of higher uncertainty, as proxied by the VIX index. This finding suggests that emotional communication becomes more informative and influential when the information environment is more uncertain, highlighting the role of global risk conditions as a key channel that amplifies the transmission of emotional signals.

Beyond our theoretical and methodological contributions, this study also carries several important policy implications. The findings suggest that central bank communication should be viewed as a multidimensional policy instrument and textual, vocal, and facial signals may each contribute to the way financial markets interpret central bank communication, although their effects differ substantially across modalities. These insights may assist policymakers, including Fed Chairs, in refining their communication strategies and the delivery of policy messages, particularly during periods of heightened uncertainty. Greater awareness of non-verbal communication may therefore help policymakers better understand how their communication is interpreted by financial markets. Moreover, because emotional spillovers are stronger in more vulnerable economies, policymakers and financial regulators outside the U.S. should incorporate communication-induced spillovers into their financial stability monitoring frameworks. In particular, emerging markets may benefit from strengthening market surveillance and liquidity management around major Federal Reserve communication events.

While our results show that the effects of Fed Chair communication extend beyond verbal content in global financial markets, several questions remain for future research. The analysis provides evidence on the moderating roles of macro-financial vulnerability and global risk conditions, but further work could examine the mechanisms through which emotional signals influence investors' beliefs, risk perceptions, and portfolio decisions.

There is also scope to improve the measurement and interpretation of nonverbal signals. The alternative measures used in the robustness analysis provide further support for the consistency of the main findings across different algorithms. Building on these results, future research could further refine these measures by using more detailed emotion categories and training data tailored more closely to real-world financial communication. Facial measures based on basic emotions, for example, may compress distinct facial movements into the same category and overlook short-lived changes in expression. The Facial Action Coding System (FACS) and micro-expression analysis offer a more granular approach by identifying individual Action Units and preserving their temporal dynamics (Hamm et al., 2011;Dong et al., 2022). Measured at a higher sampling frequency, these features could be aligned more closely with high-frequency financial-market responses, consistent with recent research using dynamic facial information (Banker et al., 2024).

Vocal measurement could be developed in a similar direction. Evidence from corporate conference calls highlights the value of models trained on naturally occurring financial communication, together with human-coded validation samples and consideration of differences across speakers, accents, and dialects (Ewertz et al., 2026). Bringing these developments together within a domain-adapted multimodal framework would allow future research to examine whether textual, vocal, and facial signals reinforce or contradict one another, providing a richer account of how investors interpret emotional signals in central bank communication.

References

Ahmed, S., Akıncı, Ö., and Queralto, A. (2026). 'U.S. monetary spillovers to emerging markets: Both policy drivers and vulnerabilities matter', Journal of International Economics, 162, 104288.

Ahrens, M. et al. (2025). 'Mind your language: Market responses to central bank speeches', Journal of Econometrics, 249, p. 105921.

Akansu, A. et al. (2017). 'Firm Performance in the Face of Fear: How CEO Moods Affect Firm Performance', Journal of Behavioral Finance, 18(4), pp. 373–389.

Alexopoulos, M., Han, X., Kryvtsov, O., and Zhang, X. (2024). 'More than words: Fed chairs' communication during congressional testimonies', Journal of Monetary Economics, 142, p. 103515.

Araci, D. and Genc, Z. (2020). 'FINBERT: Financial sentiment analysis with pre-trained language models', in Proceedings of the International Conference on Learning Representations (ICLR 2020).

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

Baltrušaitis, T., Ahuja, C., and Morency, L. (2019). 'Multimodal machine learning: A survey and taxonomy', IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), pp. 423–443.

Banker, R., Ding, H., Huang, R., and Li, X. (2024). 'Market reaction to CEOs' dynamic hemifacial asymmetry of expressions', Management Science, 70(7), pp. 4847–4874.

Bauer, M. and Neely, C. (2014). 'International channels of the Fed's unconventional monetary policy', Journal of International Money and Finance, 44, pp. 24–46.

Belin, P., Boehme, B., and McAleer, P. (2017). 'The sound of trustworthiness: Acoustic-based modulation of perceived voice personality', PLOS ONE, 12(10), e0185651.

Bertsch, C. et al. (2025). 'Central bank mandates and monetary policy stances: Through the lens of Federal Reserve speeches', Journal of Econometrics, 249, p. 105948.

Binder, C., Kuang, P., and Tang, L. (2026). 'Central bank communication and house price expectations', Journal of the European Economic Association, 24(2), pp. 769–799.

Bitouk, D., Verma, R., and Nenkova, A. (2010). 'Class-level spectral features for emotion recognition', Speech Communication, 52(7–8), pp. 613–625.

Blinder, A., Ehrmann, M., Fratzscher, M., De Haan, J., and Jansen, D. (2008). 'Central bank communication and monetary policy: A survey of theory and evidence', Journal of Economic Literature, 46(4), pp. 910–945.

Boersma, P. and Weenink, D. (2001). 'Praat, a system for doing phonetics by computer', Glot International, 5, pp. 341–345.

Bruno, V. and Shin, H. (2015). 'Capital flows and the risk-taking channel of monetary policy', Journal of Monetary Economics, 71, pp. 119–132.

Caballero, R. and Simsek, A. (2026). 'FCI-plot: Central bank communication through financial conditions', Working paper, 26 March 2026.

Cade, N. (2020). 'Using video to disclose forward-looking information: The effect of nonverbal cues on investors' judgments', Review of Accounting Studies, 25(4), pp. 1444–1474.

Campbell, J., Evans, C., Fisher, J., and Justiniano, A. (2012). 'Macroeconomic effects of Federal Reserve forward guidance', Brookings Papers on Economic Activity, 2012(1), pp. 1–80.

Chen, K., Granville, B., and Matousek, R. (2026). 'Decoding central bank communications with large language models', Journal of International Financial Markets, Institutions and Money, 109, 102325.

Cheng, P. and You, Y. (2025). 'U.S. monetary policy spillovers to emerging market countries: Responses to cost-push and natural rate shocks', Economic Modelling, 143, p. 106971.

Cho, D. and Jung, J. (2025). 'Machine learning goes beyond: time-varying monetary policy and oil price pass-through to inflation expectations', Journal of Macroeconomics, 85, p. 103702.

Cho, D. and Rho, S. (2026). 'Can the tone of central bankers' speeches help shape inflation expectations? Evidence from Japan', Journal of International Financial Markets, Institutions and Money, 107, p. 102283.

Cieslak, A. and Schrimpf, A. (2019). 'Non-monetary news in central bank communication', Journal of International Economics, 118, pp. 293–315.

Cloyne, J., Jordà, Ò., and Taylor, A. (2023). State-dependent local projections: Understanding impulse response heterogeneity. NBER Working Paper No. 30971.

Coibion, O., Gorodnichenko, Y., and Weber, M. (2025). 'The cost of the COVID-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending', Journal of Economic Behavior & Organization, 229, 106846.

Curti, F. and Kazinnik, S. (2023). 'Let's face it: Quantifying the impact of nonverbal communication in FOMC press conferences', Journal of Monetary Economics, 139, pp. 110–126.

Custódio, C., Ferreira, M., and Matos, P. (2013). 'Generalists versus specialists: Lifetime work experience and chief executive officer pay', Journal of Financial Economics, 108(2), pp. 471–492.

Dedola, L., Rivolta, G., and Stracca, L. (2017). 'If the Fed sneezes, who catches a cold?', Journal of International Economics, 108, pp. S23–S41.

Degasperi, R., Hong, S., and Ricco, G. (2026). 'The global transmission of U.S. monetary policy', Journal of International Economics, 161, 104259.

DePaulo, B., Lindsay, J., Malone, B., Muhlenbruck, L., Charlton, K., and Cooper, H. (2003). 'Cues to deception', Psychological Bulletin, 129(1), pp. 74–118.

Dong, Z., Wang, G., Lu, S., Li, J., Yan, W., and Wang, S.-J. (2022). 'Spontaneous facial expressions and micro-expressions coding: From brain to face', Frontiers in Psychology, 12, 784834.

El Ayadi, M., Kamel, M., and Karray, F. (2011). 'Survey on speech emotion recognition: Features, classification schemes, and databases', Pattern Recognition, 44(3), pp. 572–587.

Ewertz, J., Knickrehm, C., Nienhaus, M., and Reichmann, D. (2026). 'Listen closely: Measuring vocal tone in corporate disclosures', Journal of Accounting Research, 64(1), pp. 229–277.

Eyben, F., Wöllmer, M., and Schuller, B. (2010). 'OpenSMILE: The Munich versatile and fast open-source audio feature extractor', in Proceedings of the 18th ACM International Conference on Multimedia, pp. 1459–1462.

Flam, R., Green, J., and Sharp, N. (2020). 'Do investors respond to CEO facial expressions of anger during television interviews?', SSRN Electronic Journal.

Georgiadis, G. (2016). 'Determinants of global spillovers from US monetary policy', Journal of International Money and Finance, 67, pp. 41–61.

Goldin-Meadow, S. (2003). Thought before language: Do we think ergative? In: Gentner, D. and Goldin-Meadow, S. (eds.) Language in mind: Advances in the study of language and thought. Cambridge, MA: MIT Press, pp. 523–552.

Gómez-Cram, R. and Grotteria, M. (2022). 'Real-time price discovery via verbal communication: Method and application to Fedspeak', Journal of Financial Economics, 143(3), pp. 993–1025.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2023). 'The voice of monetary policy', American Economic Review, 113(2), pp. 548–584.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2025). 'Central bank communication on social media: What, to whom, and how?', Journal of Econometrics, 249, p. 105869.

Grebelsky-Lichtman, T. (2017). 'Verbal versus nonverbal primacy: Children's response to parental incongruent communication', Journal of Social and Personal Relationships, 34(5), pp. 636–661.

Gürkaynak, R., Sack, B., and Swanson, E. (2005). 'Do actions speak louder than words? The response of asset prices to monetary policy actions and statements', International Journal of Central Banking, 1(1), pp. 55–93.

Hamm, J., Kohler, C. G., Gur, R. C., and Verma, R. (2011). 'Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders', Journal of Neuroscience Methods, 200(2), pp. 237–256.

Hansen, S. and McMahon, M. (2016). 'Shocking language: Understanding the macroeconomic effects of central bank communication', Journal of International Economics, 99, pp. S114–S133.

Hausman, J. and Wongswan, J. (2011). 'Global asset prices and FOMC announcements', Journal of International Money and Finance, 30(3), pp. 547–571.

Hildebrand, C. et al. (2020). "Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications," Journal of Business Research, 121, pp. 364–374.

Ho, S.-J. and Karagedikli, Ö. (2025). 'Effects of monetary policy communication in emerging market economies: Evidence from Malaysia', Journal of International Money and Finance, 158, p. 103379.

Hoek, J., Kamin, S., and Yoldas, E. (2022). 'Are higher U.S. interest rates always bad news for emerging markets?', Journal of International Economics, 137, p. 103585.

Hüpper, F. and Kempa, B. (2023). 'Inflation targeting and inflation communication of the Federal Reserve: Words and deeds', Journal of Macroeconomics, 75, p. 103497.

Jiang, X. and Pell, M. (2017). "The sound of confidence and doubt," Speech Communication, 88, pp. 106–126.

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kamiloğlu, R., Fischer, A., and Sauter, D. (2020). "Good vibrations: A review of vocal expressions of positive emotions," Psychonomic Bulletin & Review, 27(2), pp. 237–265.

Kanelis, D. and Siklos, P. (2023). 'Emotion in euro area monetary policy communication and bond yields: the Draghi era', SSRN Electronic Journal.

Kim, S. (2001). 'International transmission of U.S. monetary policy shocks: Evidence from VARs', Journal of Monetary Economics, 48(2), pp. 339–372.

Kjeldsen, J. and Hess, A. (2021). 'Experiencing multimodal rhetoric and argumentation in political advertisements: A study of how people respond to the rhetoric of multimodal communication', Visual Communication, 20(3), pp. 327–352.

Krishnamurthy, A., Nagel, S., and Orlov, D. (2014). 'Sizing up repo', Journal of Finance, 69(6), pp. 2381–2417.

Laukka, P. et al. (2008). 'In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech', Journal of Nonverbal Behavior, 32(4), pp. 195–214.

Lehmann, J., Goussios, C., and Seufert, T. (2016). 'Working memory capacity and disfluency effect: An aptitude–treatment interaction study', Metacognition and Learning, 11, pp. 89–105.

Lin, L. (1989). 'A concordance correlation coefficient to evaluate reproducibility', Biometrics, 45(1), pp. 255–268.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019). 'RoBERTa: A robustly optimized BERT pretraining approach', arXiv preprint, arXiv:1907.11692.

Manjhi, G. and Shah, R. (2025). 'What do sentiments of budget speeches mean for stock returns?', Economic Modelling, 151, p. 107156.

Martinec, R. and Salway, A. (2005). 'A system for image–text relations in new (and old) media', Visual Communication, 4(3), pp. 337–371.

Mathur, P., Neerkaje, A., Chhibber, M., Sawhney, R., Guo, F., Dernoncourt, F., Dutta, S., and Manocha, D. (2022). MONOPOLY: Financial prediction from monetary policy conference videos using multimodal cues. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 2276–2285. Association for Computing Machinery.

Mayew, W. and Venkatachalam, M. (2012). 'The power of voice: Managerial affective states and future firm performance', The Journal of Finance, 67(1), pp. 1–44.

McCauley, R. (2012). Risk-on/risk-off, capital flows, leverage and safe assets. BIS Working Papers No. 382. Bank for International Settlements.

Mehrabian, A. (1972). Nonverbal Communication. Chicago: Aldine-Atherton.

Miranda-Agrippino, S. and Rey, H. (2020). 'U.S. monetary policy and the global financial cycle', The Review of Economic Studies, 87(6), pp. 2754–2776.

Nakamura, E. and Steinsson, J. (2018). 'High-frequency identification of monetary non-neutrality: The information effect', Quarterly Journal of Economics, 133(3), pp. 1283–1330.

Nasir, M. (2021). 'Zero lower bound and negative interest rates: Choices for monetary policy in the UK', Journal of Policy Modeling, 43(1), pp. 200–229.

Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), pp. 1–135.

Pettarin, A. (2017). aeneas: a Python/C library for forced alignment. GitHub repository.

Pon-Barry, H. and Shieber, S. (2010). "Recognizing Uncertainty in Speech," EURASIP Journal on Advances in Signal Processing, 2011(1).

Qin, Y. and Yang, Y. (2019). What you say and how you say it matters: Predicting stock volatility using verbal and vocal cues. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 390–401. Association for Computational Linguistics.

Renzhi, N. and Beirne, J. (2024). 'Global shocks and monetary policy transmission in emerging markets', ADB Economics Working Paper Series, No. 726.

Sakoe, H. and Chiba, S. (1978). 'Dynamic programming algorithm optimization for spoken word recognition', IEEE Transactions on Acoustics, Speech, and Signal Processing, 26(1), pp. 43–49.

Scherer, K. (2003). 'Vocal communication of emotion: A review of research paradigms', Speech Communication, 40(1–2), pp. 227–256.

Serengil, S. and Ozpinar, A. (2024). 'A benchmark of facial recognition pipelines and co-usability performances of modules in face recognition', The Journal of Supercomputing, 80, pp. 4875–4903.

Sweller, J. (1988). 'Cognitive load during problem solving: Effects on learning', Cognitive Science, 12(2), pp. 257–285.

Wu, J. and Xia, F. (2016). 'Measuring the macroeconomic impact of monetary policy at the zero lower bound', Journal of Money, Credit and Banking, 48(2–3), pp. 253–291.

Yang, Y. et al. (2023). "Unlocking the Power of Voice for Financial Risk Prediction: A Theory-Driven Deep Learning Design Approach," MIS Quarterly, 47(1), pp. 63–96.

Yazdani, E., Chakravarty, A., and Inman, J. (2024). "(Mis)alignment between facial and textual emotions and its effects on donors engagement behavior in online crowdsourcing platforms," Journal of the Academy of Marketing Science, 53(4), pp. 968–988.

Zehri, C., Madjd-Sadjadi, Z., and Ammar, L. (2024). 'Asymmetric impacts of US monetary policy on emerging markets: contagion and macroeconomic determinants', Journal of Economic Asymmetries, 29, e00354.

Zhang, J., Xue, W., Yu, Y., and Tan, Y. (2026). 'Debiasing ML- or AI-generated regressors in partially linear models', Information Systems Research, published online 27 May.