The Economic Research Guardian – Vol. 16(2)2026
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 16(2): 233-259

Central Bank Digital Currencies and Financial Stability Communication: Evidence from BIS Central Bank Speeches, 2016-2025

Bogdan Cristian Buzuriu*
Doctoral School, Faculty of Economics and Business Administration, West University of Timisoara, Timisoara, Romania & Management Department, Politehnica University of Timisoara, Timisoara, Romania
E-mail: bogdan.buzuriu80@e-uvt.ro

Bianca Talpas
Faculty of Management in Production and Transportation, Politehnica University of Timisoara, Timisoara, Romania
E-mail: bianca.talpas@upt.ro

Abstract
This paper examines how central banks frame the financial stability implications of central bank digital currencies (CBDCs) in their public communication. The existing literature has largely approached CBDC-related financial stability through theoretical models, design assessments and policy simulations. Although recent studies have begun to examine CBDC-related central bank communication, less attention has been paid to the separate measurement of financial-stability risk and benefit channels within this communication. To address this gap, the paper constructs a corpus of CBDC-related speeches from the BIS (Bank for International Settlements) central bank speeches database for the period 2016-2025 and develops two CBDC financial stability communication indexes, separating risk-oriented communication, captured by the Financial Stability Risks Index (FSRI), from benefit-oriented communication, captured by the Financial Stability Benefits Index (FSBI). Using an IMF-channel coverage approach, the results show a gradual change in the framing of CBDC communication. In the early phase of the debate, CBDC speeches were slightly more risk-oriented. From 2019 onward, however, benefit-oriented framing becomes more prominent. The first-order autoregressive (AR(1)) estimates additionally reveal strong year-to-year persistence in FSRI, whereas the lagged FSBI coefficient is not statistically significant. Complementing the index-based analysis, cluster analysis shows that CBDC speeches are organized around recurrent combinations of risk and benefit channels rather than isolated themes. Overall, the findings suggest that CBDC financial-stability communication has evolved toward a more selective and design-oriented narrative, in which risk-related concerns remain visible but are framed through safeguards, payment-system resilience and the role of public money.

Keywords: CBDC, Financial stability, Central bank communication, Text analysis

JEL classification: E42, E44, E58, G21, G28

Received: 27 July 2026; Received in revised form: 21 August 2026; Accepted: 25 August 2026

* Corresponding author.

1. Introduction

Central bank digital currencies have moved from the margins of monetary innovation to the centre of central banking debates. What began as a largely exploratory discussion about digitalisation, declining cash use and private digital money has become a policy agenda involving research programmes, technical experiments, pilot projects and public consultations across a wide range of jurisdictions. The CBDC debate now reaches well beyond payment technology. It touches monetary sovereignty, the future role of public money, financial inclusion, competition in payment services, bank intermediation, privacy, operational resilience and financial stability (see for example, Wronka, 2023).

The relationship between CBDCs and financial stability is not straightforward. CBDC could strengthen stability if it preserves access to central bank money in an increasingly digital economy, improves the resilience of payment systems, enhances competition and supports public trust in the monetary system. At the same time, it could generate new vulnerabilities if households and firms shift deposits away from commercial banks, if banks face higher funding costs, if credit supply is affected, or if conversion into CBDC accelerates withdrawals during periods of stress. These competing effects explain why CBDC design has become a central issue in policy debates.

Most of the literature evaluates these questions through models, simulations or institutional design analysis. This paper takes a different route. Rather than asking whether CBDCs have already affected financial stability outcomes, it asks how central banks communicate the financial stability implications of CBDCs. This distinction matters because central bank communication is not merely descriptive. Speeches signal priorities, shape expectations and reveal how policymakers understand emerging trade-offs. If central banks see CBDCs primarily as a source of instability, public communication should increasingly emphasize deposit substitution, bank funding pressure, liquidity stress and run-risk dynamics. If, instead, they view CBDC-related risks as manageable through design, one would expect a stronger emphasis on safeguards, payment-system resilience and the role of public money.

Therefore, the present paper addresses the following research question: How do central banks frame the financial stability implications of CBDCs in public speeches, and how has this framing changed over time? To answer these questions, the paper builds a CBDC communication corpus from BIS-listed central bank speeches between 2016 and 2025. CBDC-related speeches are identified through a broad dictionary that includes both generic terms, such as "CBDC" and "central bank digital currency", and project-specific expressions, such as "digital euro", "digital pound", "e-CNY", "digital yuan" and "public digital money" (for a detailed description of dictionary structure please refer to Appendix A). The empirical analysis then constructs two CBDC financial stability communication indices that separately capture risk-oriented and benefit-oriented channel coverage. The risk component is measured through the Financial Stability Risks Index (FSRI), while the benefit component is measured through the Financial Stability Benefits Index (FSBI). Both indices are constructed as speech-level channel-coverage measures. Each of the six predefined risk or benefit channels is coded as one when at least one associated term appears in a speech and zero otherwise, and the corresponding index is calculated as the proportion of activated channels. Annual values are obtained by averaging the speech-level scores for each year. To see if risks dominate benefits or vice-versa, a net index (NetCFSCI) is calculated as FSBI minus FSRI and indicates whether benefit-oriented or risk-oriented financial-stability channels are more prominent in a given year.

The main finding is that CBDC communication changes over time. In 2016-2018, the net index is slightly negative, indicating that risk-oriented channels marginally dominate benefit-oriented

channels. From 2019 onward, NetCFSCI becomes positive. This shift suggests that, as CBDC projects mature, central banks increasingly present CBDCs as instruments that may support financial stability, provided that appropriate safeguards are embedded in their design. The risk component declines over the period, while the benefit component remains relatively stable. The evidence therefore points less to a disappearance of risk concerns than to a reframing of those concerns within a broader narrative of resilience, public money and institutional design.

To identify whether the evolution of CBDC-related financial-stability communication follows a systematic temporal pattern, the analysis estimates simple annual trend regressions for the two main components of the index, FSRI and FSBI. This approach makes it possible to assess whether the observed changes in risk-oriented and benefit-oriented communication are merely descriptive fluctuations or reflect statistically significant movements over time. The regression results based on a temporal trend show that both components decline over the 2016-2025 period, but the decrease is considerably stronger for FSRI than for FSBI. This indicates that central-bank communication does not abandon benefit-oriented narratives, but progressively reduces the breadth of risk-channel coverage as CBDC projects move from conceptual discussion toward design, experimentation and institutional calibration. Further, the first-order autoregressive (AR(1)) estimates additionally reveal strong year-to-year persistence in FSRI, whereas the lagged FSBI coefficient is not statistically significant.

In addition to the temporal analysis, the study examines whether CBDC speeches are organized around recurring combinations of risk and benefit channels. For this purpose, a cluster analysis is conducted on the same corpus of 737 CBDC-related speeches, comparing k-means clustering with Density-Based Spatial Clustering of Applications with Noise (DBSCAN). K-means is used as the main typological method because it provides a parsimonious three-cluster structure for risks, benefits and topic profiles, while DBSCAN is used as a density-based robustness check that identifies recurrent exact channel configurations. The results show that CBDC communication is not distributed randomly across isolated themes, but is structured around stable bundles of bank-intermediation, systemic-stability, payment-resilience, public-money, safeguards and governance-related channels.

The paper contributes to the literature in three ways. First, to the best of our knowledge, it provides the first systematic and separate measurement of the breadth of financial-stability risk and benefit channels in central-bank CBDC communication. Unlike existing studies that primarily examine general sentiment, policy stance, thematic salience or the relationship between communication and CBDC implementation, the paper constructs distinct indices for risk-oriented and benefit-oriented channel coverage. This distinction makes it possible to examine not only whether financial stability is discussed, but also how central banks publicly balance potential vulnerabilities against stabilizing benefits and design safeguards as the CBDC debate evolves. Second, the paper provides empirical evidence on the evolution of these communication channels between 2016 and 2025. Third, it complements theoretical and policy research by showing how central banks frame CBDC-related financial-stability trade-offs in their own public communication.

The remainder of the article is organized as follows: Section 2 reviews the literature on CBDC communication, financial stability and central-bank narratives; Section 3 presents the data and describes the construction of the final corpus of CBDC-related speeches; Section 4 introduces the communication-index framework, estimates a simple annual trend model to assess whether the observed changes are statistically systematic, provides additional validation and robustness evidence; Section 5 extends the analysis through clustering methods and Section 6 concludes the article.

2. Review of the literature

The literature on CBDC and financial stability has expanded rapidly (Stoica and Dornean, 2025). Early policy work focused on the principles that should guide CBDC design and on the need for CBDCs to coexist with cash and commercial bank money. The BIS (2020) foundational principles stress that a CBDC should not compromise monetary or financial stability. Subsequent BIS work (2021) focuses more directly on financial stability implications, including bank disintermediation, liquidity pressure and potential run dynamics.

Several theoretical contributions examine how CBDCs may affect commercial banks. Bindseil (2020) proposes tiered remuneration as a way of limiting large-scale CBDC holdings and reducing disintermediation risk. Kumhof and Noone (2021) discuss design principles intended to preserve financial stability, including restrictions on abrupt conversion from bank deposits into CBDC. Adalid et al. (2022) analyse possible implications for bank intermediation in the euro area. More recently, Copple et al. (2025) show that holding limits on sterling-denominated systemic stablecoins and a potential digital pound can mitigate bank disintermediation, liquidity pressures and increased demand for central bank lending during periods of stress, although excessively restrictive limits may reduce the usability and innovation benefits of digital money. Other studies (i.e. Chen & Filippin, 2026) consider portfolio choice, deposit competition, bank market power and the interaction between CBDC demand and bank funding structures. In the same line, Katsurada (2026) shows that the Chinese digital yuan has moved from a cash-like CBDC model toward a deposit-based design. This shift is important for the CBDC financial-stability debate because it reframes the e-CNY as a commercial-bank liability that may bear interest, be covered by deposit insurance and support credit creation. The case illustrates how CBDC design can be adjusted to reduce deposit-substitution and bank-disintermediation risks, while also serving broader objectives related to cross-border payments and internationalization.

Close to our paper, Gross et al. (2025) provide a comprehensive IMF framework for evaluating the financial stability implications of CBDC. Their approach is particularly relevant for this paper because it treats financial stability as a set of channels rather than as a single outcome. CBDCs may affect liability structures, bank funding, credit supply, fee income, liquidity risk, run dynamics, information flows and payment-system resilience. The same instrument may therefore generate risks and benefits depending on its design, adoption pattern and institutional environment.

A related strand of research examines central bank communication on financial stability. Born et al. (2014) show that financial stability reports and central bank speeches can affect market returns, particularly during periods of financial stress. Using a dictionary-based textual approach, Correa et al. (2021) further demonstrate that the sentiment conveyed in financial stability reports reflects developments in the financial cycle and contains information about the probability of banking crises. Accordingly, speeches, reports, and policy statements not only shape expectations but also reveal how central banks interpret evolving economic and financial risks.

Fell et al. (2026) provide a recent methodological benchmark by comparing dictionary-based, FinBERT and prompt-based generative AI approaches across 43 issues of the ECB Financial Stability Review published between 2004 and 2025. They find broad co-movement across the resulting sentiment indicators, suggesting that transparent dictionary-based methods can capture systematic changes in financial stability communication. However, they also emphasize that dictionaries remain context-blind, whereas GPT-based methods can isolate assessment-bearing sentences and capture more nuanced shifts in the direction and temporal orientation of risk assessments. The authors therefore recommend treating dictionary-based and AI-based methods

as complementary tools rather than substitutes for expert judgement and conventional financial stability analysis.

Building on this methodological approach, the present study examines CBDC communication through a structured dictionary-based framework, complemented by topic modelling and cluster analysis. Rather than assessing whether CBDCs have caused financial instability, it investigates how central banks publicly frame the associated financial stability risks, benefits and trade-offs.

This communication perspective is useful because most CBDC projects remain in research, development or pilot phases. Since realized macrofinancial outcomes are still limited, speeches provide an early empirical window into policy thinking. They reveal whether the public narrative is dominated by risk, by benefits, or by a design-oriented attempt to reconcile the two.

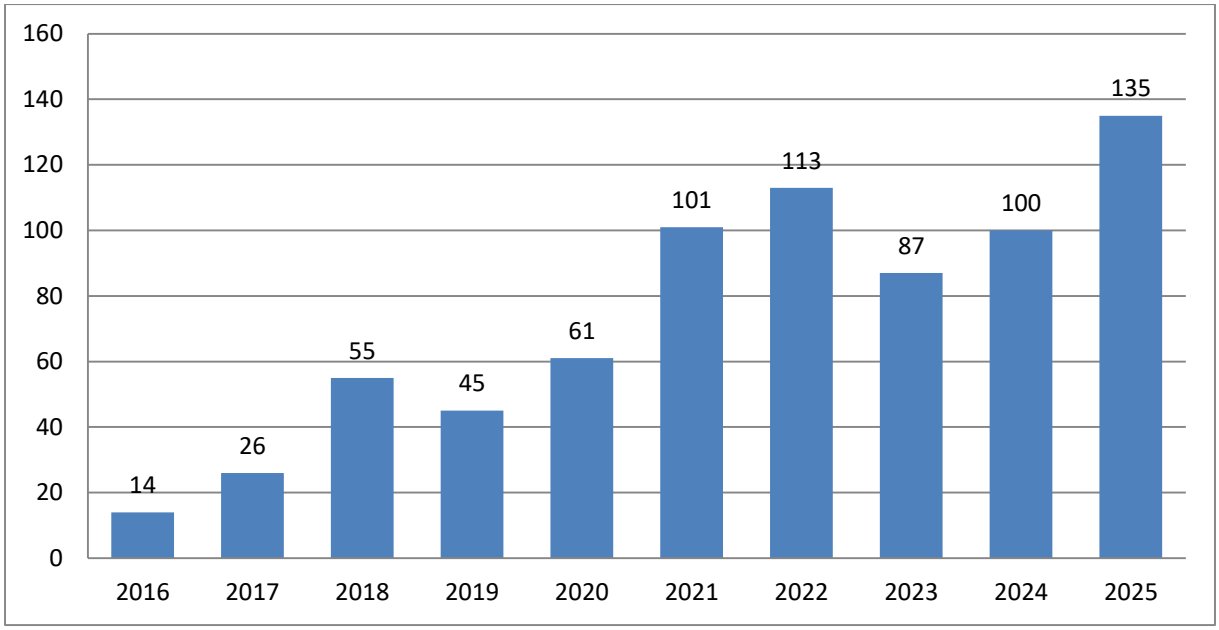
3. Data

The empirical corpus is drawn from the BIS central bank speeches database, which provides a broad repository of speeches and articles by central-bank officials. From this wider source, the analysis selects CBDC-related speeches published between 2016 and 2025. This period captures the transition from early conceptual discussion to more concrete CBDC projects, including the rise of major initiatives such as the digital euro, the digital pound and the e-CNY.

CBDC-related speeches are identified using the following broad dictionary: CBDC; central bank digital currency; central bank digital currencies; digital euro; digital pound; e-CNY; digital yuan; digital rupee; digital cash; digital currency; public digital money; retail digital currency; wholesale CBDC. The broad dictionary is necessary because central banks do not always use the acronym CBDC. In many speeches, especially those related to specific projects, the communication refers to the "digital euro", the "digital pound" or the "e-CNY" without repeatedly using the generic term. The final reproducible dictionary-based corpus used for the IMF-channel analysis contains 737 CBDC-related speeches from 2016 to 2025, identified by applying the broad CBDC dictionary to the title, description and full-text fields, followed by a consistency check that retained substantively relevant project-specific CBDC references and excluded broader digital-money mentions not directly related to central bank digital currencies.

The volume of CBDC communication increases substantially over time. The number of identified speeches rises from 14 in 2016 to 135 in 2025. Although the annual pattern is not perfectly linear, the overall trend confirms that CBDC has become a sustained topic in central bank communication. Figure 1 visualizes this increase and shows the acceleration of CBDC-related communication after 2020.

Figure 1 - CBDC-related speeches over time



Source: Authors' calculations based on the sample of CBDC-related speeches collected from the BIS Central Bank Speeches database 2016-2025.

4. Construction of the CBDC - Financial Stability Communication Indices

4.1. Methodological design

The CBDC Financial Stability Communication Index is constructed as a dictionary-based channel-coverage measure. Its purpose is to identify how broadly each CBDC-related speech covers financial-stability risks and financial-stability benefits. The index should therefore be interpreted as a measure of communication emphasis, not as a measure of realized financial stability or of the actual macro-financial effects of CBDC adoption.

The construction follows three steps. First, each speech is screened with two dictionaries: one for risk-oriented channels and one for benefit-oriented channels. Second, each channel is coded as a binary variable, taking the value of one if at least one associated term appears in the speech and zero otherwise. Third, the speech-level score is calculated as the share of active channels out of the six possible channels. This means that the index measures the breadth of channel coverage, rather than the frequency with which individual words are repeated.

FSRI captures the presence of risk-oriented financial-stability channels in each CBDC-related speech. The risk dictionary is aligned with the CBDC literature and with the IMF framework developed by Gross et al. (2025). It is organized around six channels, as reported in Table 1.

Table 1 - Risk channels

No.	Risk channel
1	Deposit substitution and deposit migration
2	Bank intermediation and bank funding
3	Credit supply and lending
4	Liquidity pressure and run risk
5	Systemic stability and contagion
6	Operational, cyber and technology risk

Source: IMF (2025).

For each speech i and risk channel c, $RiskChannel_{ic}$ equals 1 if at least one term associated with channel c appears in the speech, and 0 otherwise. The speech-level FSRI is calculated as the proportion of active risk channels:

$$FSRI_i = \frac{\sum_{c=1}^{6} RiskChannel_{ic}}{6} \quad (1)$$

where: $RiskChannel_{ic}$ is the binary indicator for risk channel c in speech i. Since the risk dictionary contains six channels, the denominator is equal to 6. A speech covering three risk channels would therefore receive an $FSRI_i$ score of $3/6 = 0.50$.

The annual FSRI is obtained by averaging the speech-level FSRI scores across all CBDC-related speeches published in year t:

$$FSRI_t = \left(\frac{1}{N_t}\right) \sum_{i \in S_t} FSRI_i \quad (2)$$

where N_t denotes the number of CBDC-related speeches in year t, and S_t denotes the set of speeches published in that year. A higher $FSRI_t$ indicates broader coverage of financial-stability risk channels in CBDC communication. It does not imply that financial instability occurred.

FSBI is constructed in the same way, but for benefit-oriented channels. The benefit dictionary captures the ways in which central banks present CBDCs as potentially supportive of financial stability. It is organized around six channels, as reported in Table 2.

Table 2 - Benefit channels

No.	Benefit channel
1	Payment-system resilience
2	Public money and monetary sovereignty
3	Financial inclusion and access
4	Efficiency, competition and innovation
5	Oversight, monitoring and transparency
6	Safeguards and design features

Source: IMF (2025).

For each speech i and benefit channel c, $BenefitChannel_{ic}$ equals 1 if at least one term associated with channel c appears in the speech, and 0 otherwise. The speech-level FSBI is calculated as the proportion of active benefit channels:

$$FSBI_i = \frac{\sum_{c=1}^{6} BenefitChannel_{ic}}{6} \quad (3)$$

The annual FSBI is then calculated by averaging the speech-level scores across all CBDC-related speeches in year t:

$$FSBI_t = \left(\frac{1}{N_t}\right) \sum_{i \in S_t} \quad FSBI_i \tag{4}$$

A higher $FSBI_t$ indicates broader coverage of benefit-oriented financial-stability channels, such as payment-system resilience, public money, safeguards or transparency. As with FSRI, the index captures communication coverage rather than the intensity of individual word use.

Finally, the net communication index NetCFSCI compares the annual benefit-oriented and risk-oriented components. It is defined as the difference between $FSBI_t$ and $FSRI_t$:

$$NetCFSCI_{\{t\}} = FSBI_{\{t\}} - FSRI_{\{t\}} \tag{5}$$

A positive NetCFSCI indicates that benefit-oriented financial-stability channels are more prominent than risk-oriented channels in a given year. A negative value indicates a risk-dominant communication profile. A value close to zero suggests that risk and benefit channels are broadly balanced in central-bank CBDC communication.

The methodology used reduces the risk that a particular index is driven by the repeated use of a small number of terms. Because each channel receives a maximum value of one in each speech, the measure captures whether a channel is present, not how often a specific word is repeated. The resulting indices are therefore best understood as channel-coverage indicators. This design also makes the annual indicators comparable over time. For each year, the index aggregates speech-level coverage across the CBDC-related speeches available in that year. As a result, changes in $FSRI_t$, $FSBI_t$ and $NetCFSCI_t$ can be interpreted as changes in the structure of central-bank communication rather than as raw changes in corpus size.

4.2. Annual evolution of CBDC financial stability communication

Table 3 presents the annual values of FSRI, FSBI and the net communication measure NetCFSCI, thereby summarizing the yearly balance between risk-oriented and benefit-oriented CBDC financial-stability communication.

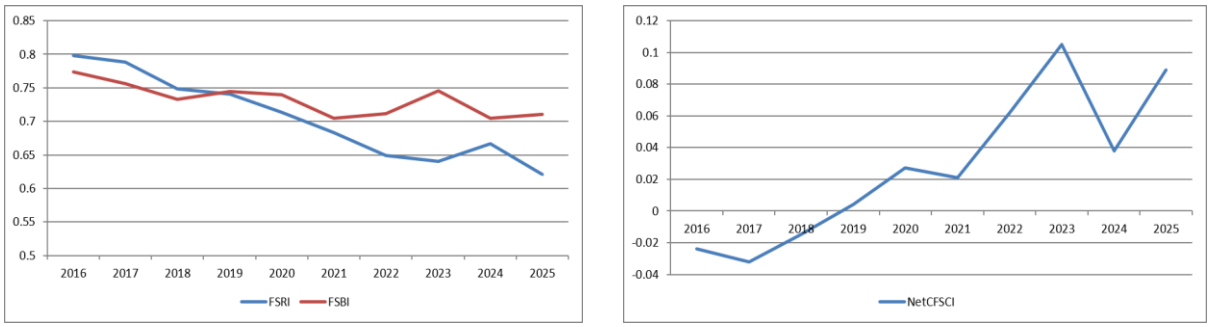
Table 3 - CBDC Financial Stability Communication Index, 2016-2025

Year	CBDC speeches	FSRI	FSBI	NetCFSCI
2016	14	0.798	0.774	-0.024
2017	26	0.788	0.756	-0.032
2018	55	0.748	0.733	-0.015
2019	45	0.741	0.744	0.004
2020	61	0.713	0.740	0.027
2021	101	0.683	0.705	0.021
2022	113	0.649	0.711	0.062
2023	87	0.640	0.745	0.105
2024	100	0.667	0.705	0.038
2025	135	0.621	0.710	0.089

Source: own computations.

Figure 2(a) illustrates the temporal evolution of FSRI and FSBI over the 2016-2025 period, whereas Figure 2(b) highlights changes in the balance between risk-oriented and benefit-oriented CBDC financial-stability narratives over time.

Figure 2 - Risk and benefit channels in CBDC communication



(a) (b)

Source: own computations.

Three results stand out analysing Figure 2. First, FSRI declines over time. It falls from 0.798 in 2016 to 0.621 in 2025. This does not mean that central banks stop discussing financial stability risks. Rather, it suggests that the range of risk channels covered in the average CBDC speech becomes narrower as the debate matures. Early speeches were more likely to cover several risk channels at once, including deposit substitution, bank intermediation, liquidity risk and systemic stability. At the same time, the decline in FSRI over the sample suggests that central banks devote relatively less attention to disintermediation, deposit substitution and run-risk concerns as CBDC projects move from conceptual debate toward design and experimentation.

Second, FSBI remains relatively stable. It stands at 0.774 in 2016 and 0.710 in 2025, with fluctuations across the intervening years. This persistence indicates that benefit-oriented narratives remain a core feature of CBDC communication. Central banks continue to link CBDC to payment resilience, the role of public money, efficiency, innovation and safeguards.

Third, the net index changes sign. NetCFSCI is negative in 2016, 2017 and 2018, but becomes positive in 2019 and remains positive through 2025. The shift is modest in magnitude, but it is substantively meaningful. It shows a movement from a slightly risk-oriented narrative toward a benefit-oriented narrative in which CBDC is presented as compatible with financial stability, conditional on careful design.

The results point to a change around 2019-2020. Before 2019, CBDC communication was more exploratory and slightly more cautious. Speeches frequently discussed the possibility that a CBDC could affect bank funding, accelerate deposit migration or create new forms of liquidity pressure. These concerns were natural in the early phase of the debate, when the institutional design of CBDCs was still highly uncertain.

After 2019, the narrative becomes more design-oriented. Major CBDC initiatives become more visible, and central banks increasingly discuss concrete arrangements such as holding limits, tiered remuneration, intermediated models, privacy safeguards and operational resilience. The policy question shifts from whether CBDCs are inherently destabilizing to how their architecture can preserve the benefits of public money while limiting risks to bank intermediation and liquidity.

The highest values of NetCFSCI are observed in 2023 and 2025. This pattern suggests that the acceleration of CBDC projects did not lead central banks to present CBDC as increasingly

destabilizing. Instead, the public narrative increasingly emphasizes resilience, safeguards and the modernization of payment systems.

For transparency and additional context, Appendix B reports an institutional comparison of CBDC financial-stability communication for selected major central banks.

4.3. Simple trend model

To complement the descriptive evidence, the analysis estimates a simple annual trend model to assess whether the observed changes in the communication indices reflect systematic temporal patterns rather than purely descriptive fluctuations.

$$Index_{\{t\}} = \alpha + \beta Trend_{\{t\}} + \varepsilon_{\{t\}} \quad (6)$$

where $Index_t$ denotes either FSRI or FSBI, α is the intercept, β is the trend coefficient and ε_t is the error term. $Trend_t$ is a linear time trend taking the value 1 in 2016 and 10 in 2025. The purpose of this model is not to establish causality, but to test whether the component-level movements are systematic rather than purely visual.

Table 4 - Annual trend regression

Dependent variable	Constant term	Trend coefficient	Trend p-value	R-squared	Adjusted R-squared	Observations
FSRI	0.8143	-0.0199	0.000006	0.932	0.924	10
FSBI	0.7659	-0.0061	0.008292	0.603	0.553	10

Notes: The models are estimated by ordinary least squares with an intercept, using the annual FSRI and FSBI mean values reported to three decimal places. The time trend takes values from 1 in 2016 to 10 in 2025.

Separate trend regressions for the two components show that FSRI declines more strongly than FSBI. In the annual series, the FSRI trend coefficient is -0.0199, while the FSBI trend coefficient is -0.0061. Both coefficients are statistically significant at the 95% confidence level, indicating systematic component-level changes over the sample period.

4.4. First-Order Lagged Regressions (AR(1))

To complement the deterministic trend analysis, two first-order autoregressive models are estimated by ordinary least squares using the annual FSRI and FSBI values reported to three decimal places. The models examine whether the breadth of risk- and benefit-oriented channel coverage in year t is associated with its value in year t−1. Because introducing one lag reduces the 2016–2025 series by one observation, each regression uses nine annual pairs covering 2017–2025. The specifications are:

$$FSRI_t = \alpha + \beta * FSRI_{t-1} + \varepsilon_t \quad (7)$$

$$FSBI_t = \alpha + \beta * FSBI_{t-1} + \varepsilon_t \quad (8)$$

where α is the intercept, β is the first-order lag coefficient and ε_t is the error term. The coefficient β measures year-to-year persistence in the respective communication index. A positive and statistically significant coefficient indicates that a higher index value in year t−1 is associated with a higher value in year t. Since the models are estimated from annual aggregate observations, β

should be interpreted as predictive persistence rather than as evidence that earlier speeches causally influence subsequent communication.

Table 5 - First-order autoregressive regressions

Dependent variable	Constant term	Lag coefficient	Lag p-value	R-squared	Adjusted R-squared	Observations
FSRI	0.0573	0.8922	0.000369	0.854	0.833	9
FSBI	0.4741	0.3451	0.268658	0.171	0.053	9

Notes: OLS estimates with an intercept, based on annual mean index values rounded to three decimal places. Reported p-values are two-sided. The inclusion of one annual lag reduces the estimation sample from ten to nine observations.

The FSRI model reveals strong persistence in risk-oriented communication. The lag coefficient is 0.8922 and is statistically significant ($p < 0.001$), while the model explains approximately 85.4 percent of the annual variation in FSRI. This result indicates that years characterized by broader risk-channel coverage tend to be followed by years with similarly broad coverage, even though the trend analysis documents a gradual decline in the overall level of FSRI.

By contrast, the FSBI lag coefficient is smaller, at 0.3451, and is not statistically significant ($p = 0.269$). Its explanatory power is also limited (R-squared = 0.171). The results therefore provide no evidence of systematic year-to-year persistence in benefit-oriented communication. This weaker temporal dependence is consistent with, although it does not directly demonstrate, greater responsiveness to contemporaneous policy developments, project milestones and changes in the CBDC debate.

These results are exploratory because the annual sample provides only nine lagged observations. Moreover, the high estimated persistence of FSRI may partly reflect the smooth downward trend in the series. The AR(1) estimates should therefore be read as a compact description of temporal dependence that complements, rather than replaces, the annual trend regressions; they do not establish causal influence or demonstrate stationarity.

4.5 Robustness analysis using topic modelling validation and a stricter CBDC corpus definition

4.5.1. Topic modelling validation methodology and results

Topic modelling is used as an unsupervised validation exercise for the dictionary-based communication indices (for a recent review of the literature regarding the use of topic modelling, please refer to Bernhard-Harrer et al., 2025). While the FSRI and FSBI are constructed from theoretically defined risk and benefit channels, the topic model provides an independent text-mining check of whether the dominant themes emerging from the corpus are consistent with the interpretation obtained from the dictionary approach.

The topic model is estimated using non-negative matrix factorization (NMF) applied to a Term Frequency- Inverse Document Frequency (TF-IDF) document-term matrix. NMF is appropriate for this validation exercise because it decomposes a sparse, non-negative text matrix into additive and interpretable components: a document-topic matrix and a topic-term matrix. Unlike a purely dictionary-based classification, NMF allows the main thematic structures to emerge from word co-occurrence patterns in the corpus.

The methodological procedure developed by Maier et al. (2018), comprises the following steps. First, the model uses the same 737 CBDC-related speeches as the baseline analysis. To reduce

noise from speeches that discuss several policy issues, the model is estimated on CBDC-relevant sentences and their immediate context within each speech. Text preprocessing includes lowercasing, removal of URLs and punctuation, removal of standard English stop words, and removal of corpus-defining terms such as CBDC, digital and currency.

Second, the cleaned text is transformed into a TF-IDF matrix. Let X denote the non-negative document-term matrix, where each row represents a speech and each column represents a retained term. The matrix includes 737 documents and 1,800 TF-IDF features, corresponding to the unigram and bigram terms retained after preprocessing and after applying a minimum document frequency of five and a maximum document frequency of 70 percent.

$$X \in \mathbb{R}_{+}^{n \times m} \tag{9}$$

This formula shows that X is the TF-IDF document-term matrix used in the topic-modelling analysis. The matrix has n rows and m columns: each row represents one speech, and each column represents one term retained after preprocessing. The symbol $\mathbb{R}_{+}$ indicates that all matrix values are non-negative, which is required for NMF.

Third, NMF matrix approximation:

$$X \approx W H \tag{10}$$

This formula expresses the central idea of NMF. The original TF-IDF matrix X is approximated by the product of two smaller non-negative matrices, W and H, where W is the document-topic matrix and H is the topic-term matrix. This means that the observed text structure is represented through a limited number of latent topics. NMF approximates the TF-IDF matrix as the product of two lower-dimensional non-negative matrices.

Fourth, dimensions of the factor matrices:

$$W \in \mathbb{R}_{+}^{n \times k}, \; H \in \mathbb{R}_{+}^{k \times m} \tag{11}$$

This formula defines the dimensions of the two matrices produced by the factorization. W contains the weight of each topic in each speech: it has n rows, one for each speech, and k columns, one for each topic. H contains the weight of each term in each topic: it has k rows, one for each topic, and m columns, one for each term in the TF-IDF vocabulary. In this study, k = 6.

The number of topics is fixed at six. This specification is retained for theoretical comparability with the six risk channels, the six benefit channels and the six topic profiles used in the cluster analysis. The purpose is not to claim that six is the only possible data-driven optimum, but to provide a parsimonious and interpretable validation layer that can be compared directly with the rest of the empirical design.

NMF optimization criterion:

$$\arg \min_{W,H \geq 0} \left|\left|X - WH\right|\right|_{F}^{2} \tag{12}$$

This formula shows the optimization problem solved by NMF. The model searches for the non-negative matrices W and H that minimize the difference between the original matrix X and its approximation WH. The term $\left|\left|X - WH\right|\right|_{F}^{2}$ is the squared Frobenius norm, which measures the reconstruction error between the original TF-IDF matrix and its low-rank NMF

approximation. A lower value means that the topic structure reconstructs the TF-IDF matrix more closely.

Topic share in the corpus:

$$TopicShare_c = \frac{sum_{i=1}^{n} W_{ic}}{sum_{c=1}^{k} sum_{i=1}^{n} W_{ic}} \quad (13)$$

This formula calculates the share of topic c in the full corpus. The numerator sums the contribution of topic c across all speeches. The denominator sums the contributions of all topics across all speeches. The resulting value indicates the relative importance of each topic in the overall thematic structure of the corpus.

NMF is useful in this study because it produces additive and interpretable topics from a sparse TF-IDF matrix. The highest-weighted terms in each row of H are used to label the topics, while the weights in W indicate how strongly each topic appears in each speech. The method is used as a validation layer for the dictionary-based indices, not as a replacement for FSRI and FSBI. Its main limitation is that the number of topics must be specified in advance and that topic labels remain interpretive.

Table 6 - NMF topic-modelling results based on the TF-IDF matrix

Topic	NMF interpretation	Share	Dominant speeches	Top terms
1	Payment infrastructure and cross-border payments	20.9%	137	payment; payments; cross-border; cross-border payments; systems; payment systems; retail; banks; instant; payment system
2	Financial stability, banking and monetary policy	20.5%	147	banks; policy; financial; monetary; monetary policy; system; stability; balance; implications; financial stability
3	Digital euro, privacy and public payment architecture	15.2%	141	euro; payment; payments; cash; phase; public; means; privacy; means payment; solutions
4	Public money, cash and commercial-bank deposits	14.8%	94	money; bank money; cash; commercial; e-krona; public; payments; private; deposits; banks
5	Wholesale CBDC, tokenised settlement and DLT	14.5%	101	settlement; bank money; wholesale; money; tokenised; financial; assets; dlt; asset; transactions
6	Stablecoins, crypto-assets and risk governance	14.3%	117	stablecoins; crypto-assets; stablecoin; financial; regulatory; risks; assets; crypto; regulation; crypto-asset

The NMF results show that CBDC communication is organized around a small number of recurrent themes rather than around a single dominant narrative. The largest topic, accounting for 20.9 percent of the topic mass, relates to payment infrastructure and cross-border payments. The second topic, at 20.5 percent, links financial stability, banking-sector language and monetary-policy framing. This confirms that financial-stability concerns are central to the corpus, but they do not dominate it.

The next two topics relate to the digital euro and public money. The digital-euro, privacy and public-payment topic accounts for 15.2 percent of the topic mass, while the public-money, cash and commercial-bank deposit topic accounts for 14.8 percent. These results are consistent with the article's interpretation that CBDC is increasingly framed as part of the modernization of payment systems and the preservation of public money in a digital environment.

The remaining topics capture wholesale CBDC and tokenised settlement, as well as stablecoins, crypto-assets and risk governance. Taken together, these topics show that financial-stability concerns are embedded in broader discussions of payment infrastructure, monetary sovereignty, private digital money and design safeguards. The NMF evidence therefore reinforces the main index-based finding: CBDC-related risks remain visible in central-bank communication, but they are articulated within a broader architecture of public money, payment resilience and regulated digital finance.

4.5.2. The use of a stricter CBDC corpus definition

The dictionary index provides a structured measure of risk and benefit channels, while topic modelling offers an independent view of the themes that dominate the corpus. The fact that both approaches point toward a design-oriented and benefit-leaning narrative strengthens the interpretation.

However, as an additional robustness check, we use a stricter CBDC corpus definition. The strict dictionary includes direct CBDC and project-specific terms, such as "CBDC", "central bank digital currency", "digital euro", "digital pound", "e-CNY", "digital yuan", "digital rupee", "retail CBDC" and "wholesale CBDC". It excludes broader expressions such as "digital money" and "digital currency", which may capture a wider set of digital-money discussions. This stricter definition identifies 667 CBDC-related speeches between 2016 and 2025.

Table 7 - Robustness check based on strict CBDC corpus

Year	Strict CBDC speeches	FSRI	FSBI	NetCFSCI
2016	7	0.857	0.833	-0.024
2017	12	0.806	0.750	-0.056
2018	40	0.788	0.717	-0.071
2019	33	0.783	0.747	-0.035
2020	57	0.705	0.737	0.032
2021	97	0.694	0.722	0.027
2022	110	0.648	0.708	0.059
2023	84	0.639	0.750	0.111
2024	96	0.661	0.696	0.035
2025	131	0.616	0.712	0.097

The robustness results in Table 8 are closely aligned with the baseline trend regressions reported in Table 4. In both specifications, the FSRI coefficient is negative, large in magnitude and statistically significant at the 95% confidence level, confirming a systematic decline in risk-channel coverage over time. The decline is even stronger in the strict-corpus robustness model (-0.0263) than in the baseline model (-0.0199), while explanatory power remains very high in both cases (R-squared of 0.927 compared with 0.932). The FSBI coefficient is also negative and statistically significant in both tables, but its magnitude remains smaller than the FSRI coefficient: -0.0087 in the robustness model compared with -0.0061 in the baseline model. This confirms that the main pattern is robust to the stricter CBDC corpus definition: both components decline over

time, but risk-oriented communication declines more sharply than benefit-oriented communication. The robustness check therefore strengthens the interpretation that the observed shift is not driven by the inclusion of broader digital-money terminology, but reflects a stable component-level change in CBDC financial stability communication.

Table 8 - Annual trend regression - robustness

Dependent variable	Constant term	Trend coefficient	Trend p-value	R-squared	Adjusted R-squared	Observations
FSRI	0.8645	-0.0263	0.000008	0.927	0.918	10
FSBI	0.7850	-0.0087	0.030068	0.464	0.397	10

4.6. Discussions

The results provide a detailed view of how CBDC-related financial stability communication changes over time. The component-level trend regressions in Table 4 show that both risk-oriented and benefit-oriented channel coverage decline over the 2016-2025 period, but the decline is much stronger for FSRI than for FSBI. The FSRI trend coefficient is -0.0199 and is statistically significant at the 95% confidence level, while the FSBI coefficient is -0.0061 and is also significant at the 95% confidence level. The AR(1) results in Table 5 capture a different temporal property: FSRI displays strong year-to-year persistence, whereas the lagged FSBI coefficient is not statistically significant. Distinguishing trend from persistence is important: FSBI declines less steeply over time, but this does not imply that its annual values are more autoregressively persistent.

Substantively, this finding qualifies a simple risk-centered reading of the CBDC debate. Theoretical models often emphasize the possibility that CBDCs could destabilize bank funding, accelerate deposit migration or intensify run dynamics. These concerns are real, and they remain present in central bank speeches. However, the annual trend evidence suggests that their relative coverage becomes less broad over time. CBDC communication therefore appears less dominated by multiple risk channels as the policy debate moves from conceptual exploration toward concrete design, experimentation and safeguards.

The robustness results in Table 8 strengthen this interpretation. Under the stricter CBDC corpus definition, the FSRI coefficient becomes even more negative (-0.0263) and remains highly significant, while the FSBI coefficient is also negative and significant but smaller in magnitude (-0.0087). The close alignment between Table 4 and Table 8 shows that the component-level trends are not an artifact of including broader digital-money terminology. The core result is stable: risk-oriented communication declines more sharply than benefit-oriented communication in both the baseline and robustness specifications.

This does not mean that central banks communicate CBDCs as risk-free instruments. Rather, the evidence points to a more conditional and design-oriented narrative. Holding limits, tiered remuneration, intermediated distribution, privacy arrangements, operational resilience and convertibility conditions are repeatedly used to show how risks can be contained. As a result, financial stability risks are increasingly discussed through the lens of institutional design and mitigation rather than as unmanaged threats. Further, as shown in Table 6, CBDC communication is organized around a small number of recurrent themes rather than around a single dominant narrative.

Therefore, in the next analysis we show that CBDC speeches are organized around recurrent combinations of risk and benefit channels rather than isolated themes.

5. Cluster analysis

5.1. Methodological framework

The cluster analysis is conducted separately for three groups of binary indicators. The first group captures the six risk channels used in the FSRI index: deposit substitution and deposit migration; bank intermediation and bank funding; credit supply and lending; liquidity pressure and run risk; systemic stability and contagion; and operational, cyber and technology risk. The second group captures the six benefit channels used in the FSBI index: payment-system resilience; public money and monetary sovereignty; financial inclusion and access; efficiency, competition and innovation; oversight, monitoring and transparency; and safeguards and design features. The third group captures the six theory-guided thematic profiles: digital money architecture; cross-border payments and infrastructure; financial stability and banking sector; regulation, governance and risk management; digital euro and public money; and monetary policy and liquidity.

Each speech is represented as a vector:

$$x_i = \left(x_{\{i1\}}, x_{\{i2\}}, \ldots, x_{\{i6\}}\right) \tag{14}$$

where $x_{ij} = 1$ if channel j is present in speech i and $x_{ij} = 0$ otherwise. The distance between speeches is computed in this six-dimensional indicator space. The purpose is to identify whether speeches are organized around stable multidimensional communication profiles.

The final presentation focuses on k-means clustering as the substantive typology. Hierarchical clustering is not retained because it produced one overwhelmingly dominant cluster in each analytical block, which confirms the existence of a broad common communication core but does not offer a useful typological separation. DBSCAN is retained as a robustness-oriented density check for the risk and benefit indicators. For DBSCAN, epsilon is fixed at 0.1 so that only exact binary profiles are grouped, while MinPts is set to 5. This specification is designed to identify repeated risk and benefit configurations without collapsing distinct binary profiles into overly broad dense groups.

K-means clustering partitions the speeches into k clusters by minimizing within-cluster dispersion (Jain et al., 1999):

$$arg\ min_{C_1,\ldots,C_k} \sum_{c=1}^{k} \sum_{x_i \in C_c} \|x_i - \mu_c\|^2 \tag{15}$$

where:

$$\mu_c = \frac{1}{|C_c|} \sum_{x_i \in C_c} x_i \tag{16}$$

C_c denotes cluster c, μ_c is the centroid of cluster C_c, and x_i is the six-dimensional vector representation of speech i.

The three-cluster solution is retained as a theory-guided and parsimonious specification, supported by interpretability and cross-block comparability, rather than as a purely data-driven optimum. Its advantage is interpretability and balanced comparison across analytical blocks. Its

limitation is that it imposes a fixed number of clusters and assigns every speech to a cluster even when a speech is relatively atypical.

Density-based clustering is implemented through DBSCAN (Sander et al. 1998). For each speech represented by the vector x_i, the epsilon-neighbourhood is defined as:

$$N_{\{\varepsilon\}(x_i)} = \{x_j \in X \mid d(x_i, x_j) \leq \varepsilon\} \quad (17)$$

A speech is classified as a core point if:

$$\left|N_{\{\varepsilon\}(x_i)}\right| \geq MinPts \quad (18)$$

where ε denotes the neighbourhood radius and MinPts represents the minimum number of observations required to form a dense region. Because the indicators are binary and limited to six dimensions, DBSCAN results are reported with full active-channel profiles rather than only top-three channels.

5.2. Risk-channel cluster analysis

5.2.1. K-means clustering results for risk channels

K-means clustering is used as the main typological specification because it offers a parsimonious and directly comparable classification framework across risk channels, benefit channels and thematic dimensions. We have opted for a K=3 cluster for two reasons. On the one hand, the inertia against various values of K drops considerably after K=3. On the other hand, working with more clusters generates a strong duplication of channels in each cluster.

For the risk-channel specification, the solution yields a silhouette coefficient of 0.323 and a within-cluster sum of squares of 489.1, indicating a moderate but interpretable degree of cluster separation (Table 9).

Table 9 - K-means clustering for the risk-channel specification

Cluster	N	Share	Top three channels	Dominant years
1	234	31.8%	bank intermediation and funding (1.00); operational, cyber and technology risk (1.00); systemic stability and contagion (0.92)	2021 (38); 2022 (38); 2025 (31); 2024 (30)
2	256	34.7%	bank intermediation and funding (0.96); systemic stability and contagion (0.88); credit supply and lending (0.87)	2025 (52); 2024 (42); 2023 (34); 2021 (29)
3	247	33.5%	bank intermediation and funding (0.90); systemic stability and contagion (0.75); credit supply and lending (0.40)	2025 (52); 2022 (48); 2021 (34); 2023 (31)

The k-means solution identifies three relatively balanced risk profiles. The first cluster combines bank intermediation and funding with operational, cyber and technology risk and systemic stability, suggesting a design-and-resilience risk profile. The second cluster emphasizes bank intermediation, systemic stability and credit supply, corresponding to the conventional macro-financial concern that CBDC may affect bank funding and lending. The third cluster is also centred on bank intermediation and systemic stability, but with a weaker credit-supply component, indicating a more general banking-sector stability profile. Overall, risk

communication is not fragmented across isolated threats; it is organized around recurring bundles in which bank intermediation remains the central channel.

5.2.2. DBSCAN results for risk channels

DBSCAN is used as a density-based robustness check. With ε = 0.1 and MinPts = 5, the method identifies 34 dense risk profiles and classifies 52 speeches as noise, representing 7.1% of the corpus. Because the lower MinPts threshold produces a more detailed profile map, Table 10 reports the five largest dense profiles.

Table 10 - Density-based clustering, DBSCAN risks

Profile	N	Share	Active channels	Dominant years
1	67	9.1%	bank intermediation and bank funding; credit supply and lending; liquidity pressure and run risk; systemic stability and contagion; operational, cyber and technology risk	2021 (14); 2018 (10); 2022 (10); 2024 (10)
2	66	9%	deposit substitution and deposit migration; bank intermediation and bank funding; credit supply and lending; liquidity pressure and run risk; systemic stability and contagion; operational, cyber and technology risk	2021 (10); 2025 (9); 2022 (8); 2023 (8)
3	54	7.3%	systemic stability and contagion; operational, cyber and technology risk	2025 (16); 2023 (9); 2021 (8); 2024 (8)
4	36	4.9%	systemic stability and contagion	2022 (11); 2025 (7); 2020 (4); 2021 (4)
5	34	4.6%	deposit substitution and deposit migration; bank intermediation and bank funding; liquidity pressure and run risk; systemic stability and contagion; operational, cyber and technology risk	2023 (6); 2025 (6); 2020 (5); 2024 (5)

The DBSCAN result confirms that the risk vocabulary is organized around repeated exact profiles rather than entirely unique speech-level combinations. The largest dense profile excludes deposit-substitution language but combines bank intermediation, credit, liquidity, systemic stability and operational or cyber risk. The second profile activates all six risk channels, showing that some speeches adopt a comprehensive risk framing. The remaining large profiles isolate systemic stability and operational/cyber concerns or combine deposit migration with bank funding, liquidity and operational risk. This supports the interpretation that financial-stability communication remains bank-centred, but the intensity and breadth of the risk bundle vary across speeches.

5.3. Benefit-channel cluster analysis

5.3.1. K-means cluster results for benefit channels

For the benefit-channel specification, the solution yields a silhouette coefficient of 0.417 and a within-cluster sum of squares of 378.0, suggesting a more clearly differentiated cluster structure than in the risk-channel analysis (Table 11).

Table 11 - K-means clustering for the benefit-channel specification

Cluster	N	Share	Top three channels	Dominant years
1	264	35.8%	oversight, monitoring and transparency (1.00); payment-system resilience (0.97); efficiency, competition and innovation (0.96)	2022 (42); 2021 (38); 2025 (38); 2024 (37)
2	237	32.2%	safeguards and design features (1.00); efficiency, competition and innovation (0.98); payment-system resilience (0.98)	2025 (45); 2023 (34); 2021 (33); 2024 (33)
3	236	32%	efficiency, competition and innovation (0.89); payment-system resilience (0.87); public money and monetary sovereignty (0.69)	2025 (52); 2022 (40); 2021 (30); 2024 (30)

The k-means solution separates three benefit narratives. The first cluster links oversight, monitoring and transparency with payment-system resilience and efficiency, reflecting a governance-oriented benefit profile. The second cluster is centred on safeguards and design features, combined with efficiency and payment-system resilience, indicating that many speeches frame CBDC benefits as conditional on institutional and technical design. The third cluster emphasizes efficiency, payment-system resilience and public money, capturing a modernization narrative in which CBDC is presented as a way to preserve public money while improving payment infrastructures.

5.3.2. DBSCAN results for benefit channels

For the benefit channels, with $\varepsilon = 0.1$ and MinPts = 5, the method identifies 19 dense benefit profiles and classifies 44 speeches as noise, representing 6% of the corpus. Because the lower MinPts threshold produces a more detailed profile map, Table 12 reports the five largest dense profiles:

Table 12 - Density-based clustering, DBSCAN benefits

Profile	N	Share	Active channels	Dominant years
1	166	22.5%	payment-system resilience; public money and monetary sovereignty; financial inclusion and access; efficiency, competition and innovation; oversight, monitoring and transparency; safeguards and design features	2022 (31); 2025 (29); 2021 (23); 2020 (22)
2	139	18.9%	payment-system resilience; public money and monetary sovereignty; efficiency, competition and innovation; oversight, monitoring and transparency; safeguards and design features	2025 (39); 2024 (20); 2021 (14); 2022 (14)
3	103	14%	payment-system resilience; financial inclusion and access; efficiency, competition and innovation; oversight, monitoring and transparency; safeguards and design features	2022 (18); 2023 (18); 2024 (17); 2018 (16)
4	74	10%	payment-system resilience; efficiency, competition and innovation; oversight, monitoring and transparency; safeguards and design features	2024 (12); 2025 (12); 2018 (10); 2022 (9)
5	29	3.9%	payment-system resilience; public money and monetary sovereignty; efficiency, competition and innovation; safeguards and design features	2025 (7); 2021 (5); 2022 (4); 2017 (3)

The DBSCAN benefit results separate dense profiles by their complete six-channel structure. The largest profile activates all benefit channels, indicating a comprehensive pro-CBDC framing. The second profile excludes financial inclusion but retains payment resilience, public money, efficiency, oversight and safeguards. The third and fourth profiles weaken the public-money and inclusion dimensions while preserving the infrastructure, efficiency, oversight and design channels. These results confirm that the benefit narrative is broad, but its most stable core is formed by payment-system resilience, efficiency, oversight and safeguards.

5.4. Thematic-profile cluster analysis

5.4.1. K-means cluster results for thematic profile

For the thematic-based specification, the solution yields a silhouette coefficient of 0.641 and a within-cluster sum of squares of 210.0, indicating a relatively strong and clearly interpretable cluster separation (Table 13).

Table 13. K-means clustering across CBDC thematics

Cluster	N	Share	Active thematic dimensions	Dominant years
1	517	70.1%	digital euro and public money (1.00); monetary policy and liquidity (1.00); financial stability and banking sector (0.96)	2025 (97); 2022 (76); 2024 (67); 2021 (66)
2	165	22.4%	digital money architecture (0.97); regulation, governance and risk management (0.95); cross-border payments and infrastructure (0.93)	2025 (34); 2021 (28); 2023 (25); 2022 (24)
3	55	7.5%	monetary policy and liquidity (1.00); financial stability and banking sector (0.91); digital money architecture (0.89)	2022 (13); 2023 (9); 2024 (9); 2021 (7)

The k-means solution identifies one large thematic cluster and two smaller thematic profiles. The first cluster links digital euro and public money, monetary policy and liquidity and financial stability and banking sector themes, showing that a large share of CBDC communication embeds financial-stability concerns within public-money and liquidity debates.

5.4.2. DBSCAN results for thematic profile

In this case, DBSCAN identifies 12 dense thematic profiles and classifies 46 speeches as noise, representing 6.2% of the corpus. Because the lower MinPts threshold produces a more detailed profile map, the table reports the five largest dense profiles.

The thematic DBSCAN results reveal a highly recurrent common core (Table 14). The largest profile activates all six thematic dimensions and contains 411 speeches (55.8% of the corpus). The second profile retains digital money architecture, cross-border payments and infrastructure, financial stability and banking, regulation and governance, and digital euro and public money, but excludes monetary policy and liquidity; it contains 108 speeches (14.7%). The next three profiles each omit one dimension from the six-theme core: digital euro and public money, cross-border payments and infrastructure, or digital money architecture. Together, these results show that thematic communication is organized around a broad institutional bundle, while the lower-density profiles reflect selective emphasis rather than a substantively separate typology. The 46 noise observations capture uncommon combinations that occur fewer than five times.

Table 14 - Density-based clustering, DBSCAN thematic profiles

Cluster	N	Share	Active thematic dimensions	Dominant years
1	411	55.8%	Digital Money Architecture; Cross-Border Payments and Infrastructure; Financial Stability and Banking Sector; Regulation, Governance and Risk Management; Digital Euro and Public Money; Monetary Policy and Liquidity	2025 (71); 2022 (59); 2024 (53); 2021 (48)
2	108	14.7%	Digital Money Architecture; Cross-Border Payments and Infrastructure; Financial Stability and Banking Sector; Regulation, Governance and Risk Management; Digital Euro and Public Money	2025 (20); 2023 (19); 2022 (16); 2024 (16)
3	38	5.2%	Digital Money Architecture; Cross-Border Payments and Infrastructure; Financial Stability and Banking Sector; Regulation, Governance and Risk Management; Monetary Policy and Liquidity	2022 (8); 2021 (6); 2023 (6); 2024 (6)

Cluster	N	Share	Active thematic dimensions	Dominant years
4	30	4.1%	Digital Money Architecture; Financial Stability and Banking Sector; Regulation, Governance and Risk Management; Digital Euro and Public Money; Monetary Policy and Liquidity	2021 (9); 2024 (6); 2025 (5); 2018 (2)
5	22	3%	Cross-Border Payments and Infrastructure; Financial Stability and Banking Sector; Regulation, Governance and Risk Management; Digital Euro and Public Money; Monetary Policy and Liquidity	2025 (8); 2022 (4); 2021 (3); 2020 (2)

The DBSCAN results provide robustness evidence for the existence of a recurrent common thematic core rather than for a sharply differentiated alternative typology. With $\varepsilon = 0.1$ and MinPts = 5, 93.8% of speeches belong to repeated exact thematic configurations, while only 6.2% are classified as noise. The dominance of the complete six-theme profile confirms that CBDC communication systematically combines monetary architecture, payment infrastructure, financial stability, governance, public money and liquidity-related considerations. At the same time, the strong concentration in the largest profile suggests that DBSCAN should be interpreted as a density-based confirmation of thematic recurrence, rather than as an independent validation of the three-cluster k-means taxonomy.

5.5. Comparative discussion

Across all three analytical blocks, the clustering evidence is consistent with the main CBDC financial stability communication indices results. The communication indices analysis shows that risk-oriented communication declines more strongly over time than benefit-oriented communication, while the net communication index becomes positive from 2019 onward. The cluster analysis clarifies the internal structure behind this trend. Risks do not disappear from CBDC communication; instead, they are increasingly embedded in broader discussions of design, safeguards, public money, payment-system resilience, governance and digital monetary architecture.

The risk clusters show that bank intermediation and systemic stability remain central communication channels. The DBSCAN analysis adds a more detailed view of this evidence, by showing which exact combinations recur most often. The benefit clusters show a stable and cohesive communication core, especially around payment-system resilience, efficiency, safeguards and oversight. The topic clusters further support this interpretation by showing that financial stability is commonly linked to digital money architecture, regulation, governance, liquidity and public-money themes.

Methodologically, the strongest substantive evidence comes from the k-means solution, because it produces comparable and interpretable profiles across the three analytical categories. DBSCAN is useful as a robustness check for risks and benefits when MinPts = 5 and full six-channel profiles are reported. Its main contribution is to show that some exact risk and benefit configurations are dense and recurrent, while a smaller number of speeches remain more heterogeneous. The overall conclusion remains aligned with the main results: the movement toward benefit-oriented financial-stability channels in central bank communication does not imply that financial-stability risks disappear, but that they are reframed within a broader design-oriented and institutionally grounded CBDC narrative.

6. Conclusions

This paper examined how central banks communicate the financial stability implications of CBDCs, using a corpus of 737 BIS central bank speeches published between 2016 and 2025. The analysis shows that CBDC communication combines both risk-oriented and benefit-oriented financial stability narratives, but that the breadth of these narratives changes over time. The annual trend regressions indicate a statistically significant temporal decline in the coverage of both FSRI and FSBI, suggesting that CBDC speeches have gradually moved from broad exploratory discussions toward more focused and selective communication. The decline is more pronounced for risk-oriented channels, whereas FSBI follows a less steep downward trend. The AR(1) estimates add a distinct dynamic result: FSRI exhibits strong and statistically significant year-to-year persistence, while the lagged FSBI coefficient is smaller and statistically insignificant. Risk-oriented communication therefore displays greater temporal continuity, whereas the evidence does not support systematic persistence in benefit-oriented communication. These results should nevertheless be interpreted as exploratory because the models contain only nine lagged annual observations and the estimated FSRI persistence may partly reflect the downward trend in the series.

The NMF topic-modelling validation reinforces this interpretation. The latent thematic structure of the corpus is organized around six major topics: payment infrastructure and cross-border payments; financial stability, banking and monetary policy; the digital euro, privacy and public payments; public money, cash and deposits; wholesale CBDC, tokenised settlement and DLT; and stablecoins, crypto-assets and risk governance. This topic structure confirms that financial stability is embedded within a broader CBDC communication agenda, linking banking-sector concerns to payment-system modernization, monetary sovereignty, public money, technological infrastructure and the governance of digital finance.

The clustering results further show that CBDC communication is not randomly distributed across individual channels, but is organized around recurrent combinations of risks and benefits. K-means clustering identifies interpretable communication profiles, while DBSCAN confirms that several dense and recurrent channel configurations appear repeatedly across the corpus. These findings indicate that central banks tend to discuss financial stability risks and benefits as bundled policy concerns rather than as isolated themes.

Overall, the results suggest that central bank communication on CBDCs has evolved from a broad financial stability debate toward a more differentiated narrative in which risks, benefits and design safeguards are increasingly discussed together. The study contributes to the CBDC literature by providing a reproducible dictionary-based communication index, validating its interpretation through NMF topic modelling, and showing through cluster analysis that CBDC-related financial stability communication follows identifiable thematic profiles. The findings should nevertheless be interpreted as evidence on communication patterns rather than as direct evidence on financial stability outcomes.

Several limitations remain. First, dictionary-based measures depend on the selection of terms. The dictionaries used here are grounded in the CBDC and financial stability literature, especially the IMF framework, but alternative dictionaries could generate different magnitudes. Second, speeches are public communication instruments. They reveal how central banks frame CBDC externally, not necessarily the full range of internal concerns. Third, the BIS speeches database may not capture all relevant CBDC communication, such as technical reports, consultation documents or legislative testimony. Fourth, the paper studies communication rather than realized financial outcomes. It therefore cannot establish whether CBDCs improve or weaken financial stability in practice.

Future research could extend the analysis in several directions. One possibility is to link the communication index to CBDC project milestones, including consultations, pilots and launch decisions. Another is to compare speeches with central bank reports and consultation papers. A third direction would be to combine the text-based index with macrofinancial variables, such as deposit growth, bank liquidity indicators or payment-system stress measures. Finally, supervised classification methods could be used to distinguish more precisely between risk, benefit and safeguard narratives.

References

Adalid, R., Alvarez-Blazquez, A., Assenmacher, K., Burlon, L., Dimou, M., Lopez-Quiles, C., Martin Fuentes, N., Meller, B., Munoz, M. A., Radulova, P., Rodriguez d'Acri, C., Shakir, T., Silova, G., Soons, O., & Ventula Veghazy, A. (2022). Central Bank Digital Currency and Bank Intermediation. ECB Occasional Paper Series No. 293.

Bank for International Settlements. (2020). Central Bank Digital Currencies: Foundational Principles and Core Features. BIS Report No. 1.

Bank for International Settlements. (2021). Central Bank Digital Currencies: Financial Stability Implications. BIS Report No. 4.

Bernhard-Harrer, J., Ashour, R., Eberl, J-M., Tolochko, P., & Boomgaarden, H. (2025). Beyond standardization: a comprehensive review of topic modelling validation methods for computational social science research. Political Science Research and Methods, 1-19.

Bindseil, U. (2020). Tiered CBDC and the financial system. ECB Working Paper Series, No. 2351.

Born, B., Ehrmann, M., & Fratzscher, M. (2014). Central bank communication on financial stability. The Economic Journal, 124(577), 701–734.

Chen, H., & Filippin, M. E. (2026). Central bank digital currency with collateral-constrained banks. Journal of Money, Credit and Banking, https://doi.org/10.1111/jmcb.70063

Copple, D., Faisal, I., Khairnar, K., Lau, J., Malik, S., Poensgen, V. (2025). The role of holding limits for sterling-denominated systemic stablecoins and a potential digital pound. Financial Stability Paper No. 53, Bank of England.

Correa, R., Garud, K., Londono, J. M., & Mislang, N. (2021). Sentiment in central banks' financial stability reports. Review of Finance, 25(1), 85–120.

Fell, J., Gardó, S., Kellner, D., Klaus, B., Lang, J. H., Nagy, L., Vendrell, P., Rusnák, M., Wendelborn, J., & Wredenborg, S. (2026). From dictionaries to AI: A new era in sentiment analysis for financial stability. Financial Stability Review, May 2026, European Central Bank.

Gross, M., Miccoli, M., Verrier, J., & Villegas-Bauer, G. (2025). Evaluating the Implications of CBDC for Financial Stability. IMF Fintech Note 2025/008.

IMF (2025). IMF's Central Bank Digital Currency (CBDC) Virtual Handbook. Available at: https://www.imf.org/en/topics/digital-payments-and-finance/central-bank-digital-currency/virtual-handbook.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data Clustering: A Review. ACM Computing Surveys, 31(3), 264-323.

Katsurada, K. (2026). The shift in China's CBDC (digital yuan) policy and key implications. JRI Research Journal, 9(13), 1-12.

Kumhof, M., & Noone, C. (2021). Central bank digital currencies: Design principles for financial stability. Economic Analysis and Policy, 71, 553-572.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., et al. (2018). Applying lda topic modeling in communication research: Toward a valid and reliable methodology. Communication Methods and Measures, 12(2–3), 93–118.

Sander, J., Ester, M., Kriegel, H-P., & Xu, X. (1998). Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications. Data Mining and Knowledge Discovery, 2(2), 169–194.

Stoica, O., & Dornean, A. (2025). From Euro to Digital Euro: A Survey of Literature. Scientific Annals of Economics and Business, 72 (SI), 19-31.

Wang, Y., Lucey, B. M., Vigne, S., & Yarovaya, L. (2022). The effects of central bank digital currencies news on financial markets. Technological Forecasting and Social Change, 180, 121715.

Wronka, C. (2023). Central bank digital currencies (CBDCs) and their potential impact on traditional banking and monetary policy: an initial analysis. Digit Finance, 5, 613–641.

Conflict of interest

No conflict of interest.

Acknowledgements

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS - UEFISCDI, project number PN-IV-P1-PCE-2023-0679, within PNCDI IV.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used AI-assisted tools to support the screening, identification and extraction of CBDC-related speeches from the BIS central bank speeches database, as well as to assist with language editing, structural refinement and improvement of academic tone. All AI-assisted outputs were independently reviewed, verified and edited by the authors, who take full responsibility for the final content, empirical results and conclusions of this manuscript.

Appendices

Appendix A. Dictionary structure

The CFSCI is constructed using channel-level dictionaries rather than raw word counts. Each channel receives a value of one when at least one associated term appears in a speech and zero otherwise. The speech-level FSRI and FSBI are then calculated as the share of risk or benefit channels present in the speech.

Table A1 - Risk-channel dictionary used for FSRI

Channel	Indicative terms
Deposit substitution and migration	deposit; deposits; deposit substitution; deposit migration; outflow; outflows; withdrawal; withdrawals
Bank intermediation and funding	disintermediation; intermediation; bank funding; funding pressure; banks; banking sector; commercial bank; commercial banks
Credit supply and lending	credit; lending; loan; loans; credit supply; credit contraction
Liquidity and run risk	bank run; runs on banks; run risk; liquidity; liquidity risk; liquidity stress; liquidity pressures
Systemic stability	financial stability; financial instability; systemic risk; systemic; contagion; stress; crisis; vulnerability; vulnerabilities
Operational and cyber risk	operational risk; cyber; cyber risk; technology risk; privacy risk; data privacy; operational resilience

Table A2. Benefit-channel dictionary used for FSBI

Channel	Indicative terms
Payment-system resilience	payment resilience; resilience; payment system; payment systems; settlement; infrastructure
Public money and monetary sovereignty	public money; central bank money; monetary sovereignty; sovereignty; anchor; trust
Inclusion and access	financial inclusion; inclusion; access; unbanked
Efficiency, competition and innovation	efficiency; efficient; competition; innovation; interoperability; cross-border
Oversight, monitoring and transparency	oversight; monitoring; transparency; traceability; integrity
Safeguards and design features	safeguards; holding limits; tiered remuneration; limits; remuneration; intermediated; design safeguards

Appendix B. Institutional comparison for selected major central banks

An institutional comparison provides additional perspective. Because the full corpus includes speeches from many central banks and the dataset does not contain a fully harmonized institution identifier, the analysis is restricted to selected major institutions for which speeches can be assigned with reasonable confidence. These results should therefore be read as evidence for selected institutions, not as a complete jurisdictional decomposition of the corpus.

The institutional comparison uses the same IMF-channel coverage method as the annual analysis. FSRI measures the share of risk channels present in each speech, while FSBI measures the share of benefit channels present. Institutional values are averages across speeches assigned to each institution.

Table B1. Selected major institutions: CBDC financial stability communication

Institution / jurisdiction	CBDC speeches	FSRI	FSBI	NetCFSCI
ECB	197	0.667	0.712	0.045
BIS	65	0.654	0.736	0.082
Bank of England	51	0.761	0.788	0.026
Federal Reserve	44	0.769	0.712	-0.057
People's Bank of China (PBoC)	7	0.429	0.524	0.095

The ECB accounts for the largest number of clearly attributable CBDC-related speeches among the selected institutions. Its NetCFSCI is positive but moderate, suggesting that risk and benefit channels are both prominent in the ECB's communication. The BIS records the highest positive net score among institutions with a relatively substantial number of speeches, consistent with its role in framing CBDC as part of a broader payment-system and monetary architecture agenda.

The Bank of England also records a positive NetCFSCI, although its FSRI is relatively high. This is consistent with the Bank's detailed public discussion of disintermediation, holding limits and financial stability safeguards. The result suggests that these risks are not presented in isolation, but are embedded in a broader design framework.

The Federal Reserve is the only selected institution with a negative NetCFSCI. This points to a more cautious communication pattern, with risk channels receiving somewhat greater emphasis than benefit channels. The PBoC result is positive, but it is based on only seven clearly attributable speeches and should therefore be interpreted with caution.

The institutional results add nuance to this interpretation. The Federal Reserve appears more cautious than the ECB, BIS or Bank of England in the selected-institution comparison, while the BIS places relatively stronger emphasis on benefit channels, consistent with its role as a global forum for payment-system innovation and central bank cooperation. These differences suggest that CBDC communication is shaped not only by technological developments, but also by institutional mandates, domestic financial structures and policy priorities. Overall, the discussion points to a mature communication pattern in which CBDC risks remain acknowledged, but are increasingly embedded within a broader framework of resilience, public money and safeguards.