IN THE FED WE TRUST? MEASURING TRUST IN CENTRAL BANKING AND ITS EFFECTS ON THE MACROECONOMY

David Aikman, Francesca Monti, Shunshun Zhang

LIDAM Discussion Paper CORE
2026 / 17

UCLouvain

LIDAM
CORE | Louvain Institute of Data Analysis and Modeling in economics and statistics

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers

In the Fed we trust? Measuring trust in central banking and its effects on the macroeconomy

David Aikman* Francesca Monti† Shunshun Zhang‡

July 14, 2026

Abstract

We develop a novel measure of trust in the Federal Reserve using Generative Artificial Intelligence to analyse millions of tweets referencing the Fed, its leadership, and its policy framework and decisions. Our measure exhibits pronounced swings that align with key events and responds intuitively to macro-financial variables and indicators of the U.S. monetary policy stance. To identify exogenous trust shocks, we use a narrative approach based on ethical scandals involving certain FOMC members. We find that trust shocks significantly affect macroeconomic conditions: news sentiment, business conditions, and the stock market all decline, while inflation expectations and market volatility (VIX) increase.

*National Institute of Economic and Social Research. E-mail: d.aikman@niesr.ac.uk

†UC Louvain and CEPR. E-mail: francesca.monti@uclouvain.be

‡King's College London. E-mail: shunshun.zhang@kcl.ac.uk

We are grateful to the editor, the associate editor and three anonymous referees for their insights and suggestions. Stephen Cecchetti, Ruben Durante, Michael Ehrmann, Domenico Giannone, Mikael Juselius, Michele Lenza and Michael McMahon for helpful comments and discussions. We also thank seminar participants at the BSE Forum, the University of Oxford, Ghent University, King's College London, Université Catholique de Louvain, the Bank of England, Banco Central do Brasil, the Central Bank of Ireland, the Central Bank of the Republic of Turkey, CBUAE, the Central Bank of Ukraine, the London Quant Group, and the 2025 Kiel-CEPR Conference for insightful comments. We thank Swarali Deepak Havaldar for excellent research assistance. This project received generous financial support from King's College London to purchase data from Twitter/X. All errors remain our own.

1 Introduction

Public trust in central banks has come under strain in recent years. Several factors have contributed to this: the unexpected inflation surge beginning in late 2021; the financial system turmoil of late 2022 and early 2023; controversies over quantitative easing and central bank losses; public attacks on policy independence, most prominently by President Trump; and questions of integrity following the resignation of senior central bankers amid alleged financial scandals.

In this paper, we propose a new method for measuring trust in central banks, collecting data from Twitter/X and using Generative Artificial Intelligence (GenAI) to assess whether posts reflect criticism or support of the central bank, its policies, and its leadership. We construct a daily, high-frequency index of public trust that can be used to study how trust evolves around key events in real time. While our method is applicable to any public media, we focus on Twitter/X because it provides a rich dataset of real-time sentiment, and on the US Federal Reserve, whose actions as the world's leading central bank attract significant public discourse.

By *trust*, we mean the public's belief in the goodwill and integrity of the central bank and its staff—essentially, an expectation that the central bank will act in the public interest and possess the technical competence to achieve its objectives.[1] This is broader than the notion of credibility, with its focus on whether the central bank will fulfil its stated commitments (Kydland and Prescott, 1977; Barro and Gordon, 1983). While credibility can be established through formal mechanisms such as legal mandates and explicit targets, trust relies on informal norms and social capital, making it more fragile. Trust is critical to the mission of central banking. Money itself is a social convention, the value of which depends on trust in the money-

[1] According to the OECD, "institutional trust is generated when citizens appraise public institutions and/or the government and individual political leaders as promise-keeping, fair and honest" (OECD, 2017).

issuing institution. There is evidence that higher trust in the central bank is associated with better anchored inflation expectations (Ehrmann et al., 2013; Christelis et al., 2020). And a central bank that commands public trust is more likely to find its actions and communications viewed as credible and reassuring in a crisis, and less prone to short-term political influence.

Despite the importance of trust to central banks' success, quantitative metrics are still few and far between. Most existing trust measures are based on surveys—such as the biannual Eurobarometer—which are infrequent, closed-ended, and subject to various biases. Our measure, by contrast, derives from a continuous stream of unfiltered public commentary, allowing us to capture trust at higher frequency and with finer nuance. We retrieve all tweets referencing the Federal Reserve or its senior leadership that also include at least one keyword from a curated list of policy-related, competence-related, or governance/ethics-related terms. This filtering yields approximately 4 million tweets spanning 2012 through December 2023. We then use GenAI (GPT) to categorise each tweet as supportive, critical, neutral or unrelated to the Fed. Validated against a sample of tweets hand-labelled by the authors, GPT's classification accuracy strongly outperformed both off-the-shelf dictionary-based methods and machine learning algorithms. Our index is the net share of supportive tweets: supportive minus critical, divided by all relevant tweets (supportive, critical and neutral), with each tweet weighted by its engagement (one plus its count of likes and retweets).

The index exhibits pronounced swings that align with key events (Figure 1). Cumulated to highlight persistent shifts,[2] it falls sharply from 2018 as President Trump's critical tweeting intensified, surges following the onset of Covid-19, and falls precipitously from mid-2021 as inflation rose well

[2]Daily net sentiment is noisy, driven as much by ephemeral news coverage as by lasting changes in trust. Figure 1 therefore plots the cumulative sum of daily changes in net sentiment, less a trailing 12-month average to remove slow-moving shifts in the average tone of discourse. This highlights persistent movements in trust rather than short-lived spikes in online conversation. The regression analysis in Section 3 uses the raw daily index.

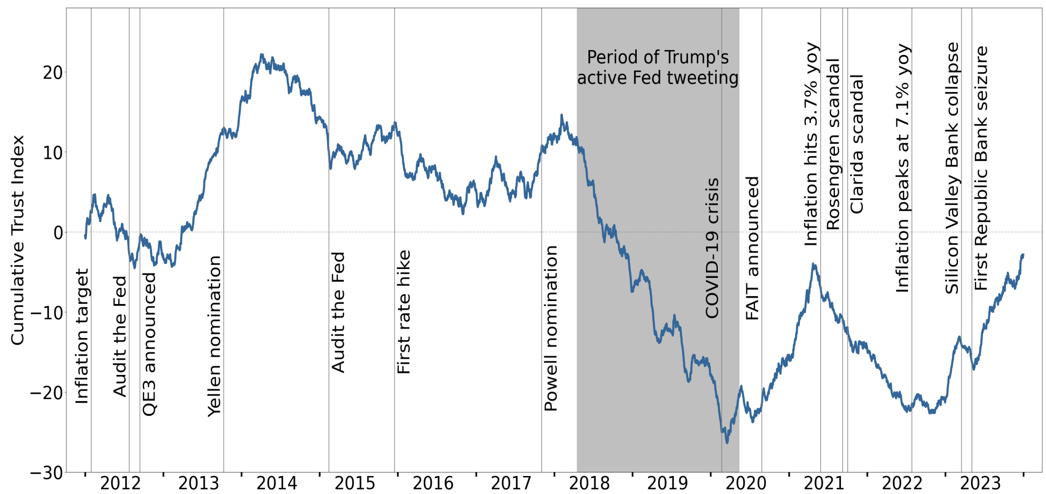

Figure 1: Index of trust in the Fed with key events

Notes: Index of trust in the Fed with key events. This figure shows the cumulative sum of the daily trust index, adjusted for shifts in the average tone of discourse via a trailing 12-month mean.

above target and a string of ethical scandals involving Fed officials came to light. Regression analysis reveals that macro-financial factors such as the yield curve, inflation expectations, and market volatility are significantly correlated with the index, as are monetary policy events: Monetary Policy Report publications, Jackson Hole speeches, and key leadership transitions all enter positively. Ethical scandals affecting senior FOMC members have the largest effect, each reducing the index by about 0.2 (on a -1 to $+1$ scale), while tweets by President Trump perceived as threatening the Fed's independence also have large negative impacts. Notably, newspaper-sentiment indices explain very little variation in our index, suggesting that Twitter/X captures a distinct channel of public discourse.

How should a loss of trust affect the macroeconomy? Theory points, first, to an *uncertainty channel*. In the model of Bursian and Faia (2018), trust is determined in equilibrium by the interaction of a monetary authority with ex-post incentives to deviate from trustworthy conduct and betrayal-averse private agents. A decline in trust is, in effect, a rise in uncertainty about the conduct of policy: agents perceive a higher probability of betrayal, demand higher risk premia, and shift towards safe, lower-yielding assets, weighing on aggregate demand and investment. At the same time, monetary policy gains less traction over the behaviour of distrustful agents, worsening the perceived inflation-output trade-off—so inflation expectations rise even as demand falls. A negative trust shock should therefore carry the footprint of an adverse supply shock: higher uncertainty and risk premia, weaker activity and equity prices, and higher inflation expectations. This mechanism connects to the broader literatures on uncertainty and real activity (Bloom, 2009; Fernández-Villaverde et al., 2015) and on political uncertainty and risk premia (Pástor and Veronesi, 2012, 2013), and its trigger is grounded in evidence that integrity violations durably erode trust in institutions and financial markets (Guiso et al., 2008; Giannetti and Wang, 2016; Gurun et al., 2018).

We assess the effects of shocks to our trust index using a narrative identification approach that exploits the news of alleged ethical scandals embroiling FOMC officials. While the alleged behaviour is potentially endogenous to the state of the economy, the dates on which the news broke—often years after the conduct in question—are plausibly unrelated to prevailing economic conditions. We treat these news dates as restrictions on the sign of the trust shock and, adopting the insight of Plagborg-Møller and Wolf (2021), recast them as a discrete-valued proxy with which to identify impulse responses in a proxy-Structural Vector Autoregression (SVAR) (Mertens and Ravn, 2013; Stock and Watson, 2018)). We find that shocks that reduce trust in the Federal Reserve have a persistent contractionary effect on the economy: business conditions, news sentiment, and the stock market decline, while the VIX rises and inflation expectations increase temporarily, as predicted by theory. Market expectations of the near-term policy rate rise only briefly and modestly.

That last result points to a second channel, operating through *expecta-*

tions of the policy path—where the type of shock matters. Our identifying events are integrity shocks: revelations of misconduct by individual officials that leave the institution's policy strategy unchanged. They carry no directional signal about future policy, so any movement in rate expectations should reflect only the endogenous response of policy to the shock's macroeconomic effects—hence the brief rise, mirroring the temporary rise in inflation expectations. Direct political pressure is different: President Trump's policy-criticism tweets, mostly calls to lower interest rates (Bianchi et al., 2023), explicitly advocate a policy direction. When we use these tweets as an alternative instrument, rate expectations decline significantly and persistently, consistent with markets attaching some probability to policymakers yielding to political pressure.

While we are the first to use Twitter/X data to assess trust in central banking, such data have been used for related purposes: to study how experts and non-experts respond to ECB communication (Ehrmann and Wabitsch, 2022), to analyse the Fed's own use of social media and its effect on inflation expectations (Gorodnichenko et al., 2025; Conti-Brown and Feinstein, 2020; Masciandaro et al., 2024), and to construct a proxy for consumer inflation expectations (Angelico et al., 2022). Closest to our work is Bianchi et al. (2023), who examine how President Trump's tweets about the Fed affected monetary policy expectations, the stock market, and bond premia; Drechsel (2026) provides complementary evidence on political pressure on the Fed using historical data. Our contribution is that we provide a new measure of trust in the institution, and study the impact of plausibly exogenous trust shocks derived from integrity-based events.

Our paper also relates to the literature on trust and communication in central banking, comprehensively reviewed by Ehrmann (2026). On communication, Coibion et al. (2022) show that the effect of monetary policy messaging on household inflation expectations depends on its medium and content, and Gorodnichenko et al. (2023) that its tone—especially in press conferences—moves asset prices and expectations. Our findings complement this literature by showing how a real-time, sentiment-based measure of trust can inform both empirical work and theoretical modelling of trust and monetary policy.

2 Data and methodology

This section describes our approach to locating relevant tweets about trust in the Fed, its leadership and its policy framework and decisions, and our use of GenAI to analyse the sentiment expressed in these data. The queries and prompts used are listed in Appendix A.1 and Appendix A.2.[3]

2.1 Twitter/X data

Twitter/X provides a timely and diverse source of public commentary on the Federal Reserve, including posts from journalists, financial-market participants, policy observers, elected officials, and other members of the public.[4] As voluntary public expressions, Twitter/X posts capture a selected but highly responsive segment of public commentary, making them particularly useful for measuring high-frequency reactions to monetary policy and institutional events. Previous research has shown that sentiment expressed on Twitter/X can influence both policy discourse and financial market dynamics (Bianchi et al., 2023).

Our objective was to extract from these noisy but timely data a measure of how the Federal Reserve is perceived by the public in terms of competence and integrity. To construct a relevant sample, we designed a multi-stage filtering process. First, we identified tweets containing references to the Fed-

[3]The Online Appendix provides further detail and analysis.

[4]According to Statista, Twitter/X had 237.8 million daily active users and 541.6 million monthly active users in 2023: `https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/`.

eral Reserve or its leadership. This initial query, which included terms such as “the Fed”, “Federal Reserve”, “FOMC”, and the names of key officials, returned 61,922,464 tweets over the 2007–2023 sample period. However, many of these were irrelevant due to homonymy: for instance, “Fed” is frequently used to refer to FBI agents (“the feds”), the federal government, or even the tennis player Roger Federer. To address this, we excluded tweets containing phrases such as “fed gov”, “the fed election”, “the feds”, and “fed ex”, and restricted instances of “fed” to those preceded by a definite article (“the fed”).[5] This step reduced the dataset to 14,104,900 tweets.

Second, we refined the subset of tweets referencing individual Federal Reserve officials. We included the names of all Fed Chairs since 2007—Ben Bernanke, Janet Yellen, and Jerome Powell—as well as prominent policymakers such as Donald Kohn, Stanley Fischer, Richard Clarida, Lael Brainard, Philip Jefferson, Mary Daly, Eric Rosengren, Robert Kaplan, Raphael Bostic, James Bullard, Jeffrey Lacker, Michelle Bowman and Lisa Cook. To reduce false positives from common surnames (e.g., “Powell”), we required that these names co-occur with relevant terms such as “market”, “economy”, “business”, “chair”, “president”, or the names of regional Federal Reserve Banks. We also include widely used nicknames where applicable (e.g., “JayPow”). This step yielded 6,608,917 tweets referencing Fed officials.

Third, we sought to isolate tweets that expressed subjective opinions about the Fed, its officials, or its policy actions. To this end, we constructed a domain-specific lexicon of terms related to trust, credibility, and institutional perception. This lexicon was based on an analysis of 40,000 tweets involving Fed officials. Using a generative AI model (ChatGPT), we identified tweets expressing support or criticism, and computed term frequencies to

[5]The filtering step is intended to remove obvious irrelevant uses of “Fed”, but it is not the only mechanism used to identify irrelevant tweets. In the subsequent GenAI classification step, each tweet is classified as positive, neutral, negative, or unrelated with respect to the Fed or its leadership. Remaining homonymous or irrelevant uses not removed by the initial keyword filters are therefore classified as unrelated.

isolate words associated with opinionated content.[6] We grouped these words into three thematic categories: policy, ethics/governance, and competence. The resulting lexicon included terms such as "credibility", "scandal", "trustworthy", "corrupt", "competent", and "foolish", as well as relevant hashtags such as "#auditthefed".[7] Table 1 lists the full set of terms.

Combining these filters, our final query included tweets that (a) referenced the Fed or a senior Fed official, and (b) contained at least one term from our trust lexicon (see Appendix A.1 for the exact query).

We used the Twitter/X API count function to examine the availability of tweets mentioning the Fed over time. The count results indicate that, from 2012 onward, relevant tweets are available in sufficiently large and stable daily volumes to support continuous daily analysis. Therefore, we restrict our Twitter/X sample to the period from 2012 to 2023, which ensures continuous daily coverage.

Table 2 reports summary statistics on daily volumes of Fed-related tweets by year. The table summarizes the final sample of 3,859,454 observations.[8] The daily volume of Fed-related tweets increased over the sample period, with particularly sharp increases in 2019 and again in 2022. The increase in 2019 coincides with heightened public attention to monetary policy, while the 2022 spike is consistent with the surge in inflation-related discussions and the rapid tightening of monetary policy. In contrast, tweet volume declines in 2023 rel-

[6] Our primary goal is to identify tweets expressing evaluative opinions. Given the subjective and ambiguous nature of such expressions, our lexicon-based filtering approach is designed to favor relevance. This conservative strategy ensures that the retained set is likely to contain meaningful content. Cross-validation with ChatGPT-labeled data confirms the effectiveness of this method.

[7] The selected hashtags are included because they are used as strong sentiment signals in the pre-classification step: #auditthefed and #endthefed are classified as negative, while #popyourcollar is classified as positive. See Online Appendix A.1 for details.

[8] Trump's Twitter/X account was suspended for part of our sample period. As a result, his tweets were not fully available via the Twitter/X API handle, even if they respected our criteria. We therefore collected these tweets separately from The American Presidency Project and added them manually. Including these 90 tweets does not affect our results at all. See the Online Appendix for more details.

Table 1: Lexicon of context-specific synonyms for confidence and trust in the Fed

Category	**Words**
Policy	aggressive, authority, bubble, credibility, credible, crisis, decision, dissent, dots, easy, fedspeak, inflation, loose, overheat, policy, qe, regulate, regulation, regulator, remit, soundness, stability, strategy, surprise, taper, tapering, testimony, threat, tight, volatility
Ethics/Governance	accountability, bitcoin, controversy, corrupt, corruption, criminal, disclosure, disruption, ethic, ethical, ethics, gold, governance, independence, influence, insider, integrity, investigation, jail, regulatebitcoin, reliability, scandal, scrutiny, trading, unethical, violate, violation
Competence	abuse, assurance, blame, bogus, bombshell, buffoon, bulls**t, catastrophic, cautious, clear, clown, clumsy, comfort, competent, composed, confident, conflict, confuse, consistent, crazy, critic, deception, denial, disgrace, disgraceful, distrust, doubt, erratic, excuse, experience, experienced, fake, fear, fickle, flop, fool, foolish, fraud, gaffe, genius, honest, idiot, ignorant, innocent, knowledgeable, lie, logical, loonie, mistake, muddle, muddy, nonsense, nuts, outrage, outrageous, overreact, panic, predictable, rational, rookie, sensible, smart, socialist, sophisticated, stable, steady, tantrum, trouble, trump, trust, trustworthiness, trustworthy, unpredictable, vague, wacky, wise
Hashtags	#auditthefed, #endthefed, #popyourcollar

Table 2: Summary statistics of daily tweets about the Fed

	min	mean	median	max	25%	75%	<10	std	count
2012	40	364.79	215.5	4,045	127.25	401.5	0	450.40	133,514
2013	45	482.25	312	5,546	196	494	0	611.77	176,022
2014	39	476.39	264	9,984	161	411	0	840.91	173,881
2015	53	532	300	7,878	180	536	0	823.98	194,180
2016	72	612.08	340	7,034	206	608.25	0	855.88	224,020
2017	82	612.15	372	9,588	237	621	0	918.23	223,433
2018	75	580.70	339	6,882	225	566	0	796.18	211,956
2019	127	804.27	556	9,089	387	913	0	777.87	293,558
2020	207	904.35	707.5	10,699	504	998.25	0	832.26	330,993
2021	263	1,102.08	867	6,311	597	1,292	0	802.92	402,258
2022	448	2,290.03	1,820	10,526	1,369	2,735	0	1,527.25	835,860
2023	112	1,807.61	1,348	11,618	990	1,923	0	1,538.17	659,779

Notes: Summary statistics by year. The columns of this table report, from left to right, the minimum, mean, median and maximum number of tweets in each year, as well as the first and the third quartiles, whether there were any days with less than 10 tweets, the standard deviation and the total number of tweets in the year of reference.

ative to 2022. This decline should be interpreted with caution. Elon Musk's takeover of Twitter/X in October 2022 was followed by substantial changes to the platform, including account suspensions and deletions, changes in user activity, and changes in data access conditions. These platform changes may have reduced both the number of observable tweets and the number of active accounts posting about the Fed. Therefore, the decline in 2023 may reflect not only changes in public attention to monetary policy, but also changes in the Twitter/X platform and in the accessibility of Twitter/X data after late 2022.

A natural concern in using social media data is the presence of automated accounts or bots. These may distort measured sentiment by rapidly amplifying certain narratives or disseminating near-identical content. Existing bot-detection tools, such as Botometer X,[9] and heuristic filters based on follower-to-following ratios, offer imperfect solutions: they do not reliably cover all users and often lack clear thresholds for classification. Moreover, it

[9] https://botometer.osome.iu.edu/

is not obvious that bot content should be excluded a priori, given its potential influence on public discourse. We therefore take an agnostic approach and show, in the Online Appendix C, that a version of our index excluding duplicate tweets—many of which appeared bot-generated—tracks the baseline index closely. We retain the full sample in our primary analysis.

2.2 Sentiment analysis

To classify the sentiment expressed in tweets about the Federal Reserve, we experimented with a range of approaches, including lexicon-based models, supervised machine learning algorithms, and large language models. In our context, traditional lexicon-based tools—such as VADER—performed poorly, often classifying tweets in ways that contradicted human judgement. Supervised models trained on hand-labelled data also struggled to generalise across the diverse and noisy language of social media. In contrast, we found that Generative Artificial Intelligence (GenAI), and specifically ChatGPT, offered significantly higher accuracy and contextual reliability.

This performance difference is corroborated by a classification exercise in which two of the authors manually labelled a set of 4,000 tweets. We then compared the ability of six lexicon-based models, eleven supervised learning models, and two versions of ChatGPT (web and API) to replicate these human judgements. ChatGPT models clearly outperformed all other approaches, with lexicon-based methods performing especially poorly. The corresponding accuracy results are reported in Appendix Section A.3.

GenAI offers several advantages for sentiment classification in this setting. First, it is highly effective at interpreting context, enabling it to distinguish sentiment directed at the Federal Reserve from sentiment aimed at other topics mentioned in a tweet. Second, it can detect implicit sentiment and handle figurative language such as sarcasm or humour, which often confounds traditional models. Third, as a pre-trained model on diverse and extensive textual corpora, it is well suited to the noisy and unstructured language typical of Twitter/X. Further details on the validation exercise, prompt specifications, and robustness checks using alternative LLMs are reported in Online Appendix B.1.

2.2.1 Classification procedure

Our goal was to assign each of the tweets in our dataset to one of four mutually exclusive categories with respect to sentiment towards the Fed: {'positive', 'neutral', 'negative', 'unrelated'}. As a first step, we pre-classified some tweets using rules-based tagging, based on the presence of unambiguous terms or phrases. For instance, any tweet containing the hashtag #endthefed was classified as negative. This step also addressed some residual issues of homonymy. Table B.4 in the Online Appendix lists the specific phrases used for this rule-based classification.

The remaining tweets in our sample were classified using ChatGPT. We employed both the API-based ChatGPT 4o-mini and the full web interface for ChatGPT 4, with the former used for bulk processing and the latter reserved for tweets with high visibility or high repetition (i.e., appearing more than 100 times). In both cases, we used a consistent prompt format: "Classify each of the following tweets according to whether they have a positive, neutral, or negative sentiment toward the Fed, the Federal Reserve System, or its leadership (FOMC members). If a tweet is unrelated to the Fed, classify it as 'unrelated'. Consider each tweet independently." We set the model's temperature parameter to zero to ensure deterministic output.[10]

We have also explored several alternative specifications, including different LLMs, few-shot prompts with author-classified tweet examples, and alternative prompt structures. These choices have very limited effects on classification performance. Details are provided in the Online Appendix Section B.1.

[10] Processing this number of tweets via ChatGPT is an extremely time-consuming task. It takes the API version of ChatGPT 4o-mini around 1 hour to process 5,000 tweets. In our case, that equates to around 760 hours of processing time for the full sample.

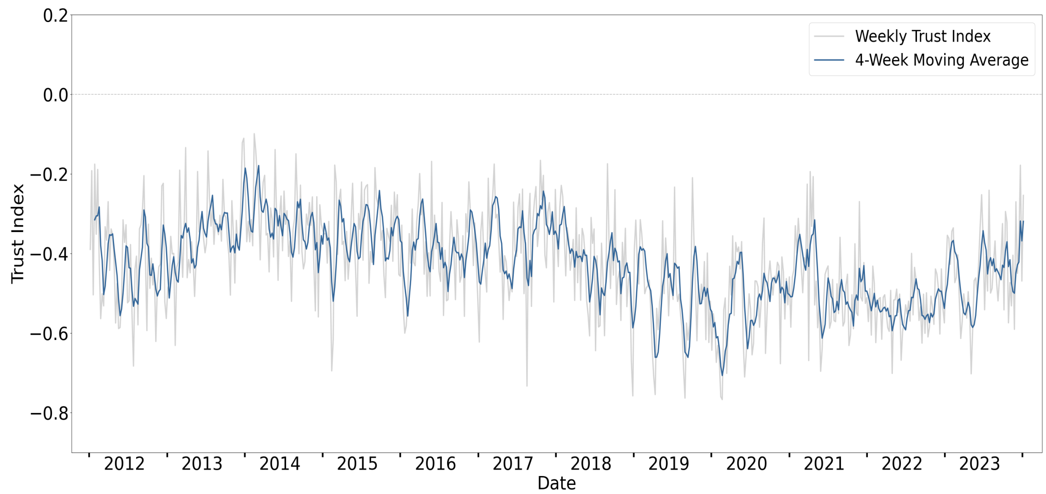

Figure 2: Index of trust in the Fed

Notes: This figure presents our index of trust in the Federal Reserve. The lightly-shaded grey line shows weekly values of the index, obtained by summing daily values Saturday-Friday. The bold blue line shows a 4-week moving average of the index. The y-axis shows the difference between the positive and negative share of tweets, measured in decimals.

This procedure resulted in the following sentiment distribution across our dataset of 3,859,454 tweets: 381,696 tweets (9.89%) were classified as expressing positive sentiment toward the Fed; 1,504,797 (38.99%) as neutral; 1,711,589 (44.35%) as negative; and 261,372 (6.77%) as unrelated. This classification provides the foundation for the construction of our daily trust index, which we describe in the next section.

3 The index of trust and its covariates

3.1 An index of trust in the Fed

Let N_t be the number of relevant tweets in day t, N_t^+ be the number of positive tweets, N_t^0 be the number of neutral tweets, and N_t^- be the number of negative tweets—that is, $N_t = N_t^+ + N_t^0 + N_t^-$. We define the trust index as the net share of positive tweets:

$$T_t = \frac{N_t^+ - N_t^-}{N_t} \tag{1}$$

Under this specification, $T_t \in [-1, 1]$. In the baseline approach, we weight tweets by engagement: tweet j receives weight $w_j = 1 + \text{likes}_j + \text{retweets}_j$. Thus, a tweet with α likes and β retweets receives weight $1 + \alpha + \beta$, while a tweet with no recorded engagement still receives weight equal to one. The weighted counts N_t^+, N_t^0, and N_t^- are then obtained by summing w_j across tweets in each sentiment category on day t. Figure 2 plots the baseline trust index for the Federal Reserve and its 4-week moving average. In the Online Appendix D, we examine whether this weighting scheme affects the results. We construct alternative versions of the trust index that exclude neutral tweets from the denominator and use the cumulative sum of trust. We also consider an unweighted specification, which removes likes and retweets from the weighting scheme by assigning equal weight to each tweet, as well as log-likes weights and combined log-likes and log-retweets weights. These variants have very similar properties to the baseline index, suggesting that the results are not driven by the denominator choice, the cumulative aggregation method, or a small number of highly engaged tweets.

3.2 Validation of the trust index

We validate our social media trust index, aggregated at a weekly frequency, by comparing its extreme values with the dominant themes in tweets during those weeks. Table 3 lists the ten weeks with the lowest and highest index values; for richer context, Online Appendix C.3 visualises their tweet-term frequencies as word clouds.

The ten lowest index values occur predominantly in the latter part of our sample, during periods of acute economic and political strain. In early February 2020 (the weeks of 8 February and 15 February), the outbreak of COVID-19 triggered unprecedented market volatility, pushing the index to

Table 3: Ten largest changes in trust over the sample

Week beginning	Rank	Index (T)	Key event
Panel A: Weeks with the lowest values of the index			
12 Aug 2017	6	-0.73	VC Fischer warns against unwinding Dodd-Frank Act.
22 Dec 2018	4	-0.76	Trump tweets about the Fed.
30 Mar 2019	8	-0.71	Trump recommends Herman Cain for Fed seat.
06 Apr 2019	5	-0.75	Herman Cain Fed nomination faces opposition.
24 Aug 2019	2	-0.76	Trump tweets about the Fed.
08 Feb 2020	3	-0.76	US stock market falls sharply during COVID-19 crisis.
15 Feb 2020	1	-0.77	US stock market falls sharply during COVID-19 crisis.
30 May 2020	7	-0.71	Publication of the June MPR detailing COVID-19 impact.
23 Apr 2022	10	-0.70	US GDP contracts.
29 Apr 2023	9	-0.70	First Republic Bank seizure.
Panel B: Weeks with the highest values of the index			
23 Mar 2013	7	-0.13	S&P 500 nears all-time high.
06 Jul 2013	4	-0.14	Bernanke reassures markets after taper concerns.
14 Dec 2013	8	-0.12	FOMC announces tapering of QE3.
21 Dec 2013	9	-0.11	Markets digest implications of QE3 tapering.
08 Feb 2014	10	-0.09	Yellen becomes Fed Chair.
15 Feb 2014	5	-0.14	Fed continues QE taper debate.
17 May 2014	6	-0.14	U.S. stocks rebound after Yellen dovish remarks.
23 Aug 2014	3	-0.15	Yellen cautious on rate hikes at Jackson Hole.
21 May 2016	1	-0.17	Yellen and Bullard speeches on policy rate normalisation.
14 Oct 2017	2	-0.17	Yellen signals gradual normalization.

its lowest level. Shortly thereafter, the week of 30 May 2020—when the June Monetary Policy Report underscored the pandemic's economic impact—also registered a steep decline. A similar trough appears in late April 2022, as the unexpected contraction in US GDP and the Fed's signal of tighter monetary policy reignited stagflation fears. In the final part of the sample, the seizure of First Republic Bank in the week of 29 April 2023 intensified financial instability concerns and further eroded trust. Political interventions compound these economic shocks: on 16 August 2017, Vice Chair Fischer warned against undoing the Dodd-Frank Act amid President Trump's repeal push; in the weeks of 22 December 2018 and 24 August 2019, sharply critical tweets by President Trump strained perceptions of central bank independence; and in early April 2019, his announcement of Herman Cain's nomination to the Fed Board added further uncertainty.

By contrast, the highest index values are associated with episodes of policy clarity, reassurance, and market optimism. In March 2013, the S&P 500 approached an all-time high, reflecting a broadly positive market environment. Although concerns over QE tapering intensified in mid-2013, Bernanke's reassurance in the week of 6 July helped calm markets after the initial taper-related volatility. By mid-December 2013 the smooth implementation of QE3 tapering signaled robust confidence in Fed decisions. Leadership transitions and affirmations of stability further buoyed trust: Janet Yellen's confirmation as Fed Chair in the week of 8 February 2014 coincided with a discernible uptick in sentiment, while the following week continued to feature discussion of QE tapering within a relatively constructive policy narrative. In May and August 2014, dovish and cautious remarks by Yellen, including her Jackson Hole communication, further supported market confidence. Speeches by Yellen and Bullard on policy normalisation in May 2016, and Yellen's emphasis on gradual normalisation in October 2017, provided clear forward-guidance signals that underpinned market optimism.

The alignment of the index's extremes with these salient events confirms the validity of our trust metric as a dynamic gauge of expressed public confidence in the Federal Reserve. This is true wether we look at our baseline trust measure or a measure of trust, which only includes experts, which we construct based on users who repeatedly post around FOMC meetings, as suggested by Ehrmann and Wabitsch (2022).[11]

3.3 What drives trust in Federal Reserve?

We augment the preceding analysis by systematically investigating a broad set of candidate drivers of our trust index. In particular, we examine the following set of potential determinants:

Macroeconomic and financial indicators. We include: the 5-year breakeven inflation rate ($E\pi^{5yr}$) and the 5-year 5-year forward inflation rate ($E\pi^{5yr5yr}$), measures of medium-term and longer-term inflation expectations respectively; the effective federal funds rate (DFF) to assess changes in the stance of monetary policy; the 30-day fed funds futures (FFF30) to capture market expectations of monetary policy actions in the near term; the 10-year Treasury yield, a key indicator of long-term interest rates and gauge of the economic outlook; and an index of the strength of the U.S. dollar vis-à-vis other currencies. We also include the VIX index, which captures market expectations for near-term volatility in the S&P 500; the Chicago Fed's National Financial Conditions Index, which provides a comprehensive measure of U.S. financial conditions; daily percentage log-returns on the S&P 500 Index and the NASDAQ index; and the daily differences of the log of the West Texas Intermediate crude oil price, as a salient measure of changes in future inflation pressures. Lastly, we include the Aruoba-Diebold-Scotti Business Conditions Index (ADS), which provides a daily real-time measure of business conditions in the U.S. economy (Aruoba et al., 2009).

[11]See Online Appendix C.2 for the series and further details.

Sentiment and policy uncertainty. We consider the relationship between our trust index and the Shapiro-Sudhof-Wilson Daily News Sentiment metric (SSW), derived from daily news reports to account for the influence of media sentiment on public perception (Shapiro et al., 2022), and the impact of the Baker-Bloom-Davis Economic Policy Uncertainty Index (EPU), which captures the level of uncertainty in economic and policy-related news (Baker et al., 2016).

Monetary policy actions and public appearances. We introduce dummy variables to capture the influence of key monetary policy events, including the publication of the semi-annual Monetary Policy Report and the Fed Chair's speech at the Jackson Hole Symposium. We also exploit existing measures of monetary policy shocks to capture the effects of unexpected shifts in policy on our trust index. In particular we consider the measure provided by Jarociński and Karadi (2020), which decomposes high frequency unexpected shifts in interest rates in a window around the monetary policy press conference into a pure monetary policy shock series and an information effect relating to the central bank's assessment of the economic outlook. We consider both series in our regression analysis. During our sample period, the Fed experienced two transitions in chairmanship—from Ben Bernanke to Janet Yellen in 2014, and from Janet Yellen to Jerome Powell in 2018. Each transition involves a set of events that could potentially influence trust in the Fed, including the nominee's announcement and Senate confirmation. These milestones are significant, as they signal shifts in leadership and policy direction, drawing substantial attention from the public and markets. We consider the impact of these events separately in our empirical analysis.

President Trump's tweets. During his first term in office (January 2017 to January 2021), President Trump frequently posted tweets criticising the Federal Reserve and its leadership. It is plausible that these tweets shaped public perceptions of the institution, until his banishment from Twitter/X in January 2021 during the final days of his term. We introduce dummies for the days in which President Trump tweeted, distinguishing tweets that explicitly pose a threat to central bank independence from his other tweets that relate to the Fed, following the classification in Bianchi et al. (2023).

Other events affecting trust. Trust in an institution and its leadership also depends on the public's view of the officials' integrity. In the later part of our sample, there were several reports of high-profile Federal Reserve officials having engaged in investment trades ahead of important policy announcements. These news plausibly raised public scepticism about the integrity of the Fed decision-making process and could therefore have dented trust in the Federal Reserve.Similarly, trust could be enhanced if a policymaker is deemed to have had a particularly positive impact on society (unrelated to policy actions). An example of this is the #popyourcollar tribute to Janet Yellen hailing her as a role model for women in STEM, launched on Twitter/X ahead of her last days in office. Table 4 lists the exogenous events we consider in our empirical analysis, while Figure 3 plots daily tweet counts by sentiment category and shows that these events generate engagement of the sign posited in the table. Moreover, the figures in Online Appendix C.4, highlight that each of these events is followed by a marked increase in Twitter/X activity specifically relating the individuals concerned in the immediate aftermath of the news breaking. We will use the shorthand "narrative proxy" for these events, as we will use them in Section 4 to construct a narrative proxy based on the shock-sign restrictions they provide.

Table 4: Exogenous events affecting trust

Date	**Sign**	**Event Details**

Continued on next page

Table 4: Exogenous events affecting trust (Continued)

04/04/2017	−	Richmond Fed President Jeffrey Lacker resigns abruptly after admitting to leaking information about confidential Fed deliberations in 2012. Published online in the Washington Post at 1.44pm ET and widely reported and shared thereafter by NYT, NPR, CNBC, CNN, WaPo, WSJ, Marketwatch and other outlets.
31/01/2018	+	#popyourcollar tribute to Janet Yellen launched by the New York Fed ahead of her last day as Fed Chair (3 February 2018).
07/09/2021	−	The Wall Street Journal reported that Kaplan was active buyer and seller of stocks in the course of 2020.
08/09/2021	−	Bloomberg reports that Rosengren made multiple purchases and sales in REITs and other securities in 2020. The Boston Fed said the trades were not done via a blind trust and were "consistent" with ethics rules.
09/09/2021	−	In response to reports that Rosengren and Kaplan had traded REITs and other stocks released the day before by the WSJ, but not picked up by other media nor the Twitterverse, the Federal Reserve Banks of Boston and Dallas simultaneously issue statements on 9 September noting that their financial transactions complied with the Fed's ethics rules. There is a big reaction after the statement. Published online in Dallas News at 2.01 pm ET and widely reported and shared thereafter by WSJ, NYT, CNBC, Reuters, etc.
27/09/2021	−	Before the market opens, Boston Fed President Rosengren announces he will pull forward his retirement by nine months. After the market closes, Dallas Fed President Kaplan says he will step down on October 8, 2021, mentioning his financial dealings. News appears on CNBC at 8.30am and it is widely reported and shared by PBS, NPR, Bloomberg, WaPo, Dallas News, etc.

Continued on next page

Table 4: Exogenous events affecting trust (Continued)

02/10/2021	–	Bloomberg reports on 1 October 2021 at 8.41pm ET that Fed Vice Chairman Clarida traded between $1 million and $5 million out of a Pimco bond fund and into two stock funds. The trade happened on February 27, 2020, a day before Powell issued a rare, impromptu statement reiterating the Fed's support of the economy. Most subsequent coverage by usual media outlets and tweets are mainly from 02/10/2021.
10/01/2022	–	Warren sends a letter to Jerome Powell, asking for more details on trades made by senior Fed officials. Later in the afternoon, Clarida sends a letter to President Joe Biden announcing that he will be stepping down on Jan. 14, instead of the Jan. 31 day that his term was originally set to expire. Press release at 4pm ET and is subequently reported that afternoon in all major outlets – CNBC, WaPo, NYT, NYPo.
14/10/2022	–	Raphael Bostic reported that he had failed to previously disclose financial transactions on his official central bank forms. Release at 3pm ET by the Washington Post, followed by CNBC, Rueters, Bloomberg and others.
20/10/2022	–	The New York Times revealed that James Bullard, president of the Federal Reserve Bank of St. Louis, had spoken on the 14th of the previous week at an off-the-record, invitation-only forum held by Citigroup, and open to clients. Bloomberg picks up the news at 2.47pm ET and other media outlets follow.
15/06/2023	–	Bostic revealed transactions involving 19 exchange-traded funds made on May 2 of previous year. At that time, officials were in a so-called blackout period during which public communications were limited and trading prohibited. News available on Bloomberg at 1.27pm ET and echo by Reuters, WSJ, etc.

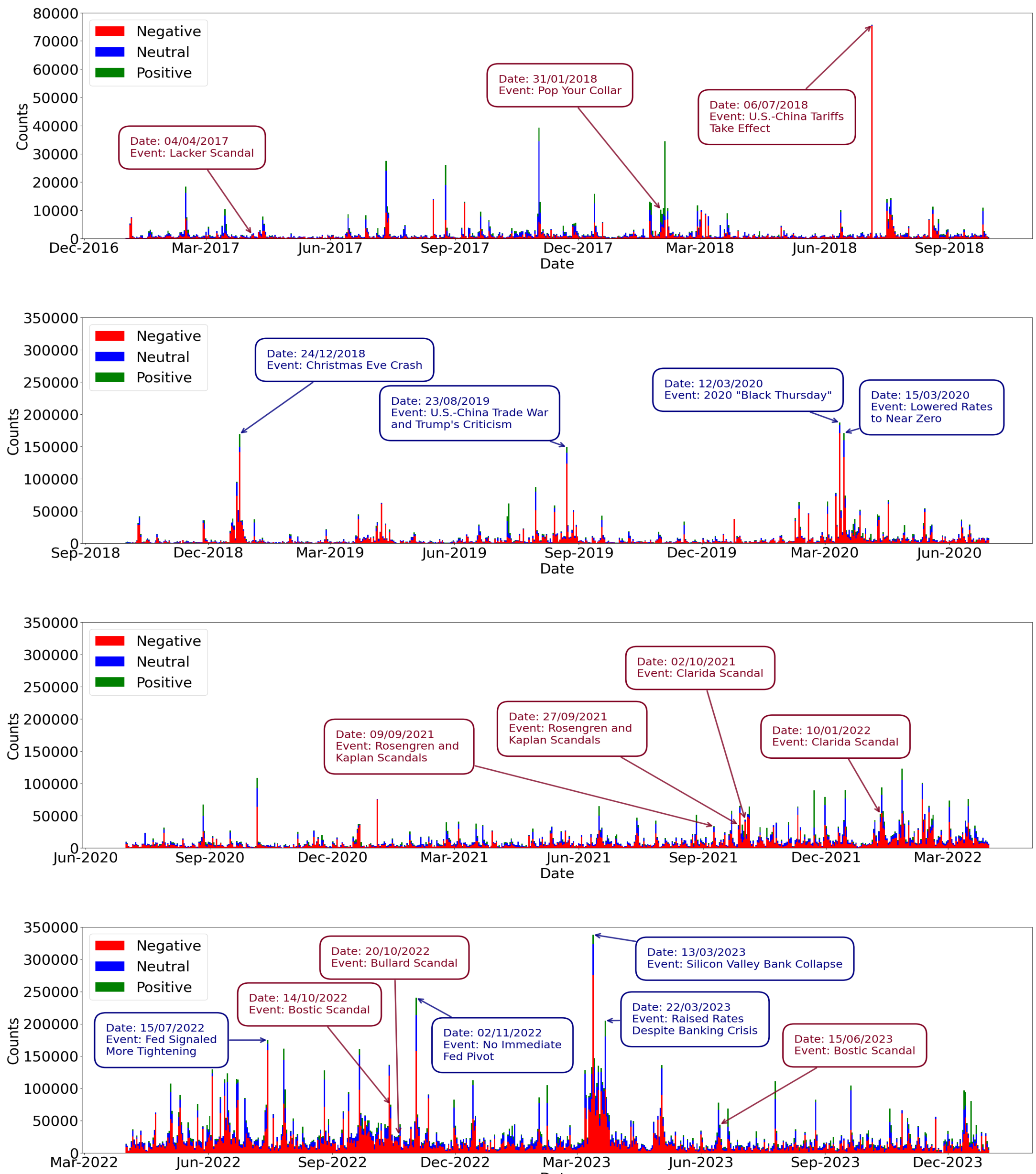

Figure 3: Posted, retweeted, and liked Tweets and exogenous events

Notes: This figure plots daily engagement-weighted tweet counts by sentiment category. Red bars denote negative tweets, blue bars denote neutral tweets, and green bars denote positive tweets. The four panels split the sample into January 2017–September 2018, October 2018–June 2020, July 2020–March 2022, and April 2022–December 2023. Annotated events identify selected dates associated with large movements in Fed-related tweet volume, including Federal Reserve official scandals, major monetary policy and market events, and episodes of heightened political or financial-market attention to the Fed.

3.4 Regression results

We regress the daily version of our raw index of trust in the Federal Reserve on the set of variables reported above, plus its own lags:

$$T_t = \alpha + \beta X_{t-1} + \gamma W_t + \sum_{i=1}^{5} \delta_i T_{t-i} + \varepsilon_t \tag{2}$$

where T_t is the trust index, X_{t-1} includes the set of macroeconomic and financial variables, and W_t incorporates dummy variables for the exogenous events listed in Table 4, plus other discrete event dates linked to monetary policy such as Jackson Hole speeches. We include 5 lags of the index to account for possible persistence in trust dynamics.

The macro-financial variables, which are standardized, are included with a one-day lag because the data we use are end-of-day values. For the other variables, we specify the timing of each variable as follows. If the event (speech, tweet or news release) happens before market closure, we consider it as information available at time t, while we treat it as information for the following day if the event happened after market closure. For example, FOMC press releases, used to construct the monetary policy surprises (Jarociński and Karadi, 2020), are issued around 2:00 pm. Since these events are publicly available before market closure, we use the event day (t) for these variables in our regressions. Regarding Fed chairmanship changes, we distinguish between the nomination announcement (around 3:00 pm) and Senate confirmation (around 5:00 pm, after market closure). Thus, in our regressions, nomination announcements are represented as Leadership nomination$_t$, while Senate confirmations are represented as Leadership confirmation$_{t-1}$. For tweets by President Trump, if a tweet was posted after the market closes, we treat it as information for the following day.[12] The results from this regression are reported in Table 5.

[12] Our results are robust when using the approach in Bianchi et al. (2023), which assigns tweets to their calendar posting date regardless of market hours.

Table 5: Regression results: daily trust index in the Fed

Macro-financial variables:					
$E_{t-1}\pi^{5yr}$	-0.0222***	-	-	-	-0.0234***
$E_{t-1}\pi^{5yr5yr}$	0.0028	-	-	-	0.0038
DFF_{t-1}	-0.1295**	-	-	-	-0.1123*
FFF30_{t-1}	0.0935*	-	-	-	0.0829
10-year Treasury yield_{t-1}	0.0402***	-	-	-	0.0345***
Financial conditions_{t-1}	-0.0017	-	-	-	-0.0023
$\text{S\&P500}_{t-1}$	-0.0155*	-	-	-	-0.0139
NASDAQ_{t-1}	0.0171*	-	-	-	0.0156*
VIX_{t-1}	-0.0142***	-	-	-	-0.0120**
US dollar_{t-1}	-0.0076	-	-	-	-0.0078
ADS_{t-1}	-0.0073*	-	-	-	-0.0073
Oil price_{t-1}	0.0059	-	-	-	0.0053
Sentiment and policy uncertainty:					
SSW_{t-1}	-	0.0175***	-	-	0.0056
EPU_{t-1}	-	-0.0037	-	-	0.0022
Monetary policy:					
JK monetary policy shock_t	-	-	-0.0033	-	-0.0046
JK monetary policy information_t	-	-	0.0079**	-	0.0068**
Monetary policy report publication_t	-	-	0.1562***	-	0.1588***
Jackson Hole speech_t	-	-	0.1553***	-	0.1668***
Leadership nomination_t	-	-	0.3244***	-	0.3404***
Leadership $\text{confirmation}_{t-1}$	-	-	0.4416***	-	0.3801**
Exogenous drivers of trust:					
Ethical scandals_t	-	-	-	-0.2945***	-0.2616***
Trump tweets Bianchi et al._t	-	-	-	-0.1643***	-0.1506***
Other Trump tweets about Fed_t	-	-	-	-0.0576	-0.0783*
Lags	Yes	Yes	Yes	Yes	Yes
α	-0.2755***	-0.2295***	-0.2166***	-0.2156***	-0.2816***
Adj. R^2	0.1364	0.1164	0.1191	0.1177	0.1522
N	4,378	4,378	4,378	4,378	4,378

Notes: The dependent variable is the daily trust index in the Fed. Continuous right-hand side variables have been standardised by subtracting the mean and dividing by the standard deviation. The regression includes 5 lags of the dependent variable. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels. Standard errors are heteroscedasticity robust (HC1).

We first regress the index on a set of macro-financial variables available at a daily frequency. Increases in break-even inflation expectations have a modest negative impact on the trust index: a one-standard-deviation rise in expected inflation ($E_{t-1}\pi^{5yr}$) reduces the index by around 2 percentage points. In contrast, longer-term inflation expectations are not statistically significant. A one-standard-deviation increase in the Fed Funds rate reduces the trust index by 11-13 percentage points. There is also some evidence that, holding constant the current Fed Funds rate and inflation expectations, expected hikes in the Fed Funds rate (increases in FFF30) tend to boost the index. One interpretation of this is that policy rate hikes tend to happen in positive economic environments. The yield on the 10-year Treasury operates in a similar manner, with higher yields boosting the trust index—consistent with higher long-term interest rates reflecting stronger growth prospects for the economy. NASDAQ stock market returns have a modest positive impact on Fed trust, and the financial market volatility, measured by the VIX index, has a negative impact. Finally, neither the exchange rate, nor the oil price have a statistically significant effect.[13]

Our trust index is systematically related to variables that are closely linked to Fed monetary policy. The publication of the biannual Monetary Policy Report, the announcement and confirmation of recent Fed chairs, and Jackson Hole speeches by the Fed Chair all have large and positive statistically significant effects. Also the Jarocinski-Karadi central bank information shock, which captures the information effects of the FOMC press statements following the release of the policy decisions, is significant in explaining movements in trust.

Finally, we note that the exogenous events listed in Table 4, defined as dummies on the date of the news release, have a large impact on the trust index. Our narrative proxy, which combines news of ethical breaches involving Fed officials and the #popyourcollar hashtag launched as a tribute to Chair Yellen (with reversed sign), has the largest measured impact on the index: taken together, these events reduce our index by 26-30% on the same day. We also find a very material impact from the set of President Trump's tweets associated with threats to the Fed's independence: these tweets generate a 15-16% point drop in the index. Overall, both events

[13] We also considered the prices of gold and bitcoin in this regression. These indicators might be considered barometers of investor beliefs about the Fed's ability to maintain price stability and a reflection of distrust in traditional financial and economic markets. Neither indicator has a statistically significant impact on our trust index.

appear to generate a substantial erosion of confidence in the Federal Reserve. The impact of other tweets by President Trump that reference the Fed is smaller and borderline significant.

The index co-moves with macroeconomic variables and sentiment, but captures something overall different, as we confirm in Online Appendix D. In it we compare the index to the Michigan survey of consumers and we show that the trust series orthogonalized with respect to the macro variables retains its characteristics and informativeness. Online Appendix D also reports a series of robustness checks – reconstructing the index using GPT-5 mini classifications, including Donald Trump's tweets, adding longer-term federal funds futures, considering monetary policy uncertainty, decomposing the index into its positive and negative components – which reconfirm the results of this section. We also assess the role of weighting by considering alternative tweet-level weighting schemes and estimating WLS regressions using the log of the daily engagement-based weight. We further examine the robustness of the results across sample choices by splitting the sample into the pre- and post-COVID periods and excluding days around FOMC announcements. Finally, we control for major platform-level changes in Twitter/X, including the Musk acquisition and the Trump ban and reinstatement periods. Across these exercises, reported in Online Appendix D, the results remain broadly consistent with our baseline findings.

4 The effects of shocks to trust in the Fed

We explore the macroeconomic and financial impact of shocks to trust by estimating a SVAR model and using a narrative identification approach based on shock-sign restrictions (Antolín-Díaz and Rubio-Ramírez, 2018). The reported ethical scandals that affected the Fed in 2017 and in 2021-2023, along with the emergence of the #popyourcollar Twitter/X trend launched on 31 January 2018 to honour Chair Yellen ahead of her stepping down as Fed

Chair, provide an ideal set of events that are relevant for the dynamics of the trust index but exogenous to the state of the economy.

4.1 Identification strategy

Our index of trust in the Fed is affected by movements in key macroeconomic and financial market indicators, as well as by policy actions and speeches by FOMC members, which reveal their views on the state of the economy and, at times, the future path of policy. It is also affected by ethics scandals that erode trust in the institution and its leadership.

Our identification strategy is based on information about alleged breaches of ethical guidelines involving FOMC members. These scandals, detailed in Table 4, can be grouped into two categories: insider trading allegations and information leak claims. In all cases, there are lags between the alleged unethical actions and the dates these events were reported in the news, ranging from 5 years to 6 days. So, while the alleged behaviour that caused the scandal is potentially endogenous to the state of the economy, the date on which these events are reported in the media is not. We assume that there is a negative shock to our trust index (indicated in the sign column in Table 4) on the day when the news of each scandal is reported. We also include a positive shock to our trust index coinciding with the #popyourcollar Twitter/X trend to honour the achievements and the legacy of Janet Yellen as she was stepping down as Fed chair in 2018. This event too is unrelated to macroeconomic conditions: it coincided with the end of Chair Yellen's mandate, but was launched a few days ahead of the end of her tenure.

Plagborg-Møller and Wolf (2021) and Giacomini et al. (2022) prove that it is possible to use information about the sign of a particular structural shock in specific periods by recasting it as a discrete-valued proxy for the structural shock. So we implement these shock sign restrictions with an indicator variable which is equal to -1 on the days indicated in Table 4 as having negative sign, i.e. the days in which the alleged ethical breaches are reported, equal to +1 on the day the #popyourcollar tribute was launched, and zero elsewhere.[14] We use this proxy as an instrument to identify impulse responses to a trust shock in a proxy-SVAR as discussed in Mertens and Ravn (2013) and Stock and Watson (2018).

The case for the validity of this instrument is very strong. It is possible that the alleged actions that contravened ethics guidelines were endogenous to the economy—for example, some trades allegedly happened just ahead of an extraordinary statement by Chair Powell in the height of the COVID-19 crisis. The dates at which the news break, by contrast, are governed by the journalistic process — filing dates, Freedom of Information Act (FOIA) timing, reporters' investigations — which has no mechanical link to the current stance of policy or the business cycle.

The results of the regressions in Section 3 confirm the instrument's relevance. We acknowledge the possibility of *looking-ahead bias* arising from the use of GPT4o-mini, since news of the alleged scandals plausibly appear in the model's pretraining data. We stress, however, that identification is threatened only if the classification error —not the correct labels, whose comovement with scandals is precisely the relevance channel we exploit— is correlated with the instrument. Because our model prompt provides only tweet text, with no temporal information, any leakage should operate through named officials and would therefore be spread across the full sample rather than concentrated at the scandal dates that drive the identification. We test this directly with a pre-scandal placebo: tweets mentioning the implicated officials in the periods before each scandal broke are not classified as more critical by the algorithm than by human evaluators, suggesting that the classifier is plausibly reading stance from the text rather than importing foreknowledge of the scandal.[15]

[14]The results are robust to using all the events or excluding the last two, which generated substantially less traffic than the others (see Figure 3).

[15]The figures in the Online Appendix C.4, which report the daily counts of tweets classified as negative, neutral, and positive around our "narrative proxy" events, further

4.2 The SVAR model

We specify a SVAR for the trust index and a set of daily variables that co-vary with it, selected among those found to be relevant in Section 3: These include the 5-year break-even inflation rate ($E\pi^{5yr}$), the ADS index, the NASDAQ index, the effective federal funds rate (DFF), 30-day federal funds futures (FFF30), the 10-year Treasury yield, the SSW index, and the VIX. All variables enter the model in levels, except for the NASDAQ index, which enters in log-levels.[16] Our baseline specification includes 28 daily lags, but our results are robust to various lag and model specifications.[17]

We estimate this model:

$$A(L)Y_t = u_t = B\varepsilon_t \tag{3}$$

where $A(L)$ is the autoregressive polynomial, Y_t is the vector of endogenous variables described above, u_t is a vector of white noise innovations with covariance matrix Σ, ε_t is the vector of white noise structural shocks, with the identity matrix as their covariance matrix, and B is matrix of their contemporaneous impacts that we aim to identify, at least in part. We estimate the model with Bayesian methods, imposing the standard Minnesota prior, which shrinks all variables towards a random walk. The tightness of the prior is selected optimally, maximising the marginal likelihood as in Giannone et al. (2015). More details on the estimation of the VAR are in Appendix E.

Assuming, without loss of generality, that the trust shock $\varepsilon_{1,t}$ is ordered first in the vector of structural shocks, the covariance between the instrument

support this argument.

[16]The Bayesian VAR, specified in levels or log-levels (NASDAQ index), is valid in the presence of unit roots and nests the cointegrated (VECM) representation (Sims et al., 1990). With Bayesian inference the posterior is a well-defined, smooth object regardless of the presence of unit-roots (Sims and Uhlig, 1991).

[17]We experimented with 7, 14 and 21 lags; with the log-differences for the NASDAQ index; with the cumulated version of the trust index; using the Secured Overnight Financing Rate instead of the fed funds rate; and we tried including the S&P500 index along with or instead of the NASDAQ index.

z_t and the reduced form VAR shocks is then $E(z_t u_t) = \alpha B_1$. This relationship can be used to identify B_1 up to a scaling constant by regressing $u_t^{2:N}$ on u_t^1 using z_t as an instrument. To pin down the constant, we set the size of the shock to imply a one standard deviation move to the trust index.

In our Bayesian estimation procedure, we obtain the credible set for the impulse responses by identifying a B matrix from each MCMC draw of the posterior distribution of the lag coefficients and the error covariance matrix of our BVAR, as done for example in Lenza and Slacalek (2024) and in Afonso et al. (2022). In principle, the Bayesian procedure should automatically be robust to weak instruments, as argued in Afonso et al. (2022) and Giannone et al. (2026), but we can still evaluate the robust F-statistic for each MCMC draw of the from the posterior distribution obtaining a distribution of F-statistics (see Figure E.14 in Appendix), the mean of which is 23.6. The draw of the series of identified shocks corresponding to the posterior mean of B can be seen in Figure A.1 in the paper's appendix.

Given the importance of Trump's tweets in moving the trust index, we also consider an indicator variable that has sign -1 on the dates of the Trump tweets identified by Bianchi et al. (2023) as threats to central bank independence and zero elsewhere. For the most part, these tweets explicitly advocate a policy direction (a lowering of rates), so that the associated shock could carry information about the expected path of the funds rate. A revelation of individual misconduct carries no such directional signal: it perturbs public trust in the institution without mechanically repricing the expected policy path.[18] These differences explain why the resulting impulse responses do not coincide with those identified from directional political pressure.

For robustness, we report in the Online Appendix the impulse response from a SVAR estimated with a flat prior,[19] with the only difference that stock

[18]It is also possible that these tweets are a function of the monetary policy stance and the state of the macroeconomy—they are issued around the time of rate increases, a strengthening dollar, and trade tensions, etc...

[19]Estimating a Bayesian VAR with a flat or uniformative prior is equivalent to estimating

prices (NASDAQ) enter the model as differences of the log-price. The impulse responses from this model, reported in the Online Appendix in Figure E.15, are very similar to those presented here.[20]

4.3 Responses to a trust shock

Figure 4 presents the dynamic effects of negative shocks to trust in the Federal Reserve, using our preferred identification approach based on the narrative proxy described in Table 4.

A negative one-standard-deviation shock to the trust in the Fed causes a sharp drop in the trust index. A loss of trust increases financial uncertainty, as measured by the VIX and dampens sentiment, here measured by the news sentiment index, SSW. This is in line with the predictions of Bursian and Faia (2018), whose model implies an increased price of risk as trust falls. This in turn leads to cautious behaviour and causes a slump in aggregate demand and investment. Our results confirm this mechanism, as we see the ADS index of business conditions and the NASDAQ both falling, the latter quite persistently.

The break-even inflation rate over the next 5 years increases significantly on impact, but this move is reabsorbed within approximately a month. This is also compatible with the results of Bursian and Faia (2018). In their model, as trust declines, the reduction in demand is accompanied by a steepening of the Phillips curve, caused by the increase in the risk premia, and consequently the monetary transmission mechanism becomes less effective. The agents perceive a worsening in the inflation-output trade-off and therefore inflation and inflation expectations increase in anticipation, despite the fall

it with maximum likelihood methods.

[20] We also report the impulse responses to a shock to trust identified with a Cholesky decomposition (Figure E.16) in which we assume that trust responds only to its own shock on impact. This is justified by the fact that the other variables in the VAR are measured as end-of-day values and so cannot be fully reflected in the trust index at time t. This shock picks up all the residual movement in the index that is not explained by lags of the model's variables.

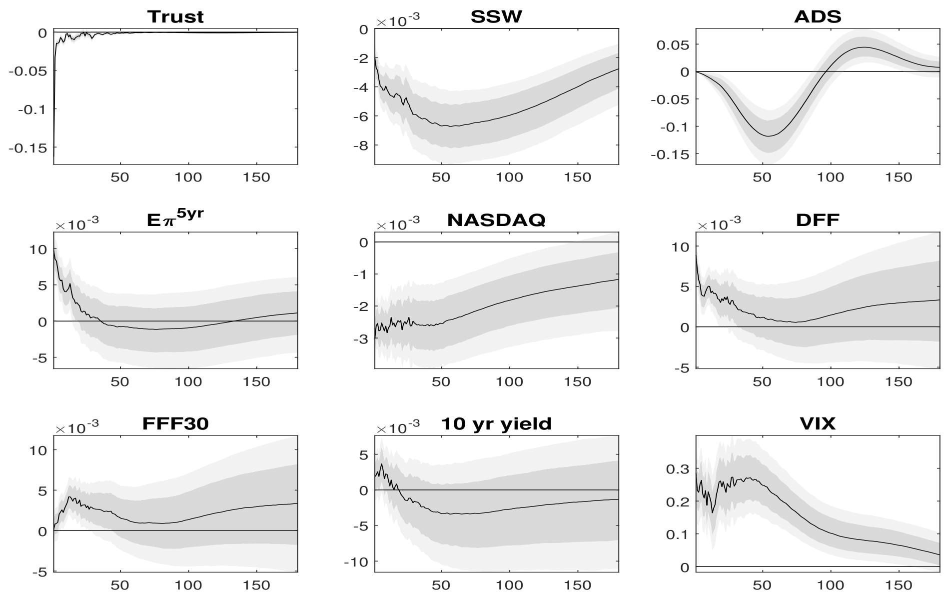

Figure 4: Narrative proxy

Notes: Impulse response functions to a trust shock identified using our narrative proxy (scandals + #popyourcollar) as an instrument. The dark and light shaded areas correspond to the 68% and the 90% high probability density (HPD) sets, respectively. The x-axis measures days.

in aggregate demand.

The fed funds rate (DFF) and the 30-day fed funds futures (FFF30) also increase briefly. In contrast, when using as an alternative instrument President Trump's tweets (Figure 5), the majority of which are calls to lower the interest rate, a significant and persistent decline in FFF30 is observed, followed by the daily effective Fed Funds rate, as the market accounts for the possibility that the policymakers cave in to political pressure. This difference is consistent with the different natures of the shocks picked up by these instruments: a shock to the perception of the policymakers' integrity vs. an explicit criticism of the central bank's policy, with a push for lowering rates.

The dynamic responses of the shocks to trust in the Fed identified using only President Trump's tweets are consistent with the evidence reported

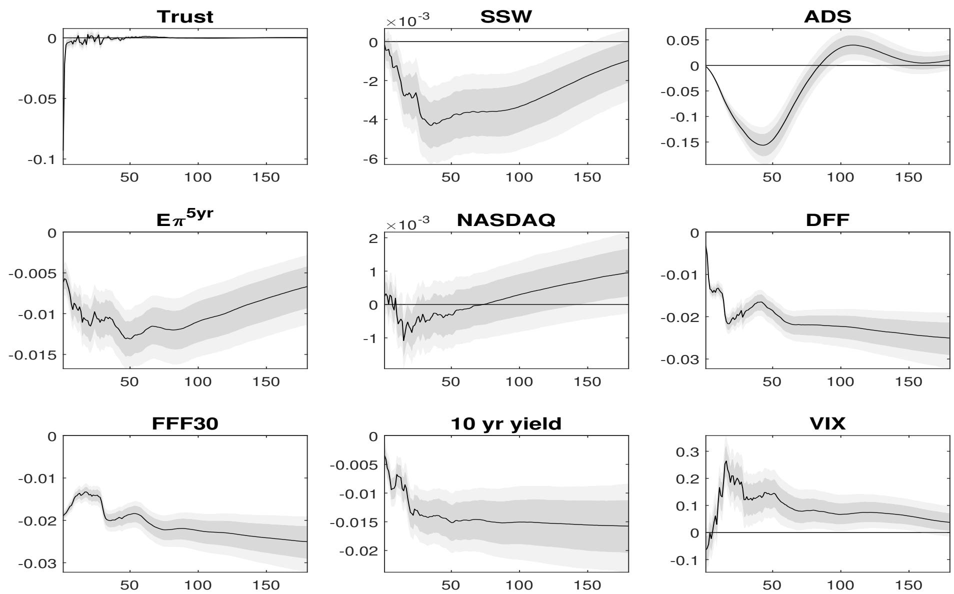

Figure 5: Trump tweets

Notes: Impulse response functions to a trust shock identified Trump tweets. The dark and light shaded areas correspond to the 68% and the 90% high probability density (HPD) sets, respectively. The x-axis measures days.

by Bianchi et al. (2023): President Trump's tweets affected market expectations about future monetary policy and inflation expectations. The daily news sentiment remains negative for almost 6 months, while the ADS index of business conditions falls but recovers somewhat more quickly. After an initial decline, the VIX Index increases quite persistently. The impact on the NASDAQ is positive, given the decrease in rates.

Table 6 reports the percentage of the variance of the variables in our VAR that is explained by the trust shock at different horizons (columns). Trust shocks play a non-negligible role in the dynamics of the stock market and can explain 10.9% of the variability in the NASDAQ after 3 months and up to 13.9% of moves in the VIX index at the 6 months horizon.

Table 6: Percentage of variance explained by trust shocks

	Days				
	0	15	30	90	180
Trust	64.6	64.2	63.0	62.3	61.8
News sentiment (SSW)	3.0	6.4	7.9	16.7	20.7
Business conditions (ADS)	1.7	1.1	2.3	11.0	11.1
$E\pi_t^{5y}$	9.6	3.8	2.2	0.9	0.7
NASDAQ	8.7	9.1	9.2	10.9	9.4
Fed Funds rate	8.1	5.2	3.2	1.1	0.7
Fed Funds Futures (FFF30)	0.0	1.8	2.3	0.8	0.6
10-year T-Bill yield	0.2	0.3	0.3	0.5	0.7
VIX	4.3	4.8	6.9	13.3	13.9

Notes: Each row reports the percentage of the variance of each variable that is explained by the trust shock at different horizons (columns).

5 Conclusion

Our research proposes a novel method to measure trust in central banking. By leveraging Generative AI to assess the sentiment in millions of tweets about the Federal Reserve and its leadership, our measure provides a more timely and comprehensive perspective of trust in the Fed than has hitherto been available. Our measure of trust, calculated as the daily difference between positive and negative sentiment tweets, correlates with key macroeconomic and financial aggregates and with relevant monetary policy events. Negative shocks to trust—identified using a narrative identification approach that maps the news of alleged financial scandals involving FOMC members into negative shocks to trust—reduce economic activity and increase inflation expectations, i.e., they operate akin to a supply shock.

What message should central bank policymakers take from our results? First, ensuring integrity and accountability of central bank officials is key to maintaining trust. The events that cause the largest erosion in trust in our sample are scandals that bring into question the integrity of key officials, so an appropriate formalised code of conduct, including restrictions on actively trading financial instruments, is paramount. Second, our results reinforce the notion that pro-active and transparent communication are important tools for bolstering the public's trust in the central bank. We find large positive effects on trust from the publication of the Monetary Policy Report and the annual Jackson Hole speech by the Fed Chair.

Finally, our analysis highlights the need for central banks to develop timely measures of the public's trust. Shocks to trust depress economic activity and may cause inflation expectations to increase. While these shocks have clearly not been of sufficient magnitude to influence U.S. business cycles materially over the recent past, there is no guarantee that this will continue to be the case in future. With the potential for greater political interference in U.S. monetary policy in future, understanding how central banks should act to retain public trust have become ever more urgent.

References

Afonso, G., D. Giannone, G. La Spada, and J. C. Williams (2022): "Scarce, Abundant, or Ample? A Time-Varying Model of the Reserve Demand Curve," Staff Report 1019, Federal Reserve Bank of New York, revised November 2025.

Angelico, C., J. Marcucci, M. Miccoli, and F. Quarta (2022): "Can we measure inflation expectations using Twitter?" *Journal of Econometrics*, 228, 259–277.

Antolín-Díaz, J. and J. F. Rubio-Ramírez (2018): "Narrative Sign Restrictions for SVARs," *American Economic Review*, 108, 2802–29.

Aruoba, B., F. X. Diebold, and C. Scotti (2009): "Real-Time Measurement of Business Conditions," *Journal of Business & Economic Statistics*, 27, 417–427.

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

Barro, R. J. and D. B. Gordon (1983): "Rules, discretion and reputation in a model of monetary policy," *Journal of Monetary Economics*, 12, 101–121.

Bianchi, F., R. Gómez-Cram, T. Kind, and H. Kung (2023): "Threats to central bank independence: High-frequency identification with twitter," *Journal of Monetary Economics*, 135, 37–54.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bursian, D. and E. Faia (2018): "Trust in the monetary authority," *Journal of Monetary Economics*, 98, 66–79.

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2020): "Trust in the central bank and inflation expectations," *International Journal of Central Banking*, 16, 1–37.

Coibion, O., Y. Gorodnichenko, and M. Weber (2022): "Monetary Policy Communications and Their Effects on Household Inflation Expectations," *Journal of Political Economy*, 130, 1537–1584.

Conti-Brown, P. and B. D. Feinstein (2020): "Twitter and the Federal Reserve: How the U.S. central bank is (and is not) surviving social media," *Brookings Economic Studies*.

Drechsel, T. (2026): "Political Pressure on the Fed," *The Review of Economic Studies*, rdag030.

Ehrmann, M. (2026): "Trust in Central Banks," *Journal of Economic Surveys*, 40, 492–504.

Ehrmann, M., M. Soudan, and L. Stracca (2013): "Explaining European Union Citizens' Trust in the European Central Bank in Normal and Crisis Times," *The Scandinavian Journal of Economics*, 115, 781–807.

Ehrmann, M. and A. Wabitsch (2022): "Central bank communication with non-experts – A road to nowhere?" *Journal of Monetary Economics*, 127, 69–85.

Fernández-Villaverde, J., P. Guerrón-Quintana, K. Kuester, and J. Rubio-Ramírez (2015): "Fiscal Volatility Shocks and Economic Activity," *American Economic Review*, 105, 3352–84.

Giacomini, R., T. Kitagawa, and M. Read (2022): "Narrative Restrictions and Proxies," *Journal of Business & Economic Statistics*, 40, 1415–1425.

Giannetti, M. and T. Y. Wang (2016): "Corporate scandals and household stock market participation," *The Journal of Finance*, 71, 2591–2636.

Giannone, D., M. Lenza, and G. Primiceri (2026): "Bayesian Inference in IV Regressions," Tech. rep., NBER Working Paper Series 34648.

Giannone, D., M. Lenza, and G. E. Primiceri (2015): "Prior Selection for Vector Autoregressions," *The Review of Economics and Statistics*, 97, 436–451.

Gorodnichenko, Y., T. Pham, and O. Talavera (2023): "The Voice of Monetary Policy," *American Economic Review*, 113, 548–84.

——— (2025): "Central bank communication on social media: What, to whom, and how?" *Journal of Econometrics*, 249, 105869.

Guiso, L., P. Sapienza, and L. Zingales (2008): "Trusting the Stock Market," *The Journal of Finance*, 63, 2557–2600.

Gurun, U. G., N. Stoffman, and S. E. Yonker (2018): "Trust Busting: The Effect of Fraud on Investor Behavior," *The Review of Financial Studies*, 31, 1341–1376.

Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Kydland, F. E. and E. C. Prescott (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, 473–491.

Lenza, M. and J. Slacalek (2024): "How does monetary policy affect income and wealth inequality? Evidence from quantitative easing in the euro area," *Journal of Applied Econometrics*, 39, 746–765.

Masciandaro, D., O. Peia, and D. Romelli (2024): "Central bank communication and social media: From silence to Twitter," *Journal of Economic Surveys*, 38, 365–388.

Mertens, K. and M. O. Ravn (2013): "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, 103, 1212–47.

OECD (2017): *Trust and Public Policy: How Better Governance Can Help Rebuild Public Trust*, OECD Public Governance Reviews, Paris: OECD Publishing.

Plagborg-Møller, M. and C. K. Wolf (2021): "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 89, 955–980.

Pástor, L. and P. Veronesi (2012): "Uncertainty about Government Policy and Stock Prices," *The Journal of Finance*, 67, 1219–1264.

——— (2013): "Political uncertainty and risk premia," *Journal of Financial Economics*, 110, 520–545.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2022): "Measuring news sentiment," *Journal of Econometrics*, 228, 221–243.

Sims, C. A., J. H. Stock, and M. W. Watson (1990): "Inference in Linear Time Series Models with some Unit Roots," *Econometrica*, 58, 113–144.

Sims, C. A. and H. Uhlig (1991): "Understanding Unit Rooters: A Helicopter Tour," *Econometrica*, 59, 1591–1599.

Stock, J. H. and M. W. Watson (2018): "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *The Economic Journal*, 128, 917–948.

A Appendix A

A.1 Details of the query used for downloading relevant tweets

The query we use in the Twitter/X API has three components. In particular, it returns a tweet if it:

EITHER contains the phrase "the fed".

OR contains the name of a Fed Chair or key Fed official *AND* one element from {"market", "economy", "business"} *OR* references to specific regional Federal Reserve Banks *OR* to the individual's title, e.g., "chair" or "president".

AND contains one of the trust-related words listed in the lexicon in Table 1 in the main text.

Further details can be found in the Online Appendix.

A.2 ChatGPT prompt

The full prompt we use is as follows:

"Please classify each of the following tweets according to whether they have a positive, neutral, or negative sentiment toward the Fed, the Federal Reserve System, or its leadership (FOMC members). If a tweet is unrelated to the Fed, please classify it as 'unrelated'. Consider each tweet independently.

Each tweet will be provided with a number. For each tweet, provide the sentiment classification as one of the following: [positive, neutral, negative, unrelated]. Format your response as follows: 'Tweet number', 'First 2 words of the tweet', 'Sentiment: [positive, neutral, negative, unrelated]'. Response format example: 'Tweet number 1' 'First 2 words' 'Sentiment: positive', ‘Tweet number 2’ ‘First 2 words’ ‘Sentiment: neutral’, ‘Tweet number 3’ ‘First 2 words’ ‘Sentiment: negative’, ‘Tweet number 4’ ‘First 2 words’ ‘Sentiment: unrelated’.

Please start from tweet number X and end with tweet number X+20.”

As explained in the main text, we set the “temperature” parameter—a model property that controls response determinacy—to 0, ensuring that the same prompts consistently produce identical answers.

A.3 Classification accuracy

This section compares the performance of lexicon-based methods, supervised machine learning models, and GenAI in categorising tweet sentiment. We report a comparison that excludes tweets manually classified as unrelated, focusing only on sentiment classification among Fed-related tweets. The Online Appendix reports the weighted F1 score, explainsthe complementary comparison including unrelated tweets, and extends the validation exercise by comparing GPT-4o-mini with several alternative LLMs.

Table A.1: Classification accuracy using manually coded tweets

Lexicon	**Accuracy**	**SML**	**Accuracy**	**GenAI**	**Accuracy**
LM	52.40%	RF	64.39%	GPT-4	75.71%
FS	50.34%	MNB	63.41%	GPT-4o-mini	72.67%
Opinion	47.31%	BNB	60.49%		
AFINN	46.52%	MLP	60.00%		
VADER	42.90%	LR	59.02%		
Harvard-IV	42.70%	Ridge	57.07%		
		GNB	54.15%		
		DT	54.15%		
		Linear SVC	52.20%		
		LS	47.32%		
		ET	43.90%		

Notes: Full lexicon and supervised machine learning (SML) model names, together with implementation details, are provided in the Online Appendix.

A.4 Identified trust shocks

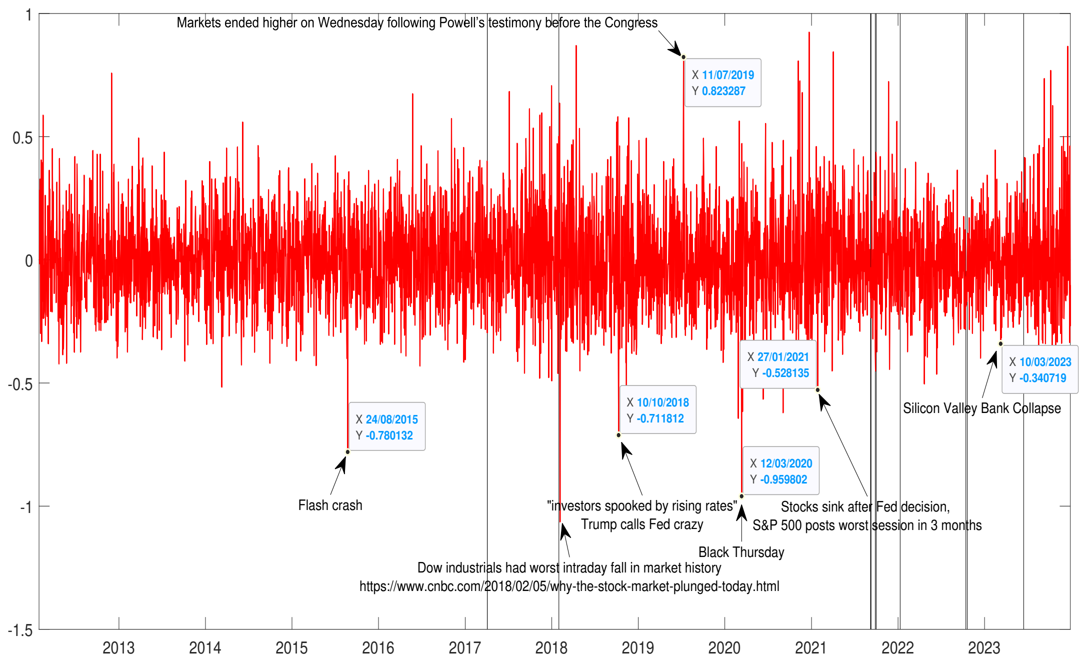

Figure A.1: Trust shocks

Notes: The red line portrays the trust shocks identified using the news of the alleged ethical scandals (and the #popyourcollar trend) as an instrumental variable. The vertical black lines correspond to the dates the news of the scandals were reported.