Central Bank Communication and CBDC Implementation

Fatih Kansoy*
University of Oxford

This Version: March 25, 2026

Abstract

CBDC issuance remains rare even though central banks discuss it extensively. Using 37,814 speeches from 134 jurisdictions and a scored 2015–2025 sample of 15,906 speeches from 122 jurisdictions, the analysis builds country-year measures of CBDC commitment, technology salience, and stated constraints. Public commitment tracks agenda participation but adds little information about subsequent advancement beyond research once prior project stage is controlled for. More open capital accounts and, less consistently, greater central bank independence are negatively associated with advancement beyond research. Legal-regulatory concerns are the most consistently negative speech-based correlate.

Keywords: Central bank digital currency, central bank communication, financial innovation, institutional conservatism, text as data

JEL Classification: E42, E58, G28, O33

*Department of Economics, University of Oxford. fatih.kansoy@economics.ox.ac.uk

1. Introduction

Few ideas in modern central banking have moved from technical curiosity to mainstream policy debate as quickly as central bank digital currency. In speeches, reports, and public consultations, central banks increasingly describe digital public money as part of the future monetary system. In practice, however, retail issuance remains rare. Recent BIS evidence shows broad engagement with CBDC research and experimentation, even as expectations of near-term retail issuance have softened and several jurisdictions have shifted attention towards wholesale designs or paused retail projects (BIS 2025, Kosse and Mattei 2022, 2023, Di Iorio et al. 2024). The central fact is therefore not simply slow diffusion. It is a widening gap between what central banks publicly discuss and how rarely jurisdictions advance beyond research. Existing cross-country studies measure formal project stages but cannot capture how central banks frame the constraints they face or when rhetoric and action come apart (Auer et al. 2020, Maryaningsih et al. 2022, Corbet et al. 2024, Le et al. 2023). The question is why central banks talk so much about CBDC but move so cautiously towards advanced project stages.

The question matters because retail CBDC would change the boundary between public and private money. A digital liability of the central bank that is widely available to households and firms could reshape deposit intermediation, bank funding, and competition in the financial system (Keister and Sanches 2023, Whited et al. 2023, Chiu et al. 2023, Andolfatto 2021). It also raises first-order questions about privacy and state surveillance (Sun et al. 2024), about cross-border spillovers, capital-flow volatility, and currency substitution (Ferrari Minesso et al. 2022, Lukonga 2023), and about whether existing legal frameworks even permit retail issuance in the first place (Bossu et al. 2020, Bechara et al. 2025). The puzzle of why central banks deliberate so extensively but advance so selectively is therefore not only a story about payments technology. It is a question about the political

economy of monetary institutions.

Central bank speeches enter the empirical design directly. Using a full corpus of 37,814 speeches from 134 jurisdictions, the analysis constructs a scored 2015–2025 sample of 15,906 speeches from 122 jurisdictions and builds country-year measures of CBDC commitment, technology salience, and stated constraints. The interpretation is guided by a simple institutional argument: public deliberation can be valuable even when advancement is unattractive, because discussing CBDC allows central banks to signal awareness and alignment with international norms at low cost (DiMaggio and Powell 1983, Bromley and Powell 2012). Advancement beyond research is different. It requires absorbing open-economy and financial-stability risks (Ferrari Minesso et al. 2022, Lukonga 2023), resolving legal-authority gaps (Bossu et al. 2020), managing resistance from incumbent financial actors (Pfister 2024), and expanding central bank activity into domains that sit uneasily with delegated mandates (Tucker 2018, Rouanet and Salter 2025). If the costs of advancement are high and heterogeneous while the costs of deliberation are low, the determinants of talk and action should differ systematically.

The evidence supports that prediction. In the strict ordered bridge, the pooled association between speech commitment and next-year retail CBDC stage attenuates sharply once year effects and lagged stage are included, falling from 3.7 to 0.4 percentage points in the preferred specification. In the dynamic benchmark, commitment is null while capital account openness, central bank independence, and legal-regulatory constraint language are all negatively associated with advancement beyond research. Deliberation-side equations confirm that the same institutional variables do not load negatively on public commitment: KAOPEN is effectively zero for talk but negative for action, and legal-regulatory language switches sign. The institutional constraints that deter advancement do not deter deliberation.

These findings contribute to three literatures. They add to the comparative CBDC liter-

ature by showing that the correlates of public engagement need not match the correlates of retail-stage progression (Auer et al. 2020, Maryaningsih et al. 2022, Corbet et al. 2024, Le et al. 2023). They contribute to the central bank communication literature by developing a speech-based panel that captures how central banks frame the constraints they face, and by showing that CBDC communication is closer to agenda participation than to credible advancement commitment (Gurkaynak et al. 2005, Swanson 2021, Bauer and Swanson 2023, Hansen and McMahon 2016, Blinder et al. 2024, Aruoba and Drechsel 2024). Finally, they speak to the political economy of monetary institutions by documenting a setting in which legitimacy-seeking discourse coexists with persistent cross-country variation in advancement when open-economy, legal, and mandate-related constraints are binding (Meyer and Rowan 1977, DiMaggio and Powell 1983, Bromley and Powell 2012).

2. Related Literature

The study sits at the intersection of three research strands. The first studies cross-country patterns in CBDC progress using ordinal stage outcomes. The second demonstrates that central bank communication can be measured credibly with text-as-data methods and that these measures contain information not captured by standard quantitative indicators. The third analyses the institutional, legal, and macroeconomic constraints that make monetary innovation costly and therefore potentially prone to a systematic wedge between public commitment and concrete project advancement. The contribution connects these strands by measuring CBDC deliberation and its content directly from central bank speech, and by testing whether rhetoric predicts advancement beyond research once persistence and common shocks are absorbed.

2.1 Cross-Country Evidence on CBDC Progress

The most systematic evidence on what central banks report doing comes from the BIS survey series on CBDCs and cryptoassets (Kosse and Mattei 2022, 2023, Di Iorio et al. 2024, BIS 2025). Across four waves, the number of respondents varies from 81 to 93 central banks. Engagement rates remain high while stated timelines to issuance have grown more cautious, and several jurisdictions have shifted emphasis from retail to wholesale CBDC. The surveys are informative about stated motivations and constraints but are not designed to explain why some jurisdictions translate exploration into advancement beyond research while others do not.

Alongside these institutional snapshots, an empirical literature uses cross-country regressions to explain differences in CBDC project stages. Foundational evidence links stage advancement to informality and mobile penetration (Auer et al. 2020), while later work emphasises distinct wholesale-versus-retail drivers, institutional quality, and stage-dependent legal capacity (Maryaningsih et al. 2022, Corbet et al. 2024, Le et al. 2023). Across studies, legal and political constraints recur as major barriers.

These studies share a measurement strategy that shapes both their findings and their limitations. CBDC progress is captured through ordinal stage classifications, and the empirical question is framed as what predicts whether a country reaches a higher stage. The approach is informative for identifying structural correlates of formal project activity. It cannot, however, distinguish jurisdictions that deliberate extensively but make limited formal progress from those that advance with minimal prior public engagement. Nor can it capture the content of deliberation, including what central banks publicly commit to, how prominent CBDC is within their broader modernisation agenda, or which constraints they emphasise when explaining caution. If the core puzzle is a divergence between talk and action, measuring only action leaves the divergence invisible.

2.2 Central Bank Communication as Data

A separate literature demonstrates that central bank speech contains economically meaningful information that goes beyond, and sometimes contradicts, formal policy indicators (Lucca and Trebbi 2012, Hansen and McMahon 2016, Hansen et al. 2018). High-frequency identification studies show that monetary-policy announcements operate through multiple channels and that the informational content of communication depends on context and decomposition method (Gurkaynak et al. 2005, Swanson 2021, Bauer and Swanson 2023). Recent work recovers monetary-policy shocks directly from central bank text (Aruoba and Drechsel 2024). Methodological foundations are surveyed in Gentzkow et al. (2019).

These findings establish speech as a productive empirical object, but most applications remain domestic and monetary-policy focused. Recent work extends the scope to cross-country central bank communication using large multilingual corpora and specialised central-banking language models (Silva et al. 2025, Campiglio et al. 2025, Gambacorta et al. 2024, Baumgartner and Zahner 2025).

This research builds on this frontier but differs from it in purpose. The international text-as-data papers primarily treat communication measures as outcomes, namely what central banks talk about and how sentiment evolves over time. Speech-based measures enter here as regressors for advancement beyond research, and the extraction targets not only the intensity of CBDC discussion but also the specific constraints that central bankers articulate. These measures map directly onto the theoretical mechanisms developed below. The issue is not simply what central banks talk about, but whether what they say predicts what they do, and if not, what does.

2.3 Institutional and Macroeconomic Constraints on Monetary Innovation

Several literatures identify mechanisms through which institutional and macroeconomic features could produce a wedge between CBDC deliberation and implementation. These are reviewed briefly here, and Section 3 develops them formally.

Institutional theory provides the generic logic. Organisations facing complex and contradictory demands frequently adopt symbolic policies that signal conformity with prevailing norms while decoupling formal commitments from operational practice (DiMaggio and Powell 1983, Meyer and Rowan 1977). Bromley and Powell (2012) refine this by distinguishing policy-practice decoupling (adopting policies without implementing them) from means-ends decoupling (implementing policies designed not to achieve their stated objectives). In a CBDC context, deliberation can generate legitimacy at low cost, while implementation requires absorbing material risks that vary across jurisdictions.

The most direct constraint runs through central bank independence. Institutions designed to resist short-term political pressure may also resist large-scale innovations with uncertain consequences (Rogoff 1985). Mission-boundary arguments likewise imply that expansion into retail payments and data governance can threaten the political compact supporting autonomy (Tucker 2018, Lastra 2024, Rouanet and Salter 2025). If this logic holds, independent central banks should be less likely to implement CBDC even when they deliberate about it extensively.

Capital account openness creates a structurally different constraint. In open economies, a frictionless digital instrument may facilitate rapid capital outflows during stress. Ferrari Minesso et al. (2022) show in a two-country model that CBDC amplifies international transmission of monetary and technology shocks. Lukonga (2023) documents that CBDC-induced currency substitution can impair monetary policy transmission. For emerging

markets with weak currencies and limited institutional credibility, these risks amount to threats to monetary sovereignty (Adrian and Mancini-Griffoli 2021). Because these vulnerabilities are structural features of the economy, not time-varying policy choices, the constraint should manifest as a between-country regularity, not a within-country dynamic.

On the domestic side, the disintermediation literature provides concrete reasons for caution. CBDC can crowd out deposits and alter bank funding conditions (Keister and Sanches 2023, Whited et al. 2023), but effects depend on design and market structure, and can include competition gains in concentrated banking markets (Andolfatto 2021, Chiu et al. 2023, Burlon et al. 2024). The resulting uncertainty itself can deter risk-averse institutions.

Legal authority represents a qualitatively different barrier. Unlike economic trade-offs that can be managed through design choices, the absence of statutory authority is binary. Issuance either has a legal basis or it does not. Bossu et al. (2020) find that most central bank statutes do not authorise digital currency issuance to the general public, and that prevailing monetary law does not clearly grant currency status to digital tokens. Legislative reform requires securing agreement among multiple veto players (Tsebelis 2002), a process that is slow, uncertain, and outside the central bank's direct control. BIS surveys consistently report legal authority as a primary obstacle (Kosse and Mattei 2023, BIS 2025), and Soderberg et al. (2023) and Bechara et al. (2025) document the specific gaps that remain.

Finally, what appears as CBDC inaction may in some cases reflect policy substitution instead of failure to act. Central banks may pursue modernisation through fast payment systems, wholesale tokenisation, or related infrastructures that deliver overlapping user-facing benefits (BIS 2024a, Chen et al. 2022, International Monetary Fund 2025). This substitution margin matters for interpretation because weak links between salience and retail-stage progression can reflect rational instrument choice instead of decoupling.

Together, these literatures converge on a puzzle they individually cannot resolve. Cross-country studies identify structural correlates of CBDC progress but measure only formal stages and therefore leave the deliberation dimension unobserved. Communication studies demonstrate that central bank speech carries real information but have not tested whether CBDC rhetoric predicts advancement beyond research. Theoretical work on independence, openness, disintermediation, and legal constraints identifies mechanisms that should produce divergence between talk and action, but no existing paper measures both outcomes from a common data source and tests which mechanisms bind. The theoretical framework below organises those predictions.

3. Theoretical Framework

CBDC adoption is modelled as a political economy choice in which public deliberation and implementation can systematically decouple. A central bank can obtain legitimacy by signalling engagement with the global CBDC agenda at relatively low cost, while implementation requires bearing material financial, legal, and political risks that vary sharply across jurisdictions. This asymmetry mirrors institutional accounts of isomorphic conformity and decoupling (DiMaggio and Powell 1983, Meyer and Rowan 1977, Boxenbaum and Jonsson 2008, Bromley and Powell 2012) and motivates treating talk and action as distinct outcomes with potentially different determinants.

3.1 Deliberation and Implementation as Sequential Choices

Let i index jurisdictions and t years. The central bank faces two sequential choices. First, it decides whether to engage in public CBDC deliberation (speeches, consultations, research programmes, official statements), denoted $D_{it} \in \{0, 1\}$. Second, conditional on deliberation, it decides whether to take an implementation step that moves the project into an advanced

stage (proof of concept, pilot, or launch), denoted $A_{it} \in \{0, 1\}$.

Empirically, deliberation is observed through speech-based outcomes instead of a directly observed binary D_{it}. The empirical proxy for deliberation therefore uses two operationalisations. The first is commitment intensity (0–4), which captures the level of stated intent in public communication. The second, used as a leakage check, caps commitment at 2 so that "development/pilot" and "implementation/launch" language does not drive the deliberation proxy. This mapping allows the theory to remain binary while the data use richer ordinal speech variation.

The net payoff from deliberation is

$$\Pi_{it}^{D} = B_{it}^{L} - c_D, \tag{1}$$

where c_D is a low fixed cost of participation in CBDC discourse and B_{it}^{L} captures legitimacy benefits. A central bank deliberates when $\Pi_{it}^{D} > 0$. Legitimacy benefits arise because central banks operate in a dense professional field and face isomorphic pressures to appear aligned with evolving international norms (DiMaggio and Powell 1983, Boxenbaum and Jonsson 2008, Scott 2014). These benefits are modelled as

$$B_{it}^{L} = \alpha_g \, g_t + \alpha_T \, T_{it} + \alpha_K \, K_{it}, \tag{2}$$

where g_t captures global time-varying pressure (shifts in international discourse around digital money, landmark events such as the Libra announcement or e-CNY pilot expansion), T_{it} is the central bank's technology orientation, and K_{it} is the jurisdiction's degree of international financial integration. The global component g_t is critical. It implies that deliberation can rise in a synchronised way across countries even if heterogeneous domestic constraints keep implementation sparse. Signalling benefits increase with technology orientation ($\alpha_T > 0$) because CBDC discussion aligns naturally with existing communi-

cations about financial innovation and payment modernisation. The sign on openness for deliberation is theoretically ambiguous ex ante. International integration can raise legitimacy pressure to engage, but high external exposure can also induce caution even at the discourse stage. The key feature of the deliberation payoff is that c_D is small. Research programmes, speeches, and working papers are inexpensive relative to the legitimacy they generate.

The net payoff from implementation, conditional on deliberation, is

$$\Pi_{it}^{A} = B_{it}^{E} - (1 + \lambda I_{it})\, C_{it}, \tag{3}$$

where B_{it}^{E} denotes economic and strategic benefits from payment efficiency and potential formalisation, C_{it} is the total implementation cost, I_{it} measures central bank independence, and $\lambda > 0$ captures how independence amplifies the effective weight placed on downside risks. The central bank implements when $\Pi_{it}^{A} > 0$.

Equation (3) formalises the talk-action wedge. Factors that raise legitimacy benefits can encourage deliberation even when the implementation calculus remains unfavourable, and independence can deter implementation without reducing the incentive to deliberate.

3.2 Implementation Cost Channels

Total implementation costs are decomposed into five channels that correspond to the constraint families central bankers routinely cite and that the speech-based measures are designed to capture.

$$C_{it} = C_{\text{impl}}(T_{it}) + C_{\text{stab}}(K_{it}) + C_{\text{bank}}(B_{it}) + C_{\text{legal}}(L_{it}) + C_{\text{priv}}(P_{it}). \tag{4}$$

Operational and capability costs, denoted $C_{\text{impl}}(T_{it})$, capture project execution bur-

dens, including technology build-out, cyber resilience, and the organisational fixed costs of operating customer-facing digital infrastructure. Technology capacity reduces these costs, so C_{impl} is decreasing in T_{it}. However, technology-oriented central banks can also modernise payments through fast payment systems or wholesale tokenisation instead of retail CBDC (BIS 2024a, Chen et al. 2022), which means that higher T does not unambiguously favour implementation.

Open-economy and monetary-stability costs, denoted $C_{\text{stab}}(K_{it})$, capture macroeconomic vulnerabilities amplified by capital account openness. In financially open economies, a widely accessible digital instrument can facilitate rapid portfolio shifts, currency substitution, and volatile capital flows during stress (Ferrari Minesso et al. 2022, Lukonga 2023). For emerging markets with weak currencies and limited institutional credibility, these risks amount to threats to monetary sovereignty (Adrian and Mancini-Griffoli 2021). Because these vulnerabilities are largely structural features of external exposure, the resulting deterrent is expected to appear predominantly as a cross-country regularity, not a strong within-country dynamic.

Banking-sector and intermediation costs, denoted $C_{\text{bank}}(B_{it})$, capture disintermediation risk and incumbent resistance rooted in bank funding structure and market power. CBDC can crowd out deposits, alter banks' funding mix, and in some market structures intensify competition, so the net banking effect is design-sensitive and ambiguous in sign (Keister and Sanches 2023, Whited et al. 2023, Burlon et al. 2024, Andolfatto 2021, Chiu et al. 2023).

Legal-authority and regulatory costs, denoted $C_{\text{legal}}(L_{it})$, capture constraints arising from the absence of statutory authority for issuance. Unlike the other cost channels, the legal basis is closer to a binary constraint. Where statutes do not authorise issuance, implementation requires legislative change and coordination across multiple institutional veto players (Bossu et al. 2020, Tsebelis 2002). Private-law questions about transfer fi-

nality and the treatment of token-based claims add further complexity (Bechara et al. 2025). The binary character of legal authority makes it a particularly credible blocker of implementation.

Privacy, data-governance, and political-salience costs, denoted $C_{\text{priv}}(P_{it})$, capture the costs of resolving privacy and traceability trade-offs. Because these costs depend on social preferences, institutional accountability, and credible limits on state data access, privacy discussion can enter with ambiguous sign for deliberation (Sun et al. 2024).

Economic benefits B_{it}^{E} increase with technology capacity and potentially with informal sector size if CBDC facilitates financial inclusion. The inclusion channel is contested, however. The IMF argues CBDC is "not a panacea" and many inclusion features can be delivered through simpler infrastructure (Lannquist and Tan 2023). If inclusion rhetoric is adopted for isomorphic motives instead of substantive ones, implementation should not respond.

3.3 Central Bank Independence as Risk Amplification

Central bank independence occupies a distinct position in the framework. It does not enter as a separate cost channel but conditions how the institution weights the entire cost structure through the term $(1 + \lambda I_{it})$ in equation (3). Three mechanisms motivate this specification.

Conservative-institution logic provides the first mechanism. In canonical models, delegation to an independent central bank dampens short-term political incentives and stabilisation biases (Rogoff 1985). The same institutional conservatism that protects monetary policy credibility implies a higher threshold for adopting innovations with uncertain tail risks. The conservatism that serves price stability also raises the bar for operational experimentation.

Mandate-boundary logic provides the second. Independence is sustained by a political compact. Broad discretion is legitimate when exercised within a bounded mandate. Retail CBDC expands central bank activity into retail payments, customer-facing infrastructure, and data governance, all of which invite distributional conflict and democratic accountability pressures (Tucker 2018). Lastra (2024) and Rouanet and Salter (2025) argue that such mission expansion erodes the credibility on which independence depends. An independent central bank may therefore rationally prefer deliberation, which preserves legitimacy and optionality, over implementation, which expands accountability demands.

Balance-sheet logic points in the same direction. Successful retail CBDC adoption can require a reallocation, and in some designs an expansion, of central bank liabilities. To maintain balance, the central bank may need either to increase holdings of government securities, which can blur the monetary-fiscal boundary, or to extend wholesale lending facilities to commercial banks, which can create new dependencies (Fernández-Villaverde et al. 2021). Independent central banks have strong institutional reasons to avoid either outcome.

The critical implication is that independence affects Π^A_{it} without affecting Π^D_{it}. Deliberation does not itself require balance-sheet expansion, mandate extension, or operational exposure. An independent central bank can deliberate freely about CBDC, satisfying isomorphic pressures and preserving optionality, while the same independence raises the bar for taking the next step.

3.4 Predictions

The decision rules $\Pi^D_{it} > 0$ and $\Pi^A_{it} > 0$ generate predictions about when talk and action align and when they diverge.

The theory is written in terms of the implementation choice A_{it}. Empirically, the paper

proxies that choice with advancement beyond research to stage 2, a coarser outcome that pools proof of concept, pilot, and launch.

Prediction 1 (Global pressure raises talk more than action). Because legitimacy benefits include a global component g_t, deliberation can rise in a correlated way across jurisdictions while heterogeneous domestic constraints keep implementation sparse. The informational content of speech commitment for implementation should therefore weaken once common time shocks are absorbed. Empirically, this means that the association between commitment and stage progression should attenuate under year fixed effects.

Prediction 2 (Technology predicts deliberation more than implementation). Technology orientation T_{it} raises legitimacy returns from engaging in CBDC discourse and reduces operational frictions. Its effect on implementation is ambiguous, however, because technology-oriented central banks can modernise payments through non-CBDC alternatives, which can reduce the marginal benefit of retail issuance (BIS 2024a, Chen et al. 2022). Technology salience should predict deliberation more strongly than advancement beyond research. Empirically, the salience index captures overall digital-finance discourse, which includes CBDC alongside other technology domains. The test therefore asks whether broad technology engagement, not CBDC commitment specifically, predicts deliberation and advancement beyond research differentially.

Prediction 3 (Independence deters implementation, not deliberation). Independence amplifies the weight on all implementation cost components through $(1 + \lambda I_{it})$. This reduces Π^A_{it} without affecting Π^D_{it}. Independence should be negatively associated with advancement beyond research while leaving deliberation unaffected.

Prediction 4 (Openness discourages implementation, deliberation sign is ambiguous). International financial integration raises stability-related implementation costs ($\partial C_{\text{stab}}/\partial K > 0$), so the advancement-beyond-research association should be negative. For deliberation, openness generates opposing forces. One is legitimacy pressure to engage. The other is caution under external vulnerability. The net deliberation association is therefore ambiguous ex ante. Because the vulnerability is structural, the openness constraint should appear primarily as a between-country regularity.

Prediction 5 (Hard constraints bind, rhetorical motivations may decouple). Legal uncertainty should be the most direct barrier to advancement beyond research because it is not design-solvable and requires external political action (Bossu et al. 2020). Financial stability and banking costs should affect advancement beyond research more than deliberation because disintermediation risks materialise only upon deployment (Keister and Sanches 2023). Financial inclusion rhetoric, if adopted for isomorphic motives instead of substantive ones, should fail to predict either outcome. That pattern would be consistent with means-ends decoupling (Bromley and Powell 2012).

The framework is falsifiable in joint-pattern terms. It would be inconsistent with the model if technology salience predicted implementation as strongly as deliberation, or if openness showed the same stable sign-strength pattern for deliberation as for implementation. The framework does not assert that any single factor causes CBDC adoption. It implies that the determinants of deliberation and implementation differ systematically because the two choices deliver different payoffs and expose the institution to different risks. The predictions are therefore not about whether particular factors matter for CBDC, but about whether they matter *differently* for what central banks say and what they do. Our speech-based measures are designed to test this.

4. Data and Measurement Strategy

The theoretical framework distinguishes deliberation from action-side progression and generates predictions about when the two should diverge. Testing them requires measures of what central banks do (advancement beyond research), measures of what they say and which constraints they emphasise (deliberation), and institutional and macroeconomic variables that proxy for the cost channels identified above. The resulting country-year panel spans 2015 to 2025.

4.1 Measuring Advancement Beyond Research

Advancement beyond research is measured by an ordinal CBDC stage indicator based on publicly documented project milestones. Project-stage information comes from the Atlantic Council CBDC Tracker[1] and codes the stage for country i in year t as $S_{it} \in \{0, 1, 2\}$. The categories are no formal activity (0), research or early exploration (1), and advanced stages including proof of concept, pilot, or live retail deployment (2). Consistent with BIS survey evidence on widespread engagement but limited issuance (Kosse and Mattei 2023, BIS 2025), the distribution of S_{it} is heavily concentrated at stages 0 and 1.

The stage-2 category pools heterogeneous project types. In the 2015–2022 speech-observed sample, stage-2 country-years are 58.8% retail-only, 22.1% wholesale-only, 14.7% mixed retail-wholesale, and 4.4% uncategorised from available project-type tags. Because the theoretical argument concerns retail CBDC, the baseline below recodes the outcome to a retail-only stage variable based on the tracker's retail-status field (`cbdcr_num`). Pooled-outcome variants are reported in the appendix for comparison.

[1]The tracker is available at `https://cbdctracker.org/` and is maintained by the Atlantic Council's GeoEconomics Center (accessed 12 March 2026). It is used here as a structured compilation of publicly reported CBDC project milestones.

The stage series is annual. We treat each country-year value as year-end status and define transitions only on contiguous t to $t + 1$ pairs. Because milestone timing is not harmonised at monthly frequency across jurisdictions, this annualisation can blur within-year ordering between speeches and stage-advancement announcements. The baseline therefore uses speech measures dated $t - 1$ to predict retail stage outcomes in $t + 1$, with contemporaneous variants reported in the appendix.

Stage dynamics are highly persistent. In contiguous country-years, the probability of remaining at the research stage is 0.888 and the research-to-advanced transition rate is 10.3 per cent. This sparsity is the empirical expression of the puzzle itself. Advancement beyond research is rare precisely because the constraints identified in our framework are binding. The transition frequencies motivate an empirical strategy centred on transition models and dynamic specifications (Section 5) instead of pooled level regressions.

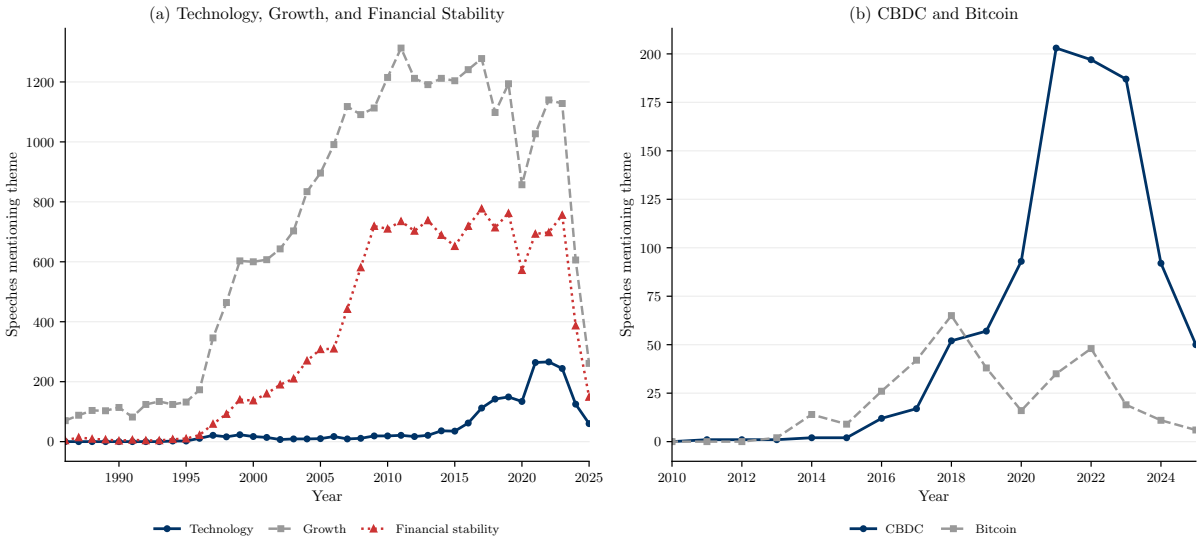
4.2 Measuring Deliberation from Central Bank Speech

Deliberation is measured using a full corpus of 37,814 speeches from 134 jurisdictions, drawn from the BIS central bank speeches download extract (BIS 2024b) and supplemented by the speech dataset compiled by Campiglio et al. (2025), which expands coverage by collecting speeches directly from central bank websites and related archives. Speeches appearing in both sources are deduplicated by jurisdiction, speech date, speaker, title, and near-exact text overlap, with the BIS record retained when duplicate metadata resolve to the same speech. The BIS extract is a curated selection provided without guarantees of completeness (BIS 2024b). Speech availability therefore varies across countries and years. The supplemented corpus mitigates coverage gaps for the analysis window, during which BIS collection became more systematic and CBDC-relevant speeches were predominantly delivered in English or with available translations. Remaining unevenness in country representation is addressed econometrically through inverse probability weighting in

Section 5.

For the scored 2015–2025 analysis window we retain 15,906 speeches from 122 jurisdictions with valid text and complete metadata (country, date, speaker). The speech-linked panel includes only those country-years with at least one observed and scored speech.

FIGURE 1: Central Bank Speech Themes, 1986-2025



Notes: Panel (a) shows annual speech counts in the full corpus for technology, growth, and financial-stability themes. Panel (b) shows annual counts for CBDC- and Bitcoin-related terms, from 2010 onward for readability.

Figure 1 places the CBDC discourse in context. Panel (a) shows that growth and financial-stability themes have dominated central bank speeches for decades, peaking at over 1,300 and 770 speeches respectively. Technology-related language was negligible before the mid-1990s but accelerated sharply from 2015 onward, reaching 266 speeches by 2022, the fastest proportional growth of any theme in the corpus despite remaining smaller in absolute count. Panel (b) narrows the lens to digital-currency terms from 2010 onward. Bitcoin references peak in 2018 and fade thereafter. CBDC-specific language follows a distinct trajectory. It rises steeply to 203 speeches in 2021 before declining by 75 per cent to 50 by 2025. The divergence confirms that central bank engagement with digital currency reflects

an institutional dynamic separate from private cryptocurrency markets.

Each speech is scored using a hybrid architecture combining keyword breadth with large language model semantic classification. The keyword component employs lexicons spanning eight technology domains (fintech, distributed ledger technology, cybersecurity, and others) to produce transparent baseline indicators of topic coverage. The LLM component (via API, GPT-5.1, zero-shot classification) applies rubric-based prompts that instruct the model to evaluate strictly on the provided text, recognise country-specific project names ("digital euro," "e-CNY") as CBDC-relevant even when the generic term is absent, and distinguish public CBDC projects from private cryptoasset regulation. The full prompt specification appears in the online appendix.

The hybrid structure addresses a cross-country measurement problem that neither component solves alone. Keyword matching is stable and auditable but yields false negatives when jurisdictions use branded project terminology. LLM classification captures such references semantically but can be noisy in ambiguous contexts. The two are reconciled through an asymmetric rule targeted at the branded-name problem. When keywords detect zero commitment but the LLM assigns strictly positive commitment, the LLM score is retained. Otherwise the two are combined by equal weighting with downward rounding through the floor operator. Among the 15,906 speeches in 2015 to 2025, 1.7 per cent trigger this override.

The scoring pipeline produces three families of measures at the speech level.

CBDC commitment is an ordinal score on a 0 to 4 scale capturing the strongest stated intent. The levels are no discussion (0), awareness (1), research or exploration (2), development or pilot (3), and implementation or launch (4). The strict baseline uses a capped variant (commitment ≤ 2) so that pilot and launch language does not enter the deliberation proxy. The full 0–4 scale is reported alongside it as a comparison specification.

Technology salience is a 0 to 100 index capturing engagement with financial technology

topics, including but not limited to CBDC. The keyword lexicon spans eight technology domains (described below), one of which covers digital currency. The index therefore measures overall digital-finance discourse orientation instead of CBDC-specific attention, though the two are not fully separable. The baseline no-floor measure also includes modest adjustments for explicit advanced-project planning, so it is best read as a broad digital-modernisation indicator rather than a pure technology-orientation proxy. In the theory, it proxies for the technology-orientation component of legitimacy benefits (Prediction 2), subject to this action-side loading. To avoid mechanically amplifying deliberation-side associations, interpretation uses the no-floor salience variant as the preferred baseline and treats the floored index as a sensitivity check reported in the appendix.

Articulated constraints are five 0 to 100 scores capturing the intensity with which a speech emphasises distinct retail-CBDC progression challenges, namely financial stability, privacy and data governance, legal and regulatory authority, financial inclusion, and competitive substitutes. These correspond to the cost channels in the action payoff (Section 3) and enable direct tests of Prediction 5.

Speech-level scores are aggregated to match the country-year advancement outcome. For CBDC commitment the within-country-year maximum is used,

$$\text{Commitment}_{it}^{\max} = \max_{s \in (i,t)} \text{Commitment}_s,$$

because a single public statement announcing a pilot or launch constitutes a revealed commitment for that year even if other speeches from the same jurisdiction are silent on CBDC. For technology salience and constraint scores, the baseline uses the simple mean across speeches in the country-year, with robustness variants based on the median, maximum, and fixed-k random speech resampling. Section 5 reports that all aggregation variants produce substantively identical conclusions.

TABLE 1: Schematic Speech Illustrations for the Measurement Rubric

Category	Case	Excerpt (abridged)	Final	KW	LLM
Panel A. Technology Salience (digital-finance discourse orientation)					
Very low (0-20)	ECB 2016	"…price stability as the ECB primary objective"	5	0	10
Medium (41-60)	ECB 2023	"…monitoring stablecoins, instant payments, and distributed ledger applications in the euro area"	55	55	55
Panel B. CBDC Commitment (stated intent)					
None (0)	IRL 2019	"…consumer protection …regulatory coordination"	0	0	0
Research (2)	ECB 2021	"…whether to issue a digital euro …complement to cash"	2	0	2
Development (3)	NPL 2023	"…working on the introduction of …CBDC in a phased manner …"	3	4	3

Notes: Final is the hybrid score, KW the keyword-only score, and LLM the model score. The excerpts are schematic, shortened, and lightly normalised for readability; they are not used in validation or estimation. In Panel B, the ECB example illustrates the asymmetric rule. When KW = 0 but branded terminology indicates CBDC activity, the final score follows the LLM score.

Table 1 provides schematic illustrations of the scoring logic and the role of the hybrid rule. The examples are based on corpus patterns and abbreviated for presentation. The ECB "digital euro" case (Panel B) uses no generic CBDC keyword yet expresses research-stage engagement. The LLM component captures this, and the asymmetric rule prevents systematic under-scoring of jurisdictions that rely on branded terminology.

A fixed GPT-5.1 run on a random 200-speech validation sample scored under the same

rubric as the full corpus returned successful, parseable outputs for every speech. A completed single-coder full-text benchmark on a 100-speech subset coded with the same 0–4 commitment rubric yields the summary results reported in Table 2.

The commitment benchmark provides the strongest human-reference evidence for the commitment scale. The 100-speech commitment benchmark reports 83% exact and 97% adjacent agreement for human versus GPT-5.1, and 84% exact and 94% adjacent agreement for human versus the hybrid score. Separate zero-versus-nonzero error rates show that GPT-5.1 has a 3.6% false-positive rate and a 5.9% false-negative rate, while the hybrid score lowers the false-positive rate to 1.2% but raises the false-negative rate to 29.4% on the benchmarked subsample. That trade-off is by design: the hybrid rule is intended to reduce false positives in the full corpus, where they would contaminate country-year aggregation, even at the cost of higher false negatives in the benchmarked subsample. A smaller 60-speech legal-intensity check yields 27% exact and 87% adjacent agreement, with weighted $\kappa = 0.668$. The 27% exact rate is low in isolation, but exact agreement is a stringent criterion on a five-bin ordinal scale (0, 1–20, 21–40, 41–60, 61–100); for this application, adjacent agreement is more informative because a one-bin difference on the mapped 0–100 scale does not materially change the substantive coding. Technology salience and the remaining four articulated-constraint channels do not have human-reference coding in the present analysis and are therefore evaluated through robustness and generated-regressor diagnostics. In the 2015–2025 scored speech sample, the final hybrid commitment differs from the raw LLM commitment in 308 of 15,906 speeches (1.9%). These are floor-based combination cases where keyword and LLM scores are both positive but unequal. Detailed human-reference benchmark tables and override counts are reported in the online appendix.

TABLE 2: Validation Summary (Human-Reference Diagnostics)

Measure	*N*	Exact agreement	Adjacent agreement	Weighted κ (quad.)
Commitment (human vs GPT-5.1)	100	0.83	0.97	0.787
Commitment (human vs hybrid)	100	0.84	0.94	0.706
Legal intensity (human vs pipeline level)	60	0.27	0.87	0.668

Notes: Commitment benchmarking uses a 100-speech single-coder subsample; legal benchmarking uses a 60-speech CBDC-relevant subsample. Adjacent agreement is defined as absolute category difference ≤ 1. For zero-versus-nonzero commitment classification, GPT-5.1 has false-positive and false-negative rates of 3.6% and 5.9%; the corresponding hybrid rates are 1.2% and 29.4%.

Because the speech measures are estimated rather than directly observed, using them as regressors introduces a generated-regressor problem. Standard errors that ignore first-stage estimation uncertainty may be too small (Pagan 1984, Murphy and Topel 1985), and Battaglia et al. (2025) show that the bias for AI-generated variables takes a specific form distinct from classical attenuation. Appendix Tables A3 and A4 report uncertainty-propagated exercises for the progression-hazard commitment estimate and the full speech-regressor dynamic benchmark. In the progression hazard, propagating first-stage uncertainty does not alter the commitment null (generated-regressor estimate 0.019 versus baseline 0.050). In the dynamic benchmark, uncertainty-propagated estimates preserve the qualitative pattern that salience and most constraint channels are null while legal language remains negative.

4.3 Institutional and Macroeconomic Variables

The theoretical framework identifies central bank independence, capital account openness, and legal-institutional quality as determinants of the talk-action wedge. Each is measured with standard sources.

Central bank independence is measured with the CBIE index from Romelli (2024), which covers 155 countries through 2023 with annual updates incorporating 370 documented reforms. The sample mean is 0.68 (range 0.26 to 0.91). In the framework, higher independence amplifies the weight on action-side costs and should deter advancement beyond research more than deliberation (Prediction 3).

International financial integration is measured by the Chinn-Ito KAOPEN index (Chinn and Ito 2008), available through 2022. In the framework, openness increases deliberation incentives through international integration pressures but raises action-side costs through the stability channel (Prediction 4). Because the vulnerability is structural, the Mundlak decomposition in our empirical design allows us to test whether the openness effect operates between countries or within countries over time.

Government effectiveness from the World Bank Worldwide Governance Indicators captures broader state capacity. As controls for economic and technological context, we include GDP per capita (IMF, current prices in USD), the Global Innovation Index (WIPO, through 2024) for technological capacity, and mobile cellular subscriptions per 100 inhabitants (ITU, through 2023) for digital infrastructure. Definitions and summary statistics are reported in Appendix Table A1.

For an auxiliary substitution check, an annual FPS-live indicator is coded from the World Bank Project FASTT Global Tracker[2] and missing tracker launch years are supplemented with official central-bank or system-operator launch sources archived in the replication files. The indicator equals one from the first year in which a live retail fast-payment system is recorded and zero otherwise. It is used only in the FPS-control specification reported below.

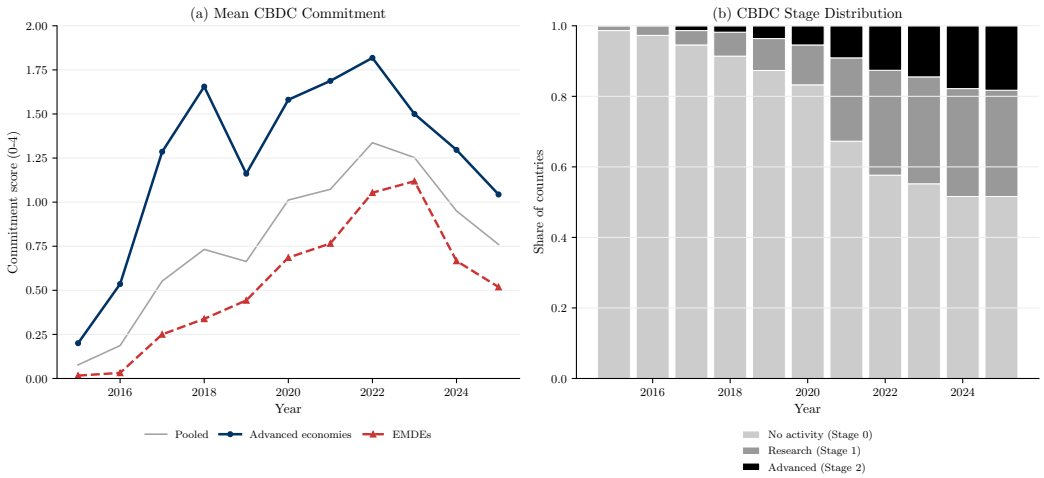
[2] https://fastpayments.worldbank.org/global-tracker, accessed 12 March 2026.

4.4 Panel Construction and Estimation Samples

The master panel contains 2,427 country-years spanning 2015 to 2025. The empirical design uses four nested samples. The full speech corpus contains 37,814 speeches from 134 jurisdictions. The scored 2015–2025 speech sample contains 15,906 speeches from 122 jurisdictions. After merging annual stage data and speech availability, the speech-observed country-year panel contains 947 country-years across 121 countries; one jurisdiction in the scored speech sample does not enter this merged panel because it lacks a matched annual stage observation after the speech-stage merge. The strict estimation sample contains 460 country-years across 82 countries after intersecting speech availability, contiguous retail transitions, lag structure, and covariate completeness.

The core empirical specifications focus on transitions in CBDC stage. From contiguous t to t+1 observations we construct two risk sets. The 0-to-1 risk set (entry into research) contains 1,728 country-years, and the 1-to-2 risk set (progression to advanced stages) contains 330 country-years. Requiring complete covariate coverage yields estimation samples of 406 country-years for entry transitions and 104 country-years for progression transitions, both covering 2015 to 2022 (the intersection of the risk sets with KAOPEN availability). The small progression sample reflects the empirical fact that few jurisdictions move beyond research, a constraint that is inherent to the phenomenon rather than to the data collection.

FIGURE 2: Evolution of Speech-Based CBDC Measures, 2015-2025



Notes: Panel (a) shows annual mean commitment in the speech-observed panel, by development group and pooled. Panel (b) shows annual country shares in stages 0, 1, and 2 in the transition panel.

Figure 2 summarises the talk-action puzzle. Panel (a) plots mean commitment by development group. Pooled commitment peaks at 1.34 in 2022, then declines 43 per cent to 0.76 by 2025, with the post-2022 retreat steeper in emerging market and developing economies (EMDEs, 50.8 per cent) than in advanced economies (AEs, 42.6 per cent). Panel (b) shows the annual stage distribution across countries. While rhetoric rises and falls in Panel (a), the stage profile in Panel (b) shifts slowly, with stage 2 reaching only 18.3 per cent of countries by 2025. Advancement moves much more slowly than rhetoric. Appendix Figure A1 provides the country-level descriptive scatter and shows that high average commitment is neither necessary nor sufficient for reaching an advanced CBDC stage.

5. Empirical Strategy and Results

The framework in Section 3 implies that deliberation and action-side progression are distinct choices with partially overlapping determinants and different exposure to persistence

and global common shocks. The empirical analysis therefore separates apparent speech-stage associations from persistence and common shocks, then asks what speech adds beyond structural predictors. It begins with stage persistence, samples, and estimands, then turns to the ordered bridge, a dynamic benchmark and within-between decomposition, forecasting comparisons, and the limited within-country evidence from transition hazards.

5.1 Persistence, Samples, and Estimands

Two features of the data motivate the estimation strategy. Stage persistence is strong. Among countries at the research stage, the probability of remaining at stage 1 in the subsequent year is 0.888, while the research-to-advanced transition rate is 10.3 per cent. Stage 0 is similarly persistent, with a 91.8 per cent retention rate. In addition, global diffusion pressures create common shocks that can mechanically correlate speech measures with stage outcomes if year effects are omitted. Together, these features mean that pooled regressions will be dominated by the between-country cross-section and by slow-moving global trends. Those patterns can generate associations that may not reflect within-country dynamics.

These concerns are addressed through two complementary estimation approaches.

Entry and progression transition dummies are defined from contiguous year pairs:

$$Y^{01}_{i,t+1} = \begin{cases} 1 & \text{if } S_{it} = 0 \text{ and } S_{i,t+1} \geq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$Y^{12}_{i,t+1} = \begin{cases} 1 & \text{if } S_{it} = 1 \text{ and } S_{i,t+1} = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Each outcome therefore equals one when the relevant transition occurs and zero otherwise. Linear probability models with country and year fixed effects,

$$Y^{k}_{i,t+1} = \alpha_i + \tau_t + \beta' X_{it} + \varepsilon_{it}, \quad k \in \{01, 12\}, \tag{5}$$

exploit only within-country variation over time. The entry risk set contains 406 country-years across 83 countries after requiring complete covariates. The progression risk set contains 104 country-years across 46 countries. Both cover 2015 to 2022, bounded by KAOPEN availability. The progression sample is the binding power constraint throughout the analysis.

Ordered probit models are estimated for next-year retail stage $S^{R}_{i,t+1}$ on a common sample, progressively adding year fixed effects and lagged retail-stage dummies. The resulting ordered bridge isolates how much of the apparent talk-action association is attributable to common time shocks and persistence (Prediction 1). Because ordered probit with many country fixed effects is inconsistent in short panels, the ordered specifications omit country fixed effects. Within-country evidence is provided separately by the transition hazards. Since ordered-probit coefficients are not directly comparable across nested models, attenuation is evaluated in model-implied probability space. Coefficient movements are reported only as descriptive context.

5.2 The Ordered Bridge

Tables 3 and 4 report the ordered bridge for the strict retail-only baseline. Every specification uses the next-year retail CBDC stage as the outcome, speech regressors dated $t-1$, and a common sample of 460 country-years across 82 countries. The preferred bridge uses commitment capped at 2 so that pilot and launch language does not enter the deliberation proxy; the full 0–4 commitment scale is shown alongside it for comparison.

TABLE 3: Ordered Bridge under the Strict Retail-Only Baseline

	Pooled	Year FE	Dynamic + Year FE
Panel A. Preferred deliberation proxy (commitment capped at 2)			
Commitment$_{t-1}$ (capped at 2)	0.288***	0.139	0.129
	(0.001)	(0.157)	(0.284)
Lagged retail stage 1$_t$	–	–	2.351***
	–	–	(<0.001)
Lagged retail stage 2$_t$	–	–	9.512***
	–	–	(<0.001)
Panel B. Full commitment comparison (0–4 scale)			
Commitment$_{t-1}$ (full 0–4 scale)	0.241***	0.119	0.129
	(<0.001)	(0.104)	(0.161)
Lagged retail stage 1$_t$	–	–	2.367***
	–	–	(<0.001)
Lagged retail stage 2$_t$	–	–	9.514***
	–	–	(<0.001)
N		460	
Countries		82	

Notes: Ordered probit models for next-year retail CBDC stage on the strict retail-only common sample (460 country-years, 82 countries, 2016–2022). Speech regressors are lagged to $t-1$; entries are coefficients with country-clustered p-values in parentheses. Panel A uses the preferred capped commitment proxy; Panel B restores the full 0–4 scale. All columns also include salience, KAOPEN, CBIE, GDP per capita, mobile subscriptions, innovation index, and government effectiveness; the dynamic column additionally includes lagged retail-stage dummies and year fixed effects. Because ordered-probit coefficients are not directly comparable across nested models, attenuation is interpreted with the probability-scale results in Table 4.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Because ordered-probit coefficients are subject to latent-scale normalisation, coefficient magnitudes are not directly comparable across nested specifications. Attenuation is therefore evaluated in probability space. Table 4 reports the average discrete change in $\Pr(S^R_{t+1} = 2)$ from a one-point increase in commitment under each bridge step.

TABLE 4: Probability-Scale Ordered Bridge under the Strict Retail-Only Baseline

Specification	Baseline $\Pr(S^R_{t+1} = 2)$	$\Delta \Pr(S^R_{t+1} = 2)$	Bootstrapped SE	95% CI
Preferred capped commitment				
Pooled	0.120	0.037	0.015	[0.012, 0.068]
Year FE	0.121	0.014	0.012	[-0.004, 0.043]
Dynamic + Year FE	0.124	0.004	0.004	[-0.003, 0.013]
Full commitment comparison				
Pooled	0.120	0.048	0.016	[0.018, 0.082]
Year FE	0.121	0.019	0.015	[-0.004, 0.054]
Dynamic + Year FE	0.124	0.006	0.005	[-0.004, 0.017]

Notes: Model-implied averages from Table 3 on the strict retail-only common sample (460 country-years, 82 countries, 2016–2022). $\Delta \Pr(S^R_{t+1} = 2)$ is the average discrete change from a one-point increase in commitment, evaluated at observed covariates. Preferred rows use capped commitment; comparison rows use the full 0–4 scale. Standard errors and confidence intervals use a country-cluster bootstrap with 500 resamples.

FIGURE 3: The Ordered Bridge under the Strict Retail-Only Baseline

Notes: Points show the average discrete effect of a one-point commitment increase on $\Pr(S^R_{t+1} = 2)$, in percentage points, with 95% confidence intervals. Filled markers use capped commitment; hollow markers use the full 0–4 scale. Intervals use a country-cluster bootstrap.

Figure 3 visualises the strict-baseline bridge. The preferred capped proxy attenuates from 3.7 percentage points in the pooled model to 1.4 points with year effects and 0.4 points in the dynamic specification. The full commitment comparison yields 4.8, 1.9, and 0.6 percentage points. In both cases, most attenuation occurs when year effects enter, and the residual dynamic effect is close to zero.

Prediction 1. Common shocks and persistence undermine the talk-action link. In probability space, commitment's predictive content attenuates sharply across bridge steps. Under the preferred capped proxy, a one-point increase in commitment raises $\Pr(S^R_{t+1} = 2)$ by 3.7 percentage points in the pooled model, 1.4 points with year effects, and only 0.4 points once lagged retail stage is included. Under the full commitment scale, the corresponding

sequence is 4.8, 1.9, and 0.6 percentage points. The year-effects step absorbs most of the pooled association in both cases. This pattern indicates that common global shocks and synchronised discourse account for most of the apparent talk-action link. Appendix event-study diagnostics further indicate that attenuation should be read as a composite of common-shock absorption and pre-existing trajectory differences rather than a single mechanism.

5.3 Primary Dynamic Benchmark and Within-Between Decomposition

Table 5 reports the reduced dynamic benchmark on the same strict sample. The preferred column keeps lagged commitment and lagged legal-regulatory language because the theoretical framework maps them most directly to the implementation cost structure: commitment tests the talk-action gap itself, while legal language captures the hard-constraint prediction. These are also the two speech measures for which human-reference benchmark evidence is reported above. Commitment remains statistically indistinguishable from zero (0.103, clustered $p = 0.378$). By contrast, capital account openness (–0.228, clustered $p = 0.011$), central bank independence (–1.114, clustered $p = 0.027$), and legal-regulatory language (–0.014, clustered $p = 0.015$) are all negative in the benchmark. Restoring the full 0–4 commitment scale or adding the FPS control leaves this hierarchy unchanged. Extending the speech block to salience and four additional constraint channels without human-reference coding preserves the qualitative sign-significance hierarchy, although coefficient magnitudes shift, especially for commitment and CBIE; Appendix Table A2 reports the full strict-lag comparator.

TABLE 5: Primary Dynamic Benchmark with Theory-Prioritised Speech Measures

	Primary capped spec.	Full commitment	+ FPS control
Lagged retail stage 1_t	2.336***	2.348***	2.334***
	(<0.001)	(<0.001)	(<0.001)
Lagged retail stage 2_t	9.355***	9.316***	9.402***
	(<0.001)	(<0.001)	(<0.001)
Commitment$_{t-1}$ (capped at 2)	0.103	–	0.104
	(0.378)	–	(0.373)
Commitment$_{t-1}$ (full 0–4 scale)	–	0.105	–
	–	(0.236)	–
KAOPEN$_t$	-0.228**	-0.228**	-0.226**
	(0.011)	(0.011)	(0.014)
CBIE$_t$	-1.114**	-1.091**	-1.112**
	(0.027)	(0.031)	(0.026)
Legal-regulatory language$_{t-1}$	-0.014**	-0.014**	-0.013**
	(0.015)	(0.016)	(0.016)
FPS live$_t$	–	–	-0.062
	–	–	(0.671)
N		460	
Countries		82	

Notes: Dynamic ordered probit models for next-year retail CBDC stage on the strict retail-only sample (460 country-years, 82 countries, 2016–2022). Speech regressors are lagged to $t-1$; all columns include lagged retail-stage dummies and year fixed effects. The preferred column keeps commitment capped at 2 and legal-regulatory language because they map most directly to the theory's implementation margin and both have human-reference benchmark evidence. The middle column restores the full 0–4 commitment scale, and the right column adds an FPS-live indicator. GDP per capita, mobile subscriptions, innovation index, and government effectiveness are included but omitted from display. Entries are coefficients with country-clustered p-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.4 Deliberation-Side Evidence

Table 6 asks whether the same covariates load similarly on talk. On the commitment margin, they do not. KAOPEN is effectively zero once year effects and lagged commitment enter (0.020, $p = 0.824$); CBIE is negative but imprecise (−0.545, $p = 0.185$), at roughly half the action-side magnitude; and legal-regulatory language is weakly positive rather than negative. On the broader salience margin, the lagged outcome is highly persistent, while KAOPEN and CBIE are again not statistically distinguishable from zero in the dynamic specification.

This cross-equation comparison sharpens the paper's main claim. The stable negative loadings are concentrated on advancement beyond research, not on public commitment. That is the core pattern behind Predictions 3–5. Prediction 2 also looks stronger in joint form than in the original one-equation benchmark: broader digital-modernisation discourse is persistent in its own right, but the negative structural loadings are concentrated on action rather than on talk. The final column of Table 5 likewise shows that the FPS control does not materially alter the openness result. The KAOPEN coefficient moves only from −0.228 to −0.226, while the FPS indicator itself is near zero (clustered $p = 0.671$). The available evidence therefore does not support the view that the openness gradient is simply a payment-maturity proxy.

TABLE 6: Deliberation-Side Evidence on the Strict Action Sample

	Commitment (capped at 2)			Technology salience		
	Pooled	Year FE	Dynamic + Year FE	Pooled	Year FE	Dynamic + Year FE
Lagged outcome$_{t-1}$	–	–	0.367***	–	–	0.276***
	–	–	(<0.001)	–	–	(<0.001)
KAOPEN$_t$	-0.057	0.018	0.020	-0.041	0.233	0.345
	(0.552)	(0.859)	(0.824)	(0.973)	(0.849)	(0.715)
CBIE$_t$	-0.454	-0.716	-0.545	-11.708	-11.949	-8.834
	(0.324)	(0.139)	(0.185)	(0.115)	(0.118)	(0.150)
Legal-regulatory language$_t$	0.002	0.007	0.007	–	–	–
	(0.706)	(0.185)	(0.126)	–	–	–
N	460			460		
Countries	82			82		

Notes: All columns use the same strict action sample as Table 5 (460 country-years, 82 countries, 2016–2022). Commitment columns use ordered probit; salience columns use OLS. Year-FE columns add year fixed effects; dynamic columns additionally add the lagged dependent variable. Commitment equations include current-year legal-regulatory language; salience equations focus on the structural covariates to avoid compositional overlap with the broad topic-intensity index. GDP per capita, mobile subscriptions, innovation index, and government effectiveness are included in every specification but omitted from display. Entries are coefficients with country-clustered *p*-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 places the key speech and structural variables on a within-between footing using the same reduced speech block as Table 5. The KAOPEN association is predominantly between-country. The country-mean component is negative and statistically distinguishable from zero, while the within-country deviation is essentially nil. CBIE is negative in the benchmark but not localised in the decomposition. Legal-regulatory language is likewise negative and statistically distinguishable from zero at the between-country level, while the within-country deviation is essentially nil. The capped commitment measure shows a

positive within-country deviation and a weak positive between-country component, which is consistent with the bridge result that commitment mostly tracks trajectory instead of providing stable cross-sectional action signal. The strong undivided CBIE coefficient and the weak decomposition are not contradictory. CBIE moves slowly within countries over this sample, so splitting limited variation into within and between components leaves the pooled benchmark precise but the separate components imprecise.

TABLE 7: Within-Between Decomposition Matching the Reduced Dynamic Benchmark

	Coefficient	*p*-value
Commitment		
Within (deviation from country mean)	0.022**	0.046
Between (country mean)	0.039*	0.078
Capital account openness		
Within (deviation from country mean)	0.004	0.855
Between (country mean)	-0.023**	0.016
Central bank independence		
Within (deviation from country mean)	-0.182	0.477
Between (country mean)	-0.010	0.869
Legal-constraint language		
Within (deviation from country mean)	0.000	0.731
Between (country mean)	-0.002**	0.013

Notes: Mundlak within-between decomposition on the strict retail-only sample (460 country-years, 82 countries, 2016–2022), using the same reduced speech block as Table 5: capped commitment and lagged legal-regulatory language. This makes the decomposition directly comparable to the primary dynamic benchmark. *p*-values use country-clustered standard errors from the linear-probability Mundlak approximation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.5 Prediction Exercise

The bridge establishes limited incremental predictive content in coefficient space. Table 8 examines the same question in forecasting space on the strict sample. Model A uses lagged retail-stage dummies, KAOPEN, CBIE, macro controls, and year fixed effects. Model B adds the full lagged speech block, namely commitment, salience, and all five constraint channels, to give speech measures every possible advantage in the forecasting comparison. The preferred out-of-sample design is leave-one-country-out cross-validation because it preserves common time effects while asking whether speech improves country ranking beyond persistence and structural controls.

TABLE 8: Predictive Performance under Baseline and Speech-Augmented Models

Panel A. AUC

Scheme	*N*	Pos.	Model A	Model B	*B* – *A*	95% CI
In-sample fit	460	55	0.962	0.968	0.006	–
Leave-one-country-out	333	52	0.945	0.948	0.003	[-0.006, 0.014]
Rolling year holdout	333	52	0.933	0.926	-0.007	[-0.022, 0.010]

Panel B. Brier score

Scheme	*N*	Pos.	Model A	Model B	*B* – *A*	95% CI
In-sample fit	460	55	0.0340	0.0326	-0.0014	–
Leave-one-country-out	333	52	0.0445	0.0436	-0.0009	[-0.0039, 0.0018]
Rolling year holdout	333	52	0.0363	0.0437	0.0074	[0.0011, 0.0165]

Panel C. Rank accuracy

Scheme	*N*	Pos.	Model A	Model B	*B* – *A*	95% CI
In-sample fit	460	55	0.642	0.666	0.024	–
Leave-one-country-out	333	52	0.716	0.720	0.004	[-0.135, 0.048]
Rolling year holdout	333	52	0.799	0.666	-0.133	[-0.236, 0.000]

Notes: Model A includes lagged retail-stage dummies, KAOPEN, CBIE, GDP per capita, mobile subscriptions, innovation index, government effectiveness, and year fixed effects. Model B adds the full lagged speech block: commitment, salience, and the five lagged constraint-language measures. *N* is evaluated country-years; Pos. is positive outcomes. *B* – *A* is Model B minus Model A; out-of-sample 95% intervals are percentile-bootstrap intervals. Leave-one-country-out is preferred because it preserves year effects. Rolling-year holdout is a stricter temporal check because year effects cannot be estimated for unseen years. Lower Brier scores are better.

The evidence suggests not. In leave-one-country-out evaluation, adding speech moves AUC from 0.945 to 0.948 and lowers the Brier score from 0.0445 to 0.0436, but bootstrap intervals for the AUC, Brier, and rank-accuracy differences all include zero. Because stage advancement is rare and lagged stage is highly persistent, the baseline AUC is mechanically high in a way that overstates how much genuinely comparable-country discrimination the model provides. In this setting the Brier score and rank accuracy are more informative diagnostics because they penalise probability calibration and ordering errors instead of class separation alone. The conclusion is unchanged across metrics. Speech adds little stable predictive content once persistence and structural controls are in the model. The rolling-year holdout comparison is harsher and, if anything, worsens predictive performance when speech is added. The forecasting result therefore sharpens the bridge interpretation. Once lagged stage and structural covariates are in the model, the lagged speech block adds little stable out-of-sample information about advancement beyond research.

5.6 Within-Country Evidence on Transition Hazards

The ordered bridge and prediction exercise are informative about attenuation and forecasting performance, but they do not exploit within-country identification. The transition hazards in equation (5) provide that evidence.

No coefficient in either column achieves conventional significance. In the entry model (406 country-years), estimates are near zero for both speech and institutional variables. In the progression model (104 country-years, 18 observed 1→2 events), point estimates for CBIE (–1.303) and commitment (0.050) remain imprecise.

A simulation-based power exercise on this exact progression design (104 country-years, 46 countries, 18 transitions; 1,000 Monte Carlo draws with country-clustered inference) implies an 80% minimum detectable effect of about 17 percentage points for a one-point commitment shift. The ordered bridge and prediction exercise therefore provide sharper

evidence on the talk-action question than the fixed-effects hazard nulls.

TABLE 9: Transition Hazards with Country and Year Fixed Effects

	Entry (0→1)	Progression (1→2)
Commitment$_t$	0.002	0.050
	(0.936)	(0.396)
Salience$_t$	0.000	-0.006
	(0.768)	(0.375)
KAOPEN$_t$	-0.037	-0.043
	(0.600)	(0.858)
CBIE$_t$	0.380	-1.303
	(0.799)	(0.541)
Legal constraint$_t$	–	0.000
		(0.967)
N	406	104
Countries	83	46

Notes: Linear probability models with country and year fixed effects; p-values are in parentheses and standard errors are country-clustered. The progression column has 104 country-years and 18 positive transitions, so confidence intervals are wide. Full controls match Section 4.

5.7 Identification Limits

Additional robustness diagnostics are reported in the appendix and the online appendix. Across these exercises, commitment continues to attenuate sharply once persistence and common shocks are absorbed, KAOPEN remains the most stable action-side correlate, and CBIE and legal-regulatory language remain benchmark-consistent but more specification-dependent.

Three sets of diagnostics bound the scope for causal interpretation. Mechanism interaction terms (CBIE × financial depth, CBIE × legal constraint, KAOPEN × financial depth) are individually insignificant in the dynamic ordered specification ($p > 0.20$ for all three), which prevents attribution of the independence or openness associations to any single theoretical channel. Event-study pre-trend tests do not reject for the progression transition ($p = 0.090$) but reject for commitment as an outcome ($p = 0.041$), which indicates that countries whose commitment rises may already differ on pre-existing trajectories. This implies that part of the ordered-bridge attenuation may reflect pre-existing trajectory differences in addition to common shocks. Put differently, the bridge attenuation likely captures a composite of common-shock absorption and trajectory-based selection, not a single mediation channel. Appendix Table A5 reports the joint pre-trend tests, and the detailed interaction coefficients are reported in the online appendix.

A related alternative is simple time-to-build. If early commitment merely reflects projects that need time to mature, early high-commitment countries should transition more often in later years. Appendix Table A6 points in that direction descriptively: among speech-observed stage-1 country-years, the late-period transition rate is 38.5% for countries with early mean commitment of at least one, compared with 8.7% for the rest. But the corresponding interaction is imprecise in both the unadjusted and adjusted LPMs ($p = 0.238$ and $p = 0.316$). The maturation interpretation is therefore plausible, but the available panel remains too small to distinguish it cleanly from persistent talk-action decoupling.

Taken together, the ordered bridge, dynamic benchmark, and robustness exercises point to stable regularities: institutional constraints predict advancement beyond research while speech commitment does not, once persistence and common shocks are absorbed.

6. Discussion

The results point to a systematic wedge between deliberation and advancement beyond research. Measures extracted from public central bank communication track engagement with the CBDC agenda and broader digital-modernisation discourse, but they provide limited incremental country-specific information about whether a jurisdiction moves from research to an advanced project stage once persistence and common time shocks are absorbed. By contrast, institutional and structural constraints remain relevant in asymmetric ways. Openness is the strongest and most stable correlate, while independence and legal language are benchmark-consistent but more specification-dependent.

6.1 Rhetorical Convergence and the Limits of Speech as a Predictor

The strict ordered bridge implies that the pooled association between speech commitment and next-year retail CBDC stage largely reflects two features of the environment that the adoption literature has increasingly emphasised but not always modelled explicitly, namely strong persistence in project status and common global shocks. Under the preferred baseline, the commitment effect falls from 3.7 percentage points in the pooled model to 1.4 points with year effects and 0.4 points once lagged retail stage is included. The full commitment scale yields the same qualitative sequence. The Mundlak decomposition points in the same direction. The between-country commitment component is weak, while the within-country deviation is modest and does not overturn the bridge result.

Central bank speech remains informative, but in CBDC-era speeches the relevant information concerns legitimacy, positioning, and optionality rather than near-term execution. Institutional accounts of organisational fields, where professional norms and peer effects drive rhetorical convergence, predict this pattern (DiMaggio and Powell 1983). In such

settings, the incentives to speak and the incentives to build are not symmetric. Public deliberation can satisfy external expectations and preserve policy flexibility, while advancement beyond research exposes the institution to hard constraints and operational risks. The evidence equally fits the decoupling mechanisms highlighted by Bromley and Powell (2012), where adoption involves complex means-ends trade-offs and organisations can adopt the language and formal trappings of action without crossing the feasibility threshold required for deployment.

For the growing text-as-data literature in central banking, one implication deserves emphasis. In canonical monetary-policy contexts, textual measures can reveal committee information, forecast policy actions, and recover monetary-policy shocks (Lucca and Trebbi 2012, Hansen and McMahon 2016, Aruoba and Drechsel 2024, Bauer and Swanson 2023). In the CBDC context, the same empirical strategy yields a different result. Speech is highly informative about engagement and framing, yet contributes little incremental predictive content for progression beyond research once persistence and common shocks are absorbed. The prediction exercise points in the same direction. Adding the lagged speech block does not materially improve leave-one-country-out AUC, Brier score, or rank accuracy relative to a model with lagged stage and structural controls alone. CBDC communication appears closer to agenda participation than to credible project-advancement commitment.

6.2 Institutional Constraints and the Advancement Threshold

Independence and openness are aligned with advancement beyond research in the strict dynamic benchmark, while commitment-side deliberation is not. Legal-regulatory constraint language is likewise negative in the action equation but not in the commitment equations. The theory motivates this pattern. Advancement beyond research bundles financial-stability exposure, mandate and accountability risks, and balance-sheet and

operational consequences that deliberation does not.

A negative association between central bank independence and advancement beyond research is consistent with multiple channels developed in the monetary-institutions literature. Because the Mundlak decomposition does not localise CBIE within or between countries (Table 7), the CBIE association is consistent with conservatism in the face of uncertain tail risks (Rogoff 1985), mandate-boundary protection (Tucker 2018, Lastra 2024, Rouanet and Salter 2025), and reluctance to create fiscal-adjacent balance-sheet dependencies (Fernández-Villaverde et al. 2021). Even in pooled AE/EMDE interaction tests, the CBIE×AE contrast is imprecisely estimated (Online Appendix Table OA1). We therefore interpret independence as a higher institutional threshold for retail-facing innovation with uncertain distributional and political consequences rather than as a sharply identified subgroup effect.

The openness association reflects a structural exposure pattern. The within-between decomposition is consistent with openness operating primarily as a cross-sectional correlate rather than a within-country transition driver, which matches the theoretical claim that the relevant risks are tied to durable macro-financial structure. Open-economy models highlight how new forms of central bank money can alter spillover and currency-substitution dynamics (Ferrari Minesso et al. 2022, Lukonga 2023). These mechanisms explain why a jurisdiction that is persistently financially open may rationally remain at the research stage even while participating actively in the global CBDC discourse.

An alternative reading is policy substitution. Jurisdictions with sophisticated payment infrastructure may rationally pursue modernisation goals through fast payment systems, settlement upgrades, or tokenised wholesale experiments instead of advancing beyond research through retail CBDC issuance. BIS comparative work explicitly frames retail CBDC and fast payment systems as potentially overlapping instruments whose desirability depends on ecosystem maturity and policy priorities (BIS 2024a). IMF policy guidance like-

wise emphasises that CBDC should be positioned within a broader payments environment and evaluated against alternatives with similar objectives (International Monetary Fund 2025). The FPS control offers a more direct test of this interpretation. Adding an annual FPS-live indicator leaves the KAOPEN coefficient nearly unchanged (−0.228 to −0.226), while the FPS term itself is near zero. That result does not eliminate substitution as an explanation, because the FPS measure is coarse and does not capture quality or coverage, but it does weaken the view that the openness gradient is simply a proxy for payment-system maturity. Read jointly, the KAOPEN and inclusion patterns remain consistent with rational substitution towards adjacent instruments as well as with institutional constraint.

6.3 Legal Feasibility as a Suggestive Constraint

Legal-regulatory constraint language is consistently negative in sign across the dynamic benchmark and the FPS-control variant. In the dynamic benchmark it is significant; in the reduced-spec Mundlak split it localises primarily to the between-country component, while the within-country deviation is essentially nil. The legal-constraint pattern is stable in sign and magnitude across specifications. This pattern remains theoretically coherent with legal scholarship documenting that many central bank statutes do not clearly authorise general-purpose digital currency issuance (Bossu et al. 2020), with political-economy arguments that statutory change is slow where veto players are numerous (Tsebelis 2002), and with BIS survey evidence reporting legal authority as a primary obstacle cited by central banks at the research stage (Kosse and Mattei 2023, BIS 2025). Bechara et al. (2025) and Soderberg et al. (2023) document the specific legal gaps that must be addressed before issuance can proceed.

Reverse causality is plausible. Jurisdictions that discover binding legal obstacles will talk about them, and stalled jurisdictions may discuss legal issues precisely because progress has become difficult. The legal-constraint coefficient is therefore best read as a

revealed feasibility marker, not an identified causal effect. A related possibility is that legal-constraint rhetoric proxies for underlying legal-infrastructure quality. The specification includes government effectiveness as a partial control, but a more targeted measure of statutory readiness might absorb part of this coefficient.

7. Conclusion

Central banks talk about CBDC far more than they advance beyond research, and the determinants of the two margins differ. Public commitment tracks global discourse waves and attenuates to near zero once persistence and common shocks are absorbed. By contrast, capital account openness, central bank independence, and legal-regulatory language are negatively associated with advancement beyond research but not with public commitment. The talk-action gap is therefore not simply noise in the speech data. It reflects an environment in which deliberation serves legitimacy and optionality at low cost, while progression beyond research exposes central banks to financial, legal, and political risks that vary sharply across jurisdictions.

The paper contributes in three ways. First, it introduces a speech-based panel for 122 jurisdictions that captures not only whether CBDC is discussed but how central banks frame the constraints they face, and it provides human-reference benchmark evidence for the commitment and legal-language dimensions. Second, it shows that the correlates of talk and action diverge in the directions predicted by a simple institutional framework, with the deliberation-side equations providing direct cross-equation evidence. Third, it connects CBDC communication to broader debates on institutional isomorphism and policy-practice decoupling by documenting a setting in which rhetorical convergence coexists with persistent cross-country variation in advancement.

For policy, the implication is that tracking what central banks say about CBDC is in-

formative about the agenda and the constraints they perceive, but not about whether or when they will move beyond research. Readiness to advance depends more directly on legal authority, open-economy vulnerability, and institutional risk tolerance. The future of public digital money will be shaped less by the intensity of public deliberation than by whether the structural conditions for progression are in place.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this work, the author used GPT-5.1 (OpenAI, via API) for automated classification of central bank speeches as part of the research methodology described in Section 4. The model scored speech excerpts for CBDC commitment, technology salience, and constraint dimensions under a structured rubric. The author reviewed and checked the outputs as described in the validation section and takes full responsibility for the content of the published article. No generative AI was used for manuscript drafting or editing.

References

Adrian, T. and Mancini-Griffoli, T. (2021). The rise of digital money. *Annual Review of Financial Economics*, 13:103–122.

Andolfatto, D. (2021). Assessing the impact of central bank digital currency on private banks. *Economic Journal*, 131(634):525–540.

Aruoba, S. B. and Drechsel, T. (2024). Identifying monetary policy shocks: A natural language approach. NBER Working Paper No. 32417.

Auer, R., Cornelli, G., and Frost, J. (2020). Rise of the central bank digital currencies: Drivers, approaches and technologies. BIS Working Papers No. 880.

Battaglia, L., Challet, D., Frazier, D., and Oomen, R. (2025). Inference with ai-generated regressors. CeMMAP Working Paper.

Bauer, M. D. and Swanson, E. T. (2023). A reassessment of monetary policy surprises and high-frequency identification. *NBER Macroeconomics Annual*, 37(1):87–155.

Baumgartner, M. and Zahner, J. (2025). Mapping central bank communication with contextual embeddings. BIS Working Papers No. 1299.

Bechara, M., Rosha, A., and Tourpe, H. (2025). Private law dimensions of token-based cbdc. IMF FinTech Note 2025/003.

BIS (2024a). Cbdcs and fast payment systems: Policy choices and trade-offs. BIS report based on central bank interviews.

BIS (2024b). Central bank speeches. https://www.bis.org/cbspeeches/download.htm.

BIS (2025). Advancing retail cbdc work while preserving optionality: Results of the 2024 bis survey. BIS Papers No. 159.

Blinder, A. S., Ehrmann, M., de Haan, J., and Jansen, D.-J. (2024). Central bank communication with the general public. *Journal of Economic Literature*, 62(2):412–476.

Bossu, W., Itatani, M., Margulis, C., Rossi, A., Weenink, H., and Yoshinaga, A. (2020). Legal aspects of central bank digital currency: Central bank and monetary law considerations. IMF FinTech Note 2020/001.

Boxenbaum, E. and Jonsson, S. (2008). Isomorphism, diffusion and decoupling. In Greenwood, R., Oliver, C., Suddaby, R., and Sahlin, K., editors, *The SAGE Handbook of Organizational Institutionalism*, pages 78–98. SAGE.

Bromley, P. and Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. *Academy of Management Annals*, 6(1):483–530.

Burlon, L., Chiu, J., Davoodalhosseini, S. M. R., and Hendry, S. (2024). Optimal cbdc quantity limits and banking stability. Working paper.

Campiglio, E., Deyris, J., Romelli, D., and Scalisi, G. (2025). Warning words in a warming world: Central bank communication and climate change. Working paper.

Chen, L., Maimbo, S., Mattoo, A., and Natarajan, H. (2022). Cbdcs versus fast payment systems in emerging markets: A survey of central banks. Policy report.

Chinn, M. D. and Ito, H. (2008). A new measure of financial openness. *Journal of Comparative Policy Analysis*, 10(3):309–322.

Chiu, J., Davoodalhosseini, S. M. R., Jiang, J. H., and Zhu, Y. (2023). Bank market power and central bank digital currency. *Journal of Political Economy*, 131(5):1213–1267.

Corbet, S., Hou, Y., Hu, Y., and Oxley, L. (2024). Understanding cbdc adoption determinants: Institutions, technology, and development. *Journal of International Financial Markets, Institutions and Money*, 92:101997.

Di Iorio, A., Kosse, A., and Mattei, I. (2024). Embracing diversity, advancing together: Results of the 2023 bis survey on central bank digital currencies and crypto. BIS Papers No. 147.

DiMaggio, P. J. and Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2):147–160.

Fernández-Villaverde, J., Sanches, D., Schilling, L., and Uhlig, H. (2021). Central bank digital currency: Central banking for all? *Review of Economic Dynamics*, 41:225–242.

Ferrari Minesso, M., Mehl, A., and Stracca, L. (2022). Central bank digital currency in an open economy. *Journal of Monetary Economics*, 127:54–68.

Gambacorta, L., Doerr, S., Gambetti, L., and Nitschke, J. (2024). Cblms: Domain-specific language models for central banking. BIS Working Papers No. 1215.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Gurkaynak, R. S., Sack, B., and Swanson, E. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1), May.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99(S1):S114–S133.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the fomc: A computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

International Monetary Fund (2025). Navigating retail cbdc choices in advanced economies. IMF Policy Paper.

Keister, T. and Sanches, D. (2023). Should central banks issue digital currency? *The Review of Economic Studies*, 90(1):404–431.

Kosse, A. and Mattei, I. (2022). Gaining momentum—results of the 2021 bis survey on central bank digital currencies. BIS Papers No. 125, Bank for International Settlements.

Kosse, A. and Mattei, I. (2023). Making headway – results of the 2022 bis survey on central bank digital currencies and crypto. BIS Papers No. 136, Bank for International Settlements.

Lannquist, A. and Tan, B. (2023). *Central bank digital currency's role in promoting financial inclusion*. International Monetary Fund.

Lastra, R. M. (2024). Mission expansion and the boundaries of central banking. Policy paper.

Le, T., Chuc, A., and Taghizadeh-Hesary, F. (2023). Degrees of cbdc adoption: Stage-dependent determinants in cross-country evidence. *Economic Analysis and Policy*, 80:1062–1078.

Lucca, D. and Trebbi, F. (2012). Measuring central bank communication: An automated approach with application to fomc statements. *Journal of Monetary Economics*, 59(6):513–528.

Lukonga, I. (2023). Monetary policy implications of cbdcs in open economies. IMF Working Paper.

Maryaningsih, N., Hermansah, H., and Faliansyah, M. (2022). Central bank digital currency: What factors determine the stages of adoption? *Bulletin of Monetary Economics and Banking*, 25(1).

Meyer, J. W. and Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2):340–363.

Murphy, K. M. and Topel, R. H. (1985). Estimation and inference in two-step econometric models. *Journal of Business and Economic Statistics*, 3(4):370–379.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, 25(1):221–247.

Pfister, C. (2024). The political economy of retail central bank digital currency. *Revue d'economie politique*, 134(4):623–648.

Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *The Quarterly Journal of Economics*, 100(4):1169–1189.

Romelli, D. (2024). Trends in central bank independence: a de-jure perspective. *BAFFI CAREFIN Centre Research Paper*, 217.

Rouanet, L. and Salter, A. (2025). Mission creep at the federal reserve since 2008. Working paper.

Scott, W. R. (2014). *Institutions and Organizations: Ideas, Interests, and Identities*. SAGE, 4 edition.

Silva, T. C., Moriya, K., and Veyrune, R. M. (2025). From text to quantified insights: A large-scale llm analysis of central bank communication. IMF Working Paper 2025/109.

Soderberg, G., Kiff, J., and Tourpe, H. (2023). How should central banks explore cbdc? strategic and legal dimensions. IMF FinTech Note 2023/008.

Sun, T., Zhou, Y. S., and Solomon, F. (2024). Cbdc and privacy: Design trade-offs and policy choices. IMF FinTech Note 2024/004.

Swanson, E. T. (2021). Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, 118:32–53.

Tsebelis, G. (2002). *Veto Players: How Political Institutions Work*. Princeton University Press.

Tucker, P. (2018). *Unelected Power: The Quest for Legitimacy in Central Banking and the Regulatory State*. Princeton University Press.

Whited, T. M., Wu, Y., and Xiao, K. (2023). Will central bank digital currency disintermediate banks? Working paper.

Appendix

S1. Summary Statistics

Table A1 reports summary statistics for the main variables used in the empirical analysis. The panel is unbalanced over 2015 to 2025. Sample sizes vary because speech-based measures require at least one observed speech per country-year, central bank independence coverage ends in 2023, and KAOPEN availability ends in 2022. The disparity between the CBDC stage mean (0.32) and the speech-based commitment mean (0.77) reflects the central pattern in the data. Verbal engagement with digital currency topics is more common than formal project advancement.

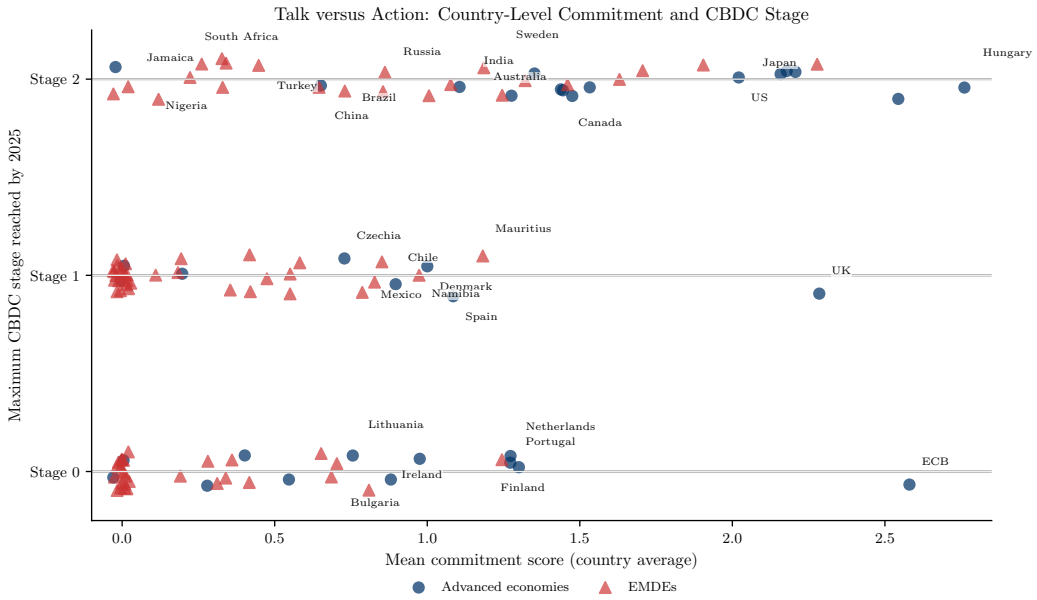
TABLE A1: Summary Statistics

Variable	N	Mean	SD	Min	Max
Panel A. Dependent Variables					
CBDC Stage (0-2)	2,427	0.32	0.61	0.00	2.00
CBDC Commitment (0-4)	947	0.77	1.25	0.00	4.00
Panel B. Speech-Based Measures					
Technology Salience (0-100)	947	45.31	18.20	0.00	100.00
Number of Speeches	947	16.71	27.35	1.00	230.00
Panel C. Institutional Variables					
Central Bank Independence	1,395	0.68	0.16	0.26	0.91
Government Effectiveness	1,901	-0.00	1.00	-2.44	2.32
Panel D. Main Controls					
GDP per Capita (current US$)	2,124	15,816	21,227	103	140,941
Capital Account Openness	1,406	0.32	1.41	-2.00	2.00
Mobile Subscriptions per 100	1,817	109.03	39.20	12.67	416.26
Innovation Index	1,311	34.73	12.89	10.20	68.40

Notes: Unbalanced panel, 2015–2025. CBDC Stage is coded 0 for no formal activity, 1 for research, and 2 for advanced activity (proof of concept, pilot, or launch). CBDC Commitment is the 0–4 speech-based ordinal scale, and Technology Salience is the 0–100 digital-modernisation index. Government Effectiveness follows the World Bank WGI (about −2.5 to +2.5), and Capital Account Openness is the Chinn–Ito KAOPEN index (−2 to +2). Speech-based variables are defined only for country-years with at least one scored speech. The scoring algorithm is described in Online Appendix Section OA3.

S2. Supplementary Descriptive Evidence

FIGURE A1: Talk versus Action in Country-Level Commitment and CBDC Stage



Notes: Each point is a country. The x-axis is mean commitment in speech-observed years; the y-axis is the highest CBDC stage reached by 2025, with slight vertical jitter for readability. Marker shape distinguishes advanced economies (AEs) from emerging market and developing economies (EMDEs).

Figure A1 provides the country-level descriptive view. Stage-2 countries have higher mean commitment on average (1.16 versus 0.35–0.40 for stages 0 and 1), but the relationship is far from deterministic. Eleven countries in the top commitment quartile remain at stage 0 or 1, including the ECB and the United Kingdom, while countries such as China and Nigeria reach stage 2 without entering the top commitment quartile. High public commitment is neither necessary nor sufficient for advancement beyond research.

S3. Robustness and Uncertainty Checks

The appendix collects the fuller strict-lag comparator promised in the main text, the aggregation and generated-regressor checks, and the identification-limit diagnostics. Secondary subgroup, timing, and validation details are reported in the online appendix below.

TABLE A2: Full Strict-Lag Speech Block in the Dynamic Benchmark

	Primary capped spec.	Full commitment	+ FPS control
Lagged retail stage 1_t	2.351***	2.367***	2.349***
	(<0.001)	(<0.001)	(<0.001)
Lagged retail stage 2_t	9.652***	9.746***	9.737***
	(<0.001)	(<0.001)	(<0.001)
Commitment$_{t-1}$ (capped at 2)	0.154	–	0.155
	(0.224)	–	(0.223)
Commitment$_{t-1}$ (full 0–4 scale)	–	0.147	–
	–	(0.127)	–
Technology salience$_{t-1}$	-0.006	-0.006	-0.006
	(0.320)	(0.287)	(0.316)
KAOPEN$_t$	-0.219**	-0.218**	-0.217**
	(0.020)	(0.020)	(0.023)
CBIE$_t$	-1.398***	-1.384***	-1.398***
	(0.004)	(0.005)	(0.004)
Financial stability language$_{t-1}$	0.012*	0.012**	0.012*
	(0.055)	(0.047)	(0.060)
Privacy language$_{t-1}$	0.015	0.015	0.016
	(0.290)	(0.307)	(0.287)
Legal-regulatory language$_{t-1}$	-0.017**	-0.017**	-0.017**
	(0.018)	(0.020)	(0.019)
Financial inclusion language$_{t-1}$	-0.002	-0.002	-0.003
	(0.584)	(0.624)	(0.576)
Substitutes language$_{t-1}$	-0.004	-0.005	-0.004
	(0.722)	(0.673)	(0.740)
FPS live$_t$	–	–	-0.049
	–	–	(0.756)
N	460		
Countries	82		

Notes: This table reports the full strict-lag speech block. All columns use the same strict retail-only sample as Table 5, but add salience and the four additional articulated-constraint channels without human-reference coding. The qualitative sign-significance hierarchy is preserved relative to the reduced theory-prioritised speech benchmark, although coefficient magnitudes shift, especially for commitment and CBIE. Entries are coefficients with country-clustered p-values in parentheses. GDP per capita, mobile subscriptions, innovation index, and government effectiveness are included but omitted from display. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A3: Commitment Aggregation Sensitivity in the Progression Hazard

Variant	Coefficient	p-value
Maximum (baseline)	0.050	0.396
Mean	0.030	0.852
Median	0.067	0.699
Share ≥ 2	0.079	0.880
Fixed-k resampling (Rubin)	0.027	0.866
Generated-regressor draws (Rubin)	0.019	0.777

Notes: Progression hazard (1→2) with country and year fixed effects. Direct estimates use country-clustered p-values; Rubin rows pool across repeated speech draws. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A4: Dynamic Benchmark with Generated-Regressor Uncertainty (Speech Terms)

Speech regressor	Rubin $\bar{\beta}$	Rubin SE	Rubin p-value
Commitment$_t$	0.052	0.079	0.510
Salience$_t$	0.000	0.006	0.958
Financial stability constraint$_t$	0.002	0.005	0.633
Privacy constraint$_t$	-0.001	0.016	0.944
Legal constraint$_t$	-0.012**	0.006	0.043
Financial inclusion constraint$_t$	0.004	0.006	0.455
Substitutes constraint$_t$	-0.000	0.012	0.982

Notes: Dynamic ordered benchmark re-estimated on 121 speech-measure draws (1 baseline + 120 within-country-year bootstrap draws), using the full speech-regressor covariate set on the broader speech-observed sample. Rubin estimates combine within-draw and between-draw variance. The common estimation sample is constant across draws (N = 574, 87 countries), which is broader than the strict retail-only benchmark sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A5: Event-Study Pre-Trend Joint Tests

Outcome	Pre-trend test statistic	Pre-trend p-value
Progression transition ($Y_{i,t+1}^{12}$)	6.489	0.090
Speech commitment (Commitment$_{it}^{\max}$)	8.252	0.041
Advanced-stage status ($S_{it} = 2$)	2.041	0.564

Notes: Joint tests of event-study lead coefficients with reference period $t = -1$. Values correspond to the diagnostics discussed in the identification-boundaries subsection.

TABLE A6: Time-to-Build Diagnostic

Group	Transition rate	Transitions	Observations
Early high commitment, 2019–2020	0.200	2	10
Early low commitment, 2019–2020	0.150	3	20
Early high commitment, 2021–2022	0.385	5	13
Early low commitment, 2021–2022	0.087	4	46
Interaction estimates for $Y_{i,t+1}^{12}$			
Early high commitment × late period (unadjusted)	0.248	$p = 0.238$	$N = 89$
Early high commitment × late period (+ controls)	0.221	$p = 0.316$	$N = 77$

Notes: Early high commitment denotes countries with mean commitment of at least one in 2015–2018. The upper panel reports raw 1→2 transition rates in the speech-observed stage-1 risk set. The lower panel reports the interaction coefficient from a simple LPM with, respectively, and without the standard structural controls. Both interaction models use country-clustered standard errors. The pattern is directionally consistent with time-to-build, but the interaction is imprecisely estimated.

S4. Measurement Validation

The human-reference benchmark tables referenced in the main text are reported here.

Human benchmark protocol. A structured single-coder protocol was applied to a 100-speech subset of the 200-speech validation sample. The sample is deliberately informative. Speeches with non-zero commitment signals and borderline classifications are prioritised, with random fill from the remaining cases. Coding uses full speech text, not the excerpted scoring input, the same 0–4 commitment rubric as the machine scoring pipeline, and archived coder notes and confidence fields in the replication files. This benchmark is single-coder and informatively sampled rather than a multi-coder random-label audit.

TABLE A7: Human-Coded Commitment Benchmark (100-Speech Subsample)

Comparison	*N*	Exact agreement	Adjacent agreement	Weighted κ (quad.)
Human vs GPT-5.1	100	0.83	0.97	0.787
Human vs Hybrid pipeline	100	0.84	0.94	0.706

Notes: Human labels come from a single-coder full-text benchmark on 100 validation speeches, coded on the same 0–4 commitment rubric using conservative criteria. Hybrid matches use Date-Country-Author-URL keys; if duplicate URL-key matches exist, the highest-confidence pipeline row is retained. Adjacent agreement is defined as absolute category difference ≤ 1. Category 4 is absent in the human-coded subsample, consistent with the rarity of implementation or launch speeches. This is a single-coder human-reference benchmark rather than a multi-coder adjudicated standard.

TABLE A8: Human-Coded Legal-Constraint Benchmark (60-Speech CBDC-Relevant Subsample)

Comparison	*N*	Exact agreement	Adjacent agreement	Weighted κ (quad.)	Weighted κ (linear)
Human vs pipeline legal level	60	0.27	0.87	0.668	0.414

Notes: Single-coder benchmark on a 60-speech subsample from the 2015–2025 scored corpus, restricted to speeches with positive CBDC commitment. Human legal levels use a 0–4 full-text rubric; pipeline legal levels map the 0–100 legal score into five bins (0, 1–20, 21–40, 41–60, 61–100). Adjacent agreement is defined as absolute difference ≤ 1. This is a single-coder benchmark rather than a multi-coder adjudicated standard.

Online Appendix

OA1. Extended Robustness Tables

The online appendix reports subgroup splits and additional specification variants.

TABLE OA1: Formal AE/EMDE Interaction Tests in the Dynamic Ordered Model

Term / Implied Effect	Coefficient	p-value
$\text{KAOPEN}_t \times$ AE	0.225	0.217
$\text{CBIE}_t \times$ AE	-1.371	0.205
Implied KAOPEN effect in EMDEs	-0.276	0.001
Implied KAOPEN effect in AEs	-0.051	0.792
Implied CBIE effect in EMDEs	-0.764	0.377
Implied CBIE effect in AEs	-2.135	0.005

Notes: Pooled dynamic ordered probit with lagged stage, year fixed effects, and the full control set on the common sample (574 country-years, 87 countries, 2015–2022). AE is the advanced-economy indicator. Entries use country-clustered inference. Interaction rows are direct coefficient tests; implied-effect rows are group-specific linear combinations.

TABLE OA2: Progression Model under Selection and Specification Robustness

	Commitment_t	KAOPEN_t	CBIE_t	Legal_t
Baseline (country + year FE)	0.050	-0.043	-1.303	0.000
	(0.396)	(0.858)	(0.541)	(0.967)
IPW-weighted	0.057	-0.008	-1.405	0.000
	(0.358)	(0.973)	(0.452)	(0.950)
Logit (year FE only)	0.234	-1.579**	-1.682	-0.040
	(0.450)	(0.031)	(0.644)	(0.173)

Notes: Progression model (1→2 risk set). For linear models, p-values in parentheses use country-clustered standard errors. IPW weights come from a probit for speech availability. The logit row reports model-based inference. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE OA3: Construct and Timing Alignment Checks for Commitment

Specification	N	Countries	Commitment coef.	p-value	$\Delta \Pr(S_{t+1} = 2)$
Strict-lag dynamic (t-1 speech, stage dummies)	460	82	0.071	0.322	0.005
Dynamic + Year FE (lagged stage dummies)	574	87	0.057	0.365	0.004
Dynamic + Year FE (lagged stage linear)	574	87	0.059	0.352	0.004
Dynamic + Year FE (commitment capped at 2)	574	87	0.124	0.147	0.005

Notes: Each row re-estimates the dynamic bridge with the listed design change. The first row is the strict-lag timing check with speech regressors lagged to $t-1$ and lagged-stage dummies; the remaining rows are contemporaneous or construct variants around the dynamic benchmark. $\Delta \Pr(S_{t+1} = 2)$ is the average discrete change from a one-point increase in commitment; in the capped row, the shift is re-capped at 2. p-values are model-based ordered-probit values for these checks. Clustered and bootstrap inference is reported in the main tables.

OA2. Stage Construction and Timing

The annual stage outcome used in the paper is `cbdc_stage3` from the country-year backbone panel. We map this directly to $S_{it} \in \{0, 1, 2\}$ and construct transitions only on contiguous annual pairs, so $S_{i,t+1}$ is observed only when both years are present for country i.

The annual coding rule is year-end status. If a country crosses multiple tracker categories within a calendar year, the annual panel retains a single end-of-year category rather than dated within-year milestones. A transition from stage 1 in year t to stage 2 in year $t+1$ is therefore interpreted as "advanced by the end of $t+1$," not as a dated within-year event.

The tracker-based stage variable does not provide harmonised within-year milestone dates across countries. This prevents a consistent cross-country implementation of alternatives such as first-observed-within-year status. As a result, annual models cannot fully separate speeches that occur before a milestone from speeches that occur after a milestone in the same calendar year. To limit this problem, the main specifications use speech measures from $t-1$ to predict retail stage outcomes in $t+1$, with lagged retail-stage dummies absorbing trajectory persistence. The contemporaneous commitment checks are reported in Online Appendix Table OA3.

Stage coding is built from publicly documented milestones and therefore shares part of its information environment with central bank communications. This source overlap can inflate contemporaneous talk-action correlations. The transition and lagged specifications reduce this overlap, and interpretation is kept at the level of structured associations rather than causal effects.

OA3. Speech Scoring Algorithm

The speech-scoring procedure combines a keyword module that yields transparent baseline scores, a large language model (LLM) classifier that scores speech excerpts under a structured rubric, and a hybrid rule that combines the two and corrects systematic failures of pure keyword matching.

OA3.1 Corpus and Preprocessing

The combined speech corpus contains 37,814 speeches from 134 jurisdictions, drawn from the BIS central bank speeches database (BIS 2024b) and supplemented by Campiglio et al. (2025). For the analysis window we retain speeches delivered on or after 1 January 2015 that have valid text exceeding 100 characters and complete country metadata. The resulting analysis sample comprises 15,906 speeches from 122 jurisdictions (2015–2025).

OA3.2 Keyword Module

The keyword module produces two speech-level scores. These are a continuous technology-salience index and an ordinal CBDC-commitment score.

Technology-salience lexicon. The salience lexicon spans eight categories. These are *CBDC* ("cbdc," "central bank digital currency," "digital fiat"); *Payments* ("instant payment," "real-time payment," "ISO 20022"); *Fintech* ("fintech," "regtech," "innovation hub"); *Digital assets* ("stablecoin," "crypto-asset," "tokenisation"); *Technology* ("blockchain," "dlt," "ai"); *Infrastructure* ("api," "open banking"); *Transformation* ("digital transformation," "cashless"); and *Security* ("cybersecurity," "operational resilience"). The displayed list is illustrative. The full replication lexicon includes both British and American spelling variants for terms such as tokenisation and modernisation.

For speech s, let h_{sj} denote the number of keyword hits in category $j \in \{1, \ldots, 8\}$ and let $|s|_{\text{KB}}$ denote speech length in kilobytes. The keyword salience score is

$$\text{Sal}_s^{KW} = 60 \cdot \frac{\#\{j : h_{sj} > 0\}}{8} + 40 \cdot \min\Big(1,\ \frac{\sum_j h_{sj}}{0.5 \cdot |s|_{\text{KB}}}\Big),$$

where $\#\{j : h_{sj} > 0\}$ counts the number of keyword categories with at least one hit. The first term therefore captures breadth across categories (range 0–60), and the second captures mention density (range 0–40).

CBDC-commitment lexicon. Commitment keywords are organised in three signal groups. These are research terms ("exploring," "feasibility study," "consultation"), development

terms ("pilot," "proof of concept," "testing"), and implementation terms ("launching," "issuance," "rollout"). The keyword commitment score is

$$\text{Com}_s^{KW} = \begin{cases} 4 & \text{if implementation keywords appear,} \\ 3 & \text{if development keywords appear but implementation keywords do not,} \\ 2 & \text{if research keywords appear but no higher-tier keywords do,} \\ 1 & \text{if generic CBDC keywords appear but no higher-tier keywords do,} \\ 0 & \text{otherwise.} \end{cases}$$

That is, the keyword score equals the highest commitment tier activated in speech s.

OA3.3 LLM Classifier

Excerpt construction. Rather than submitting full speech text, we construct a 4,500-character excerpt using three windows. These are the first 1,500 characters, one pseudo-random 1,500-character window drawn from the middle half of the speech, and the final 1,500 characters. Central bank speeches follow conventional structures in which policy announcements and risk assessments concentrate in introductions and conclusions. The tripartite sampling rule captures these segments while including a middle-text sample for speeches that develop CBDC topics at length. Computational replication relies on archived scored outputs retained in the replication package rather than redrawing these middle windows. For the 12.6 per cent of speeches shorter than 5,000 characters, the full text is submitted.

Classification and outputs. Each excerpt is submitted to GPT-5.1 via the OpenAI API in zero-shot classification mode using the rubric prompt reproduced in Section OA5. The model returns a structured JSON object containing technology salience (0–100), CBDC commitment (0–4), five constraint scores (0–100 each), a dominant constraint label, and a confidence score. When the API response is unparseable, keyword-only scoring serves as a fallback.

OA3.4 Hybrid Combination Rule

Keyword and LLM outputs are combined at the speech level and then aggregated to the country-year level.

Speech-level salience. When both components are available, speech-level salience is

$$\text{Sal}_s = \min\left(100,\ 0.5 \cdot \text{Sal}_s^{KW} + 0.5 \cdot \text{Sal}_s^{LLM} + \Delta_s\right),$$

where Δ_s applies discrete adjustments for explicit evidence of implementation planning in the speech text. These include a +35 adjustment for active programme indicators and a +5 adjustment for explicit implementation language. The resulting no-floor salience measure is therefore partly implementation-loaded and is interpreted in the main text as a broad digital-modernisation indicator rather than a pure technology-orientation proxy. A floor-imposed salience variant was also estimated as a sensitivity check; action-side results were unaffected, while deliberation-side coefficients were mechanically larger.

Speech-level commitment. Commitment uses an asymmetric rule designed to correct two failure modes. It addresses false negatives from branded project names that omit generic CBDC terminology, and false positives from ambiguous keyword triggers that are not CBDC-related in context.

$$\mathrm{Com}_s = \begin{cases} \mathrm{Com}_s^{LLM} & \text{if } \mathrm{Com}_s^{KW} = 0 \text{ and } \mathrm{Com}_s^{LLM} > 0, \\ 0 & \text{if } \mathrm{Com}_s^{KW} > 0 \text{ and } \mathrm{Com}_s^{LLM} = 0, \\ \left\lfloor 0.5 \cdot \mathrm{Com}_s^{KW} + 0.5 \cdot \mathrm{Com}_s^{LLM} \right\rfloor & \text{otherwise.} \end{cases} \tag{OA1}$$

Among the 15,906 speeches in the analysis sample, 273 (1.7 per cent) satisfy $\mathrm{Com}_s^{KW} = 0$ and $\mathrm{Com}_s^{LLM} > 0$. These cases trigger the branded-name override. This shows that the correction is empirically relevant but does not dominate the scoring. Because the override retains the LLM value, these speeches do not create a hybrid-LLM discrepancy.

Country-year aggregation. For country i in year t, commitment is aggregated as the within-year maximum,

$$\mathrm{Commitment}_{it}^{\max} = \max_{s \in (i,t)} \mathrm{Com}_s,$$

while technology salience and constraint scores are aggregated as within-year means across speeches. Alternative aggregation operators (median, share above threshold, fixed-k random resampling) are reported in the appendix tables.

Computational details. The scored 2015–2025 speech sample of 15,906 speeches was processed via the OpenAI API with 50 concurrent requests in approximately 71 minutes (4,277 seconds). In the archived processed scoring files retained for replication, keyword-only fallback is invoked for 0 of 15,906 speeches. Replication code is available in the replication package.

OA4. Additional Identification Diagnostics

Mechanism interactions referenced in the discussion of identification limits are reported here.

TABLE OA4: Mechanism Interaction Terms in Dynamic Ordered Specification

Interaction term	Coefficient	*p*-value
CBIE × Financial depth	-3.198	0.206
CBIE × Legal constraint	-0.001	0.981
KAOPEN × Financial depth	0.070	0.846

Notes: Dynamic ordered interaction specification (2015–2022, N = 407, 64 countries). *p*-values are model-based for this diagnostic. Country-clustered and bootstrap inference is reported in the main tables.

OA5. LLM Prompt Specification

The verbatim prompt submitted for each speech excerpt is reproduced below. The prompt uses zero-shot classification with structured JSON output.

```
You are analyzing a CENTRAL BANK speech.

CRITICAL INSTRUCTIONS:
1. TEXT DEPENDENCE: Evaluate STRICTLY on the provided text. DO NOT
use external knowledge about the central bank's actual projects
or status.
2. GEOGRAPHIC NEUTRALITY: Focus on the substance of actions and
intent. Recognize that CBDC projects may have specific names (e.g.,
Digital Euro, e-CNY). A named project should be scored the same
as a generic project at the same stage.
3. PUBLIC VS PRIVATE: Discussions about regulating private stable-
coins/crypto increase Salience, but do NOT count towards CBDC Com-
mitment unless a public CBDC is the focus.

SPEECH EXCERPT:

"{text}"

—

DIMENSION 1: TECHNOLOGY SALIENCE (salience_toi, 0-100)

CONTEXT: Score strategic innovation (e.g., DLT, AI, Fintech, Dig-
ital Assets, Payment Modernization).

0-20: No Focus.
20-40: Simple related observation.
40-60: Observation & Analysis (Analyzing fintech/stablecoin trends,
```

```
monitoring, no CB action).
60-75: Active Exploration (Researching new tech, modernization projects).
75-90: Concrete Initiatives/Pilots (PoC, Pilots, detailed archi-
tecture).
90-100: Implementation & Leadership (Major infrastructure rollout).

—

DIMENSION 2: CBDC COMMITMENT (commitment_cbdc, 0-4)

CONTEXT: Score the highest level of commitment regarding a public
Central Bank Digital Currency (CBDC) or equivalent digital fiat.

0: None (No mention of CBDC/digital fiat, or explicitly dismissed.
Discussion of private crypto only is 0).
1: Awareness (Peripheral mention, discussing the concept generally).
2: Research/Exploration ("Studying," "assessing feasibility," "an-
alyzing the case for," consultation).
3: Development/Pilot ("Proof of concept," "pilot program," "design-
ing architecture," "testing").
4: Implementation/Launch ("Launching," "issuance," "rollout," "live").

—

DIMENSION 3: CONSTRAINTS (0-100 each)

Score the emphasis on risks or requirements: financial_stability,
privacy, legal, inclusion, substitutes (e.g., stablecoins, instant
payments, sovereignty, currency substitution).

"dominant_constraint": highest scoring constraint or "none" if all
<20.

—

CONFIDENCE: Your overall certainty (0.0-1.0).

Return ONLY JSON:

{{"salience_toi": <int>, "commitment_cbdc": <int>, "constraints":
{{"financial_stability": <int>, "privacy": <int>, "legal": <int>,
"inclusion": <int>, "substitutes": <int>}}, "dominant_constraint":
"<string>", "confidence": <float>}}
```

The prompt embeds text-dependence, geographic-neutrality, and public-versus-private instructions to ensure scoring reflects the excerpt rather than the model's parametric knowledge, treats branded project names equivalently, and separates stablecoin regulation from public CBDC commitment.