Central Bank Mandates, Evolution and Determinants: Lessons from Large Language Models

N. Nergiz Dincer* Barry Eichengreen† Joan J. Martinez‡

Abstract

This paper examines central bank mandates, contrasting single or primary objectives (typically price stability) with multiple objectives. We compare mandate measures constructed using researcher judgment with two automated methods: string-based keyword matching and a large language model that apprehends meaning and context. While researcher and LLM methods achieve reasonable agreement, string-based coding systematically misclassifies complex mandate structures, particularly those with conditional clauses or hierarchical objectives common after 1990. Despite these measurement challenges, we identify robust determinants of mandate choice: price stability-dominant mandates are more likely in wealthier countries with independent central banks; multiple balanced objectives are more likely in trade-open economies. These patterns hold across reliable coding methods. Other results vary by method, underscoring the importance of validation when constructing institutional indices from legal texts.

JEL Classification: E52, E58, E24, J64, C81
Keywords: Central bank mandates, monetary policy, employment, inflation targeting, large language models

*TED University, Department of Economics and TEDU Trade Research Center. Email: nergiz.dincer@tedu.edu.tr

†Department of Economics, University of California, Berkeley. Corresponding author: eichengr@econ.berkeley.edu.

‡Haas School of Business, University of California, Berkeley. Email: martinezp_jj@berkeley.edu

We acknowledge generous financial support from the Clausen Center for International Business and Policy at UC Berkeley. All remaining errors are our own. This version: September 22, 2025.

1 Introduction

Central bank governance is in the news. The most prominent discussion and accompanying literature is probably that on central bank independence: whether independence from politics enables central banks to better pursue sound and stable policies and deliver desired economic outcomes. But also prominent, and closely related, is the question of how societies define the central bank's mandate, and how that mandate affects the central bank's monetary policy stance and strategy in practice. One school of thought emphasizes the advantages of a narrowly focused, often single, objective, namely the pursuit of price stability. Those who favor a single mandate argue that central banks with limited objectives can only achieve limited targets, and that clear communication is facilitated by a focus on a single mandated objective. The credibility of central bank policy may be damaged when central bankers endeavor to pursue targets that their instruments are incapable of affecting. An opposing school of thought insists that central banks are regularly tasked with pursuing multiple objectives, not just price stability but inter alia high employment, financial stability and even climate-change abatement. Proponents argue that central banks actually possess multiple instruments enabling them to pursue multiple targets. The defenders of multiple mandates suggest that consistent policy formulation, clear communication and political accountability require acknowledging this reality.

A limited literature has analyzed central bank mandates, their adoption, and their effects. Grilli et al. (1991) construct a binary measure of whether monetary stability is mandated, while more recent work such as Romelli (2024) refines the scale to five points by incorporating additional goals. Garriga (2016) provides a "human coding" of as many as 192 central bank mandates starting in 1970, based on a close reading of central bank statutes and related legislation, identifying instances where inflation and employment are mentioned in the mandate. She distinguishes cases where the central bank has a price-stability goal but not an employment goal (an "inflation mandate"), where it has both price-stability and employment goals (a "dual mandate"), and where it is also tasked with pursuing other goals.

Garriga and Rodriguez (2025) use this dataset to analyze the impact of mandate type on observed inflation and employment outcomes, finding that central banks with dual (inflation and employment) mandates tolerate higher inflation but have no differential impact on employment. Li (2024) constructs a dataset for 50 central banks in 2022 by reading through central bank statutes and legislation and placing the subject banks into mandate categories: price stability focused, exchange rate stability focused, output focused, and employment/unemployment focused. He finds that the choice of mandate depends on economic conditions, with fast-growing countries less likely to choose a price stability mandate, while richer economies are more likely to opt for an employment mandate. In contrast, he finds no apparent difference between countries with civil and common law traditions. Gabriel et al. (2022) categorize 25 advanced-country central banks according to whether they have a mandate to maintain financial stability, based on World Bank surveys of bank regulation. They find no impact of the structure of the mandate on inflation. Dikau and Volz (2021) focus on whether central banks have mandates to promote sustainability and green finance. They analyze the statutes of 135 central banks in 2017, finding that only 12 percent have explicit sustainability mandates, although an additional 40 percent are mandated to support the government's general economic objectives, some of which may be climate related.

In our own work (Dincer and Eichengreen (2008b, 2010b, 2018)), we included as criteria for evaluating central bank transparency the presence or absence of a formal statement of the objectives of monetary policy. For as many as 120 central banks starting in 1998, we used a close reading of central bank statutes to distinguish cases where no formal objective or objectives were stated, cases where there were multiple objectives without prioritization (as, for example, the United States since passage of the Humphrey-Hawkins Act), and cases where there was one primary objective or else multiple objectives with explicit priority attached to one.

An alternative to researcher-based coding, used in all the aforementioned studies, is to interrogate central bank officials regarding their mandates. Thus, the Bank for International

Settlements' Central Bank Governance Network surveyed 62 central banks in 2022, finding that all of the surveyed central banks reported explicit mandates to maintain price stability, that more than 90 percent reported explicit mandates to maintain financial stability, and that more than 40 percent reported explicit mandates to support the general economic policies of the government (Tombini, 2025). Slightly more than 20 percent of surveyed banks reported explicit mandates to support economic growth, slightly less than 20 percent explicit mandates to support employment. A limitation of this source, besides that the fact that it has responses from only a subset of central banks, is that while it tabulates different mandates it does not provide information on combinations and hierarchies: whether central banks have single or dual mandates, or whether some mandated targets receive priority over others.

Finally, there are single-country case studies. Jackson and Baily (2023) consider the Bank of England's new mandate, adopted in 2021, to facilitate the transition to net zero. Using a "discourse analysis" of Bank of England publications and speeches (a discussion of the language in these documents), along with participant interviews, they highlight the potential conflict between the central bank's traditional objective of maintaining price and financial stability and this new mandate, concluding that the Bank has in practice limited its policies to mitigating climate-related risks to financial stability. Bertsch et al. (2025) analyze changes in the Federal Reserve System's perceived mandate over time. In a departure from human-coding-based studies, they apply large language models (LLMs) to the corpus of central bank speeches since 1970, and characterize the results as showing that Federal Reserve officials regard financial stability as an important objective along with the explicit elements of their dual mandate (price stability and high employment), especially in periods when debt-to-GDP ratios are high.

These studies, while shedding light on the pattern of central bank mandates, their incidence and their effects, are subject to limitations. Most cover only limited countries or periods of time. They categorize mandates in different ways. Some consider only the pres-

ence or absence of price stability and employment mandates, whereas others consider also other mandated goals such as financial stability. Some consider whether certain objectives receive priority above others, where other such studies do not. The reliability of human-agent (researcher) coding may be affected by cognitive biases and researcher fatigue. Surveys of central bank officials may be affected by those officials' economic or political preconceptions. The reliability of large language models for categorizing central banks according to mandate is yet to be systematically analyzed.

With these limitations in mind, we assemble a new data base of central bank mandates over more than a century and construct measures of whether central banks, as a matter of statute and legislation, possess an explicit primary objective or instead are tasked with pursuing multiple goals. We analyze the incidence of these mandates in terms of the economic and financial characteristics of countries. Specifically, we analyze the objectives of 120 central banks annually over the period since 1900, documenting trends in the composition of central bank mandates over much longer time span than in preexisting studies, and associate these trends and cross-country variations with differences in country conditions.

An important innovation of this paper is that we systematically compare traditional human-agent coding, where the researcher goes through central bank statutes line by line, making judgments about the mandates described there, with automated methods made possible by recent advances in machine learning. Human-agent coding makes full use of researcher experience and intuition but may be subject to fatigue and interpretative drift that introduce systematic measurement error, especially given that we are coding 8,277 country-year observations. We compare this human coding to two automated methods.

One is simple string-based coding, using a rule-based algorithm that searches for specific hierarchical terms and individual words or strings of words assembled in comprehensive keyword dictionaries. This approach has the advantage of transparency and reproducibility, making it easy to audit and validate. However, it may miss contextual nuances and misinterpret negations and conditional language that change meaning without changing keywords.

The second method makes use of large language models (LLM) that respond to contextually enriched prompts based on the semantic content and institutional context of mandate texts, rather than relying on simple keyword matching. The LLM approach uses a zero-shot classification algorithm that processes mandates using carefully designed prompts that specify the classification tasks and theoretical framework. As in the case of string-based coding, this approach should maintain consistency across our large data set. But, in addition, it is likely that subtle institutional distinctions will be captured by this method more often than that using keyword matching. The corresponding limitation of the LLM approach is that it can be something of a black box, making it difficult to trace specific classification decisions, and reproducibility may be compromised by model updates, prompt sensitivity, and potential biases in training data.

Importantly, we completed the researcher-based coding before implementing the two automated methods, and did not revise the researcher-based coding subsequently. Nor did we use the researcher-based coding as an input for the automated models, or use the output of one automated model for pre-training the other. Thus, we can compare the results of these three mutually independent approaches to coding central bank mandates.

Comparisons of the three methods used to categorize mandates shows broad agreement between approaches on overall trends over time, which are toward price stability-dominant mandates. Looking across the entire century, we find that the share of central banks with a single objective (or a primary objective that dominates subsidiary objectives) rose first in the 1920s, with the adoption of exchange rate and price stability-related targets by newly established central banks whose statutes mandated adherence to gold standard rules (Eichengreen and Kakridis, 2024), and then again in the 1990s, with the spread of inflation targeting. After the Global Financial Crisis of 2008-9, there was then some stabilization or even a slight decline in the share of central banks with single mandates or mandates prioritizing inflation, reflecting the tendency for some central banks to be assigned additional responsibility for the maintenance of financial stability.

Broad agreement is not the same as perfect agreement. The string-based and human researcher-based methods are relatively in close agreement prior to 1990, the point in time, interestingly, when a significant number of additional central banks started receiving mandates to target inflation; thereafter our human-based scores for price-stability dominance are considerably higher than those generated by the string-based approach. LLM-based coding generated consistently higher scores for price-stability dominance prior to 1990, but after 1990 our human agent-based scores are higher than LLM-based scores. Conformance scores across methods are consistently greater for high-income countries, perhaps reflecting the adoption of relatively standardized legal frameworks that reduce interpretive differences.

In other respects, the results of applying the three coding methods differ markedly. For example, over the entire period starting in 1900, the string-based method identifies just 2.1 percent of country/year observations as dual mandates, whereas the LLM approach associates 10 percent of observations with dual mandates, the researcher-based approach fully 28.6 percent.

These pronounced differences in coding methods would seem to complicate any effort to identify systematic correlates of mandate choice. Yet we consistently find that economic development (higher GDP per capita) and greater central bank independence are associated with price stability-dominant mandates, while trade openness is associated with an increased likelihood of observing dual mandates.

Our paper contributes to several literatures. Most obviously, there is the literature directly concerned with the incidence and effects of central bank mandates, summarized earlier in this introduction. In addition, there is the recent literature applying large language models to central bank documents. Besides Bertsch et al. (2025), who analyze speeches by Federal Reserve officials, as already mentioned, Bohl et al. (2023) use text mining methods and structural topic models to analyze the influence of central bank mandates on the tone and content of speeches of central bank board members, contrasting speeches of officials of the Federal Reserve (with its dual mandate) and the European Central Bank (with its single

mandate targeting inflation). They find that ECB board members focus more on inflation than employment issues, while members of the Fed's Board of Governors focus more on employment and labor market conditions, and that expectations about the variables at the core of each central bank's mandate drive the tone of speeches delivered by these officials. Alonso-Robisco and Carbó (2023) apply large language models and dictionary methods to central bank documents as a way of gauging sentiment toward the creation of a central bank digital currency. Christiano Silva et al. (2025) apply a large language model to a dataset of central bank documents from 169 countries to develop a communication index that captures signals about future policy rate changes, forward guidance, and balance sheet operations. Here we use such models to analyze central bank statutes and specifically mandates, as opposed to speeches and related policy documents. And we provide systematic comparisons of the mandate categorizations provided by different automated methods as well as comparisons of automated and human agent-based coding.

Third, there is a relatively sparse literature on how central bank mandates affect economic outcomes. Some empirical studies suggest that central banks with dual mandates respond more aggressively to recessions, leading to faster recoveries; see for example Rodriguez-Rodriguez et al. (2024). Note, however, that this study, like others contrasting single and dual mandates, focuses narrowly on the comparison between the ECB and the Fed, where we consider a very much larger set of central banks and their mandates. Garriga and Rodriguez (2025), in the study just cited, find some evidence that dual mandates are associated with small short-term declines in unemployment, but no evidence of long-term employment benefits. Here we focus on the determinants of mandate structure rather than the effects. And our study of course differs in both the type of central bank mandates considered and the methods used to place central banks into mandate categories.

The rest of the paper is organized as follows. Section 2 introduces our classification system. Section 3 describes the data used in the analysis and the three classification methods (string-based, LLM-based, and expert coding). Section 4 examines the reliability of these

classification approaches. Section 5 analyzes the determinants of mandate choice. Section 6 concludes. The appendix that follows contains some additional technical details.

2 Conceptual Framework and Classification System

When categorizing central bank mandates, we utilize the comprehensive classification system shown in Table 1. This eleven-point scale captures the full spectrum of central bank mandates, from those focused exclusively on price stability to those with multiple objectives or without price stability as a concern.

Mandate scores range from 0.0/0.1, when there is no clear mandate or objective and when price stability and related terms are not mentioned at all, to 0.9/1.0, when price stability is the dominant or sole objective of policy. Intermediate scores are assigned when (in ascending order) the central bank is mandated to pursue multiple balanced objectives and when price stability is explicitly categorized as a secondary objective; when the central bank receives a dual mandate; and when price stability is co-equal with still other objectives.

This eleven-point mandate classification builds on the foundation of earlier indices, listed in Table 2, but introduces greater granularity to capture the actual diversity of statutory formulations. Classic coding, most notably Cukierman et al. (1992), provided a six-tier hierarchy ranging from "price stability sole objective with final authority" to "objectives excluding price stability." This framework was important for systematically distinguishing whether price stability was present, primary, or subordinated. However, it still grouped together cases that differ in meaningful ways, such as when price stability is primary alongside financial stability, versus when it is primary but explicitly paired with employment or growth. Similarly, Grilli et al. (1991) employed only a binary measure of whether monetary stability was mandated, whereas more recent work such as Romelli (2024) refined the scale to five points but still treated categories like "non-conflicting" and "conflicting" additional goals in broad strokes.

Our eleven-point scheme is designed to reflect more subtle gradations in central bank mandates. At the upper end of the hierarchy, it distinguishes between "sole" price stability mandates, "dominant" mandates with explicit conflict-resolution authority, and "primary plus secondary" formulations that differ depending on whether the second objective is compatible (financial stability) or non-compatible (employment, growth). At the lower end, it further differentiates between balanced multi-objective mandates, compatible-objective primacy, growth/employment primacy, and cases with either no clear mandate or an explicit non-price focus. This extended spectrum allows us to situate contemporary debates—where financial stability, employment, and even climate goals are increasingly invoked in central bank statutes—within a more precise taxonomy. In short, the eleven-point classification does not replace but rather extends earlier indices: it preserves comparability with Cukierman et al. (1992) and related measures while offering a richer rubric that can capture the evolving complexity of central bank objectives.

3 Data and Methodology

3.1 Data Sources

We gathered data on central bank laws from central bank annual reports, central bank websites and other central bank publications for as many central banks as possible over the period 1900–2024. The resulting sample comprises 120 central banks and a panel of 8,277 country-year observations. Most of the omissions are central banks of micro-states; our sample includes the central banks of all large, systemically important countries. Sources are as described in Dincer and Eichengreen (2008a, 2010a, 2018).

Samples. The paper employs a sub sample for estimation purposes. The panel data analysis of mandate determinants (Section 5) relies on a restrictive sample comprising 2,558 country-year observations drawn from 91 countries between 1970 and 2021. This estima-

tion sample requires countries to have a minimum of 30 years of data to ensure adequate within-country variation for fixed effects estimation, alongside complete information for central bank independence measures, macroeconomic indicators (including GDP per capita, trade flows, and inflation), and financial variables (depth and openness measures). The 2021 endpoint for the regression analysis reflects practical constraints in obtaining comprehensive macroeconomic data for more recent years.

3.2 Three Approaches to Mandate Classification

We implement our 11-point classification using three distinct approaches: traditional human-based or researcher coding; string-based coding; and LLM coding.

Human-based/Researcher coding follows established manual classification protocols where domain experts apply the scoring framework through detailed reading and interpretation of mandate documents.

String-based coding applies rule-based keyword matching algorithms that search for specific hierarchical terms ("sole," "primary," "secondary") and indicators of objectives ("price stability," "employment," "financial stability") within central bank legal texts. The algorithm employs comprehensive keyword dictionaries covering both modern monetary policy terminology and historical language from pre-1940 central banking documents, including terms such as "gold standard," "convertibility," and "sound currency" so as to capture institutional evolution across different monetary regimes (see Appendix Table A1 for a list of relevant terms).

LLM coding employs large language model analysis using the bidirectional and auto-regressive transformer (BART) architecture, a sequence-to-sequence neural network model, via zero-shot classification to systematically apply our classification scheme. This leverages recent advances in institutional text analysis that demonstrate 10-40 percent performance

improvements over traditional methods (Wang et al., 2023) while achieving inter-coder reliability coefficients above 0.990 (Chae and Davidson, 2025).[1] This automated approach addresses key limitations of human coding, such as fatigue effects, individual interpretive biases and inconsistency across large samples, while preserving contextual understanding beyond the simple keyword detection relied on in string-based coding (Heseltine and Clemm von Hohenberg, 2024).

Appendix Tables A2 and A3 provide details on this LLM implementation. The approach offers several methodological advantages, as detailed in the computer science literature on institutional text classification. Automated approaches such as large language models enable consistent application of our classification framework across 120 countries and 8,277 observations. The BART architecture receives contextually enriched prompts along the lines of the following: *"The following is a central bank mandate text that should be classified according to its objectives and hierarchy. Price stability encompasses both inflation control and exchange rate stability objectives,"* ensuring consistent interpretation of our theoretical framework.

4 Method Validation: Reliability and Agreement Across Classification Approaches

4.1 Temporal Consistency and Method Divergence

An advantage of our approach is that we can compare scoring of mandates across methods. To repeat, we completed the researcher-based coding before implementing the two automated methods, and did not revise the researcher-based coding subsequently. Similarly, we do not use the researcher-based coding as inputs for pre-training the automated models. Thus, we can treat these three approaches to classifying central bank mandates as independent from one another and compare the results on that basis.

[1]The BART implementation uses the Hugging Face transformers pipeline with 11 candidate labels corresponding to our classification scores, each with detailed descriptions to guide model interpretation.

Temporal disaggregation in Figure 1 sheds light on the reliability of the three methods. Both LLM coding and string-based coding show an uptick in mandate scores after World War I, when a cluster of newly established central banks was created with statutes that prioritized exchange rate stability under the gold standard. That uptick is more subdued in our human-agent coding. We then see a second increase in average mandate scores in the 1990s, this time coincident with the adoption of inflation targeting by a growing number of countries. In this instance the uptick is most pronounced according to the human-agent scoring.

All three coding schemes show the prevalence of single mandates then stabilizing or regressing slightly after the Global Financial Crisis (GFC). Human-agent coding and string-based scores track one another closely, especially between 1950 and 1990s, but no longer to the same extent after the GFC. After the turn of the century and especially after the GFC, the LLM-based scores are no longer higher than the human-agent based scores, as they had been for the entirety of the nine preceding decades.

The post-1990 change in the relationship between the expert-based scores on the one hand and the two automated scoring methods on the other is noticeable. It is unclear whether this divergence reflects the ability of human researchers to pick up nuances that escape automated methods, or whether human agents may be influenced other statements and aspects of the broader context affecting the behavior of central banks not specifically included in the database. The fact that the string-based method diverges from the two other (researcher- and LLM-based) methods after 1990 may indicate that this dictionary-based method fails to capture important nuances associated with the spread of inflation targeting regimes.

Table 3 provides comparative statistics on the distribution of mandate scores for the entire period of analysis (1900-2024). We observe systematic differences in mean scores: LLM (0.564), Researcher (0.493), and String (0.425). It is possible that human coders understate price stability orientation relative to LLMs due to conservative interpretation

or fatigue. Interestingly, simple string-based coding understates price stability orientation even more. It may be that the string-based approach picks up keywords associated multiple objectives, such as "employment" or "financial," in a context where these words pertain to something other than the central bank's objective strictly defined. In addition, variation across country-year observations (as captured by the coefficient of variation of the score) is less according to the LLM coding.

Table 4 shows summary statistics for the three broad classifications over the period 1970-2024. We focus on this time-frame for capturing the modern era of central banking. The main reason our remaining analysis focuses on this period is due to data availability. Consistent macroeconomic data needed for our empirical analysis, including unemployment rates, inflation measures, and financial indicators, are unavailable before 1970 for many countries.

The researcher-coded classification shows the most dual mandates (28.6 percent), while the string-based classification finds the fewest (3.1 percent), while the LLM-coded classification lies in between (11.2 percent). Despite these level differences, all three methods document the same secular trend: a pronounced shift toward price stability dominance in recent decades, following the inflationary crises of the 1970s-1980s and the subsequent adoption of inflation targeting frameworks.

Much of the interest in central bank mandates revolves around comparisons of single or multiple objectives. It is thus hard to know where to place central banks whose objectives are not specified. The prevalence of unspecified mandates varies by classification method: 653 observations (11.8 percent) for researcher-coded, 1,790 observations (32.3 percent) for string-based, and 72 observations (1.3 percent) for LLM-coded classifications out of 5,545 country-year observations from 1970-2024. In the empirical analysis of Section 5 we therefore drop these observations from the sample.

4.2 Income-Level Validation

Figure 2 shows the share of single or prioritized objectives, again disaggregated into high income, high-middle income, low-middle income, and low-income country groupings, in this case over time. All groupings display broadly similar patterns over time, with upticks in the prevalence of single mandates in the 1990s and then a leveling off or slight reversal in their prevalence after the GFC.

Again, however, the pattern of agreement and disagreement between coding methods varies by income group. For high-income countries, we see the same patterns as for the full sample of countries: human agent and string-based coding look broadly similar until the late 20th century; LLM models consistently produce higher scores until that time, after which human-agent-based scores rise sharply, to the point where they exceed LLM-based scores. Other income groupings, and especially the upper-middle-income group, generally show closer agreement between the two automated methods. An exception is the low-income-country grouping, where the two automated methods diverge toward the beginning of the period, and where the string-based coding shows no increase in the mandate score over the entire period (whereas both the LLM score and human-agent-based score do show such increases).

4.3 Cross-Method Agreement Patterns

Figure 3 shows a pairwise scatter analysis supporting the idea that these implementation differences across methods are consistent rather than random. LLM coding consistently provides higher scores than human approaches. It is possible that this automated method captures implicit price stability priorities that a conservative human interpretation may understate. String-based methods consistently and not just randomly produce the lowest scores. High-income countries cluster closer to the diagonal, indicating better cross-method agreement. Overall correlations are in the 0.24-0.43 range, confirming moderate but imperfect agreement. Deviations from diagonals again suggest systematic implementation differences

rather than random measurement error.

Cross-validation in Figure 4 confirms, as noted above, that implementation agreement varies by institutional development. High-income countries show the strongest correlations (0.42-0.53), likely reflecting standardized legal frameworks that reduce interpretive differences. Lower-income countries exhibit weaker agreement (0.13-0.38), suggesting greater institutional ambiguity that challenges both human and automated coding. That said, that the correlations are significantly positive suggests that each coding method captures different aspects of institutional complexity rather than pure measurement error.

4.4 Implications for Economic Analysis

These reliability assessments have direct implications for our economic analysis. The string-based method's have limitations to identify dual mandates means it cannot reliably test theories about employment-inclusive monetary policy. Its systematic underscoring of price stability dominance would bias any analysis toward finding weaker institutional effects.

Therefore, in Section 5, we privilege results that are robust across the two reliable methods (researcher and LLM), while presenting string-based results primarily to illustrate the dangers of naive keyword matching. When researcher and LLM methods agree on a relationship's direction and significance—as with GDP per capita, central bank independence, and trade openness—we can be confident in the finding despite measurement challenges.

5 Determinants of Central Bank Mandate Choice

5.1 Empirical Strategy

We now use all three scoring methods to analyze the determinants of differences in central bank mandates across countries and over time.

Sample Construction and Ancillary Variables The analysis employs a panel dataset of 2,558 country-year observations from 91 countries spanning 1970-2021. In contrast to tables on previous sections, we restrict the sample period to end in 2021, as this cutoff year maximizes data availability for macroeconomic indicators. Sample construction required the availability of (1) a measure of central bank independence, (2) core macroeconomic variables (lagged log GDP per capita, lagged trade openness, past inflation), (3) valid mandate classifications from at least one of the three methods, and (4) countries with at least 30 years of observations to ensure sufficient within-country variation for reliable fixed effects estimation.

Dependent Variables: Mandate Classification. For simplicity of exposition and to facilitate subsequent regression analysis, we consolidate our more comprehensive central bank mandate classification system into three principal mandate categories that capture the essential distinctions. We motivate this more compact classification scheme with reference to a simple framework that explicitly links central bank mandate structures to policy settings. Following previous literature (Clarida et al., 1999; Woodford, 2003; Walsh, 2003), we model central banks as minimizing an intertemporal loss function. We extend the standard framework to incorporate varying weights on different objectives based on central bank mandate structures.

Consider a central bank with a loss function of the form:

$$L_t = E_t \sum_{i=0}^{\infty} \beta^i \left[\alpha_\pi(\pi_{t+i} - \pi^*)^2 + \alpha_u(u_{t+i} - u^*)^2 + \alpha_f f_{t+i}^2 + \alpha_a a_{t+i}^2\right] \tag{1}$$

where π_{t+i} is inflation at time $t+i$, π^* is the central bank's inflation target, u_{t+i} is the unemployment rate, u^* is the natural rate of unemployment, f_{t+i} represents financial stability concerns, and a_{t+i} captures all other potential objectives. The parameters α_π, α_u, α_f, and α_a represent the weights that the central bank places on deviations from each objective, while β is the discount factor.

This simple framework motivates our adoption of a more compact mandate schema:

1. **Price Stability Dominant (PSD) Mandate:**

   This category encompasses classes 1.00-0.60 from the classification scale, where price stability either stands as the sole objective or has clear primacy over other considerations.

2. **Dual/Employment-Inclusive (DEI) Mandate:**

   This category encompasses classes 0.50-0.40 from the classification scale, where employment/labor market outcomes are explicitly included as objectives alongside price stability, either as co-equal (class 0.50) or with some degree of subordination (class 0.40).

3. **Multiple/Balanced Objectives (MBO) Mandate:**

   This category primarily covers classes 0.30-0.00, where multiple objectives coexist without a strict hierarchy (class 0.30), employment dominates price stability (class 0.20), or price stability is absent altogether (classes 0.10-0.00).

Estimating Equation. We estimate the following linear probability model with country fixed effects:

$$M_{i,t}^{j} = \alpha_i + \beta' \mathbf{X}_{i,t} + \epsilon_{i,t} \tag{2}$$

where $M_{i,t}^{j} \in \{0, 1\}$ is a binary indicator equal to 1 if country i at time t has mandate type $j \in \{PSD, DEI, MBO\}$ and 0 otherwise; α_i represents country fixed effects; and $\mathbf{X}_{i,t}$ is a vector of explanatory variables including lagged log GDP per capita, central bank independence (LVAU measure), lagged inflation, lagged log trade openness (imports plus exports as share of GDP), log financial depth (private credit to GDP ratio), and the Chinn-Ito index of financial openness. Extended specifications augment $\mathbf{X}_{i,t}$ with legal origin dummies and government effectiveness measures. The model is estimated separately for each mandate type and classification method using ordinary least squares with heteroskedasticity-robust standard errors.

5.2 Results

5.2.1 Descriptive Patterns Across Classification Methods

Table 5 compares the ancillary variables used in our analysis across the three mandate types (PSD, DEI and MBO) resulting from each of the three classification methods (researcher, string and LLM coded). It points to some intriguing differences across classification methods. The researcher- and string-based methods both suggest that countries where central banks have price stability-dominant mandates have higher per capita incomes, but the LLM-coding does not suggest that this is the case. The researcher- and string-based methods similarly both suggest that financial depth is associated with price stability-dominant mandates, but the LLM coding associates it with dual mandates. The researcher- and LLM-based methods suggest that countries with price stability-dominant mandates have more independent central banks, but the string-based coding instead associates dual mandates with the highest likelihood of an independent central bank. The researcher- and LLM-based methods similarly both associate low levels of trade openness with dual mandates, but the string-based coding associates it with multiple objectives. According to the two automated methods, countries with history of low inflation and that are financially open are most likely to have dual mandates; according to the researcher-based coding, low past inflation and financial openness are associated with price stability-dominant mandates.

These descriptive patterns reveal a fundamental challenge for the estimation of the role of determinants in shaping mandates: the string-based method identifies dual mandates in only 59 observations across 4 countries, compared to 823 observations across 33 countries for researcher-coding and 287 observations across 17 countries for LLM-coding.[2] Given these stark differences in both sample composition and descriptive statistics, readers would be forgiven for expecting that regression results will be inconsistent across methods. As we shall see, however, the regression analysis reveals a reasonable degree of agreement on core relationships despite these obvious measurement issues.

[2]The countries categorized as having DEI mandates are: Bolivia, Namibia, Peru and Philippines.

5.2.2 Analysis of Mandate Determinants

Tables 6, 7 and 8 report those regression results. All equations are estimated by ordinary least squares (OLS) with country fixed effects included. The dependent variables are binary indicators taking the value of 1 if a country has the specified mandate type (price stability dominant, dual/employment-inclusive, or multiple/balanced objectives, respectively) and 0 otherwise. Robust standard errors are reported in parentheses.

The most robust finding in our analysis is the GDP per capita effect across all three classification methods. Wealthier countries systematically gravitate toward price stability-dominant mandates (positive coefficients ranging from 0.037 to 0.160) while avoiding both dual mandates (negative coefficients from -0.003 to -0.073) and multiple objectives (negative coefficients from -0.046 to -0.119). This pattern holds regardless of classification method, suggesting that economic development level is the strongest and most reliable predictor of central bank mandate choice.

While the direction of central bank independence effects shows some agreement across methods, the magnitude differences are noteworthy. For price stability-dominant mandates, the researcher-coded method shows an effect size (1.863) more than twice that of the LLM method (0.789), with the string-based method in between (1.221). These magnitude differences suggest that each classification method may be capturing different dimensions of what constitutes independence in practice, with the researcher-coded approach potentially incorporating implicit aspects of independence not captured in formal legal texts.

All three methods indicate that trade openness makes multiple balanced objectives more likely and both price stability dominant and dual mandates less likely, although only in the case of the LLM coding is the reduced likelihood of a price-dominant objective statistically significant at standard confidence levels. The consistent pattern for trade openness across mandate types reveals an important economic logic: open economies appear to require more flexible mandate structures. The positive association with multiple balanced objectives coupled with negative associations (where significant) with both price stability dominance

and dual mandates suggests that exposure to external shocks necessitates broader policy flexibility rather than rigid adherence to specific targets.

Other macroeconomic variables show less consistent patterns. The researcher and LLM-based methods both suggest that increased financial depth makes a price-dominant objective more likely and a dual mandate less likely. The string-based classification suggests the opposite, however. All of these coefficient estimates are statistically significant at standard confidence levels.

Results for past inflation are inconsistent. The researcher-based coding suggests that past inflation makes price stability-dominant mandates less likely but multiple balanced objectives more likely; the string-based method that it makes price stability-dominant mandates more likely and dual mandates less likely; and the LLM method that it has no significant effect. The same is the case of financial openness, which is positively associated with price-dominant objectives according to the researcher-based coding, negatively associated according to the string-based coding, and insignificantly associated according to the LLM coding, and positively associated with the likelihood of observing a dual mandate according to the two automated methods but negatively associated with this likelihood according to the researcher-based coding. This absence of a consistent pattern may indicate that mandate choices reflect preferences and traditions over institutional design rather than responses to macroeconomic and inflationary crises.

These exceptionally high R-squared values for cross-country institutional analysis deserve emphasis. The ability to explain 73-84 percent of the variation in mandate choices suggests that these institutional arrangements are not random historical accidents or purely path-dependent outcomes, but rather systematic responses to country characteristics. The stability of these R-squared values across different model specifications further reinforces the robustness of the core relationships.

A critical consideration in interpreting these results is the stark difference in composition of the dependent variable across methods, particularly for dual mandates. The string-based

method identifies dual mandates in only 59 observations across 4 countries, compared to 823 observations across 33 countries for the researcher-coded method and 287 observations across 17 countries for the LLM method. This dramatic difference in identifying dual mandates—with the string-based method effectively failing to detect them in most countries where other methods find them—may explain many of the sign reversals and inconsistencies observed. The string-based method appears to be identifying an almost entirely different phenomenon when it classifies mandates as "dual." These disagreements between methods, particularly the string-based method's difficulty in identifying dual mandates, may themselves be informative about the challenges of extracting institutional meaning from legal texts through mechanical keyword searches versus contextual interpretation.

Despite these inconsistencies, two patterns emerge with clarity. First, economic development (GDP per capita) and central bank independence consistently predict movement toward price stability-dominant mandates across all classification methods, while trade openness consistently predicts an increased likelihood of multiple balanced objectives. Second, the systematic nature of these relationships—as evidenced by the high R^2 values—suggests that mandate choices follow predictable patterns based on country characteristics rather than representing arbitrary institutional choices.

6 Conclusions

We have provided new quantitative measures and categorizations of central bank mandates spanning the period from the dawn of the 20th century to today, for as many as 120 central banks, depending on date of establishment. We classify central banks according to statutes and other relevant legislation on an 11 point scale that captures the full spectrum of mandates, from those focused exclusively on price stability to those with multiple objectives without priority and without price stability as a concern. We use three alternative methods for identifying central bank objectives and classifying mandates: a traditional human-based

coding, where individual documents are closely read by the researcher; a string-based algorithm that searches for indicators of objectives and hierarchical terms within central bank legal texts; and a large language model employing bidirectional and auto-regressive transformer architecture.

While the results of the three classifications do not agree exactly, they broadly concur in documenting trends over time in average mandate score. All three methods show a modest increase in price stability-dominant mandates after World War I and a sharper increase in the 1990s. They suggest that the recent increase in price-dominant mandates is common to high-income, middle-income and low-income countries alike.

In terms of average mandate score, the researcher-based and string-based coding agree most strongly before 1990, researcher-based and LLM-based coding thereafter. This shift coincides with the spread of inflation-targeting mandates to additional countries. We suspect that the string-based method does least well at picking up this shift, since it likely captured related and unrelated mentions of inflation, price stability and kindred terms both before and after the advent of these new regimes, whereas the researcher coding and LLM method better pick up this nuance.

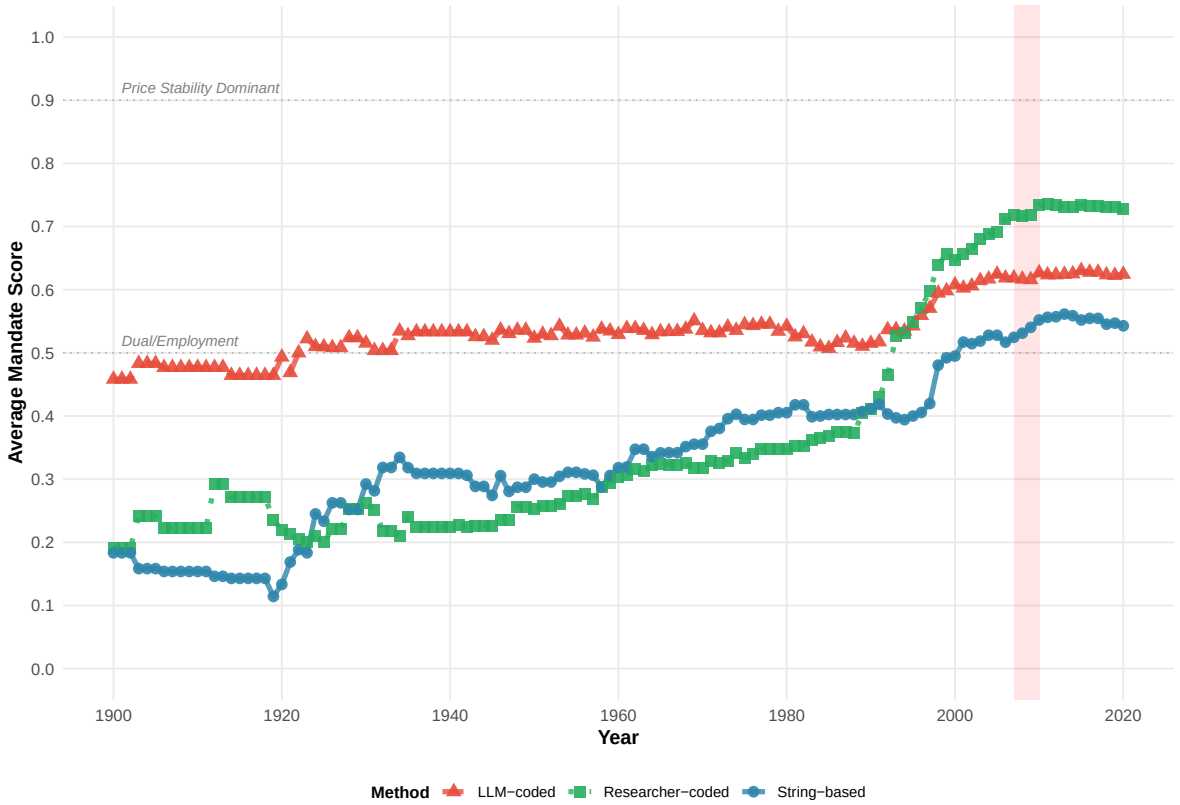
If overall time trends are broadly similar across coding methods, other aspects are not. Dividing mandates into price stability-dominant, dual and multiple-balanced categories shows the extent of disagreement between the three coding methods. The divergence is especially pronounced for dual mandates, where the string-based method identifies only 59 country/year cases, the LLM-based coding 287 cases and the researcher-based coding 823 cases for the post-1970 period.

Notwithstanding such differences, we consistently find that economic development (GDP per capita) and central bank independence predict movement toward price stability-dominant mandates across all classification methods, while trade openness predicts an increased likelihood of multiple balanced objectives. The systematic nature of these relationships—as evidenced by the high R^2 values—suggests that mandate choices follow predictable patterns

based on country characteristics rather than representing arbitrary institutional choices.

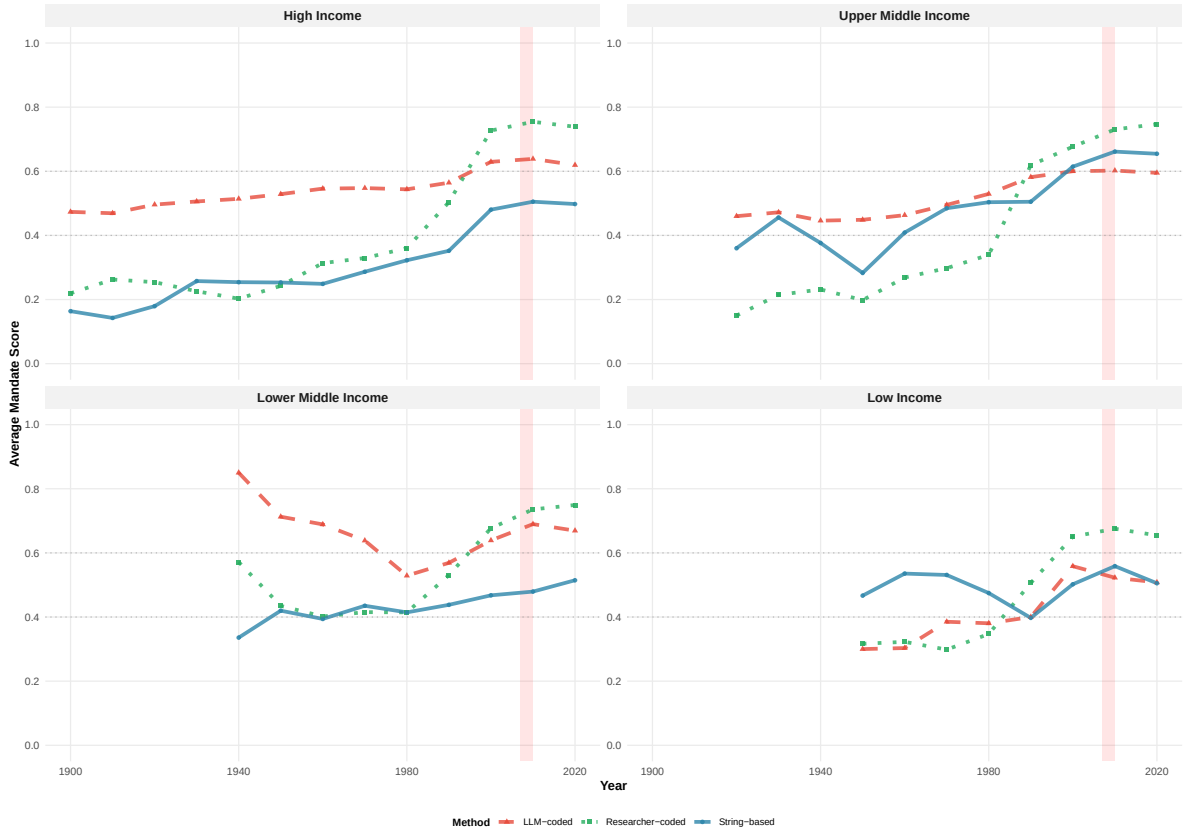
Less positively, other results are not robust across classification methods. This is reason for caution when making use of such institutional indices constructed from legal texts.

Figure 1: Average Mandate Scores Comparison by Implementation Methods, 1900-2024



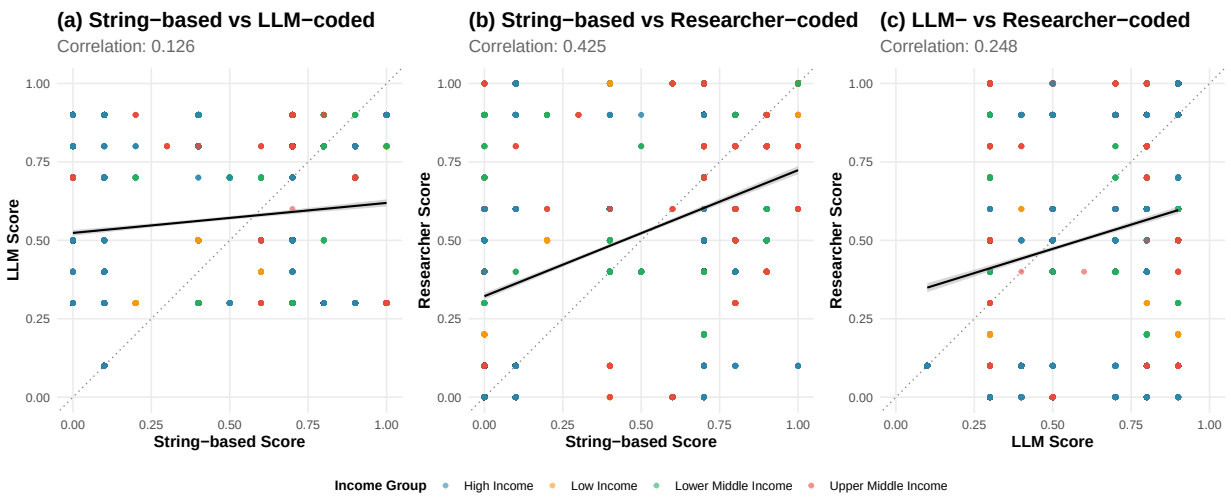
Notes: The figure shows the average mandate scores calculated per year by implementation method. Y-axis represents mandate orientation (0=employment focus, 1=price stability focus). Lines use different colors and styles by method: LLM-coded (red, dashed), Researcher-coded (green, dotted), String-based (blue, solid). Shaded region marks Global Financial Crisis period (2007-2010). Sample composition varies by data availability across time periods.

Figure 2: Implementation Comparison Across Country Income Levels, 1900-2024



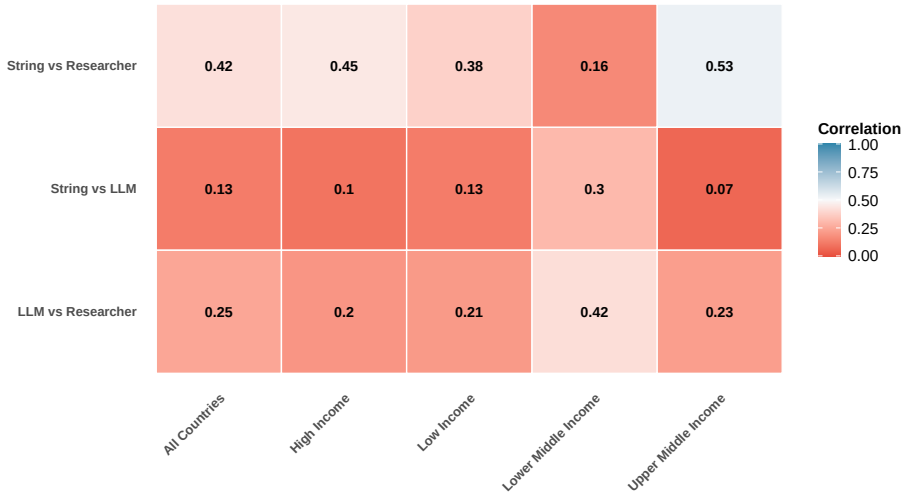
Notes: The figure displays mandate scores by World Bank income classification and implementation method calculated per decade. Each panel represents one income group with time series by coding approach. Y-axis shows mandate orientation (0=employment focus, 1=price stability focus). Shaded regions indicate Global Financial Crisis (2007-2010). Line styles match Figure 1 conventions. Sample coverage varies by income group and time period based on data availability.

Figure 3: Cross-Implementation Validation of Classification System



Notes: Scatter plots compare mandate scores across implementation method pairs. Each point represents a country-year observation with colors indicating World Bank income group (blue=high income, red=upper middle, green=lower middle, orange=low income). Diagonal reference lines show perfect agreement between methods. Correlation coefficients in subtitles quantify linear association. Sample includes observations with valid scores under both plotted methods. The period of analysis is 1900-2024.

Figure 4: Cross-Validation of Mandate Classification by Country Income Level



Notes: The figure shows a heatmap if pairwise correlations between the three implementation methods of the classification system, stratified by World Bank income groups. Color intensity indicates correlation strength (darker = higher correlation). Each cell represents the linear association between two coding approaches within the specified income group. Sample includes all country-year observations 1900-2024 with valid classifications under all three methods.

Table 1: Central Bank Mandate Classification Categories

Score	Mandate Category	Description
A. Price Stability Dominance (PSD)		
1.00	Price Stability Sole	Price stability (domestic price stability or exchange rate stability) as the sole objective, with no other objectives mentioned
0.90	Price Stability Dominant	Price stability (domestic price stability or exchange rate stability) as the primary objective, with explicit final authority over any secondary objectives
0.80	Price Stability Primary + Financial	Price stability (domestic price stability or exchange rate stability) is primary; financial stability is an explicitly secondary and compatible objective
0.70	Price Stability Primary + Employment/Other	Price stability (domestic price stability or exchange rate stability) is primary; non-compatible objectives (e.g., employment or growth) are explicitly secondary
B. Dual Employment Inclusive (DEI)		
0.60	Price Stability + Compatible (Co-equal)	Price stability (domestic price stability or exchange rate stability) is a co-equal objective alongside other compatible goals such as financial stability; no clear hierarchy
0.50	Dual Mandate	Price stability (inflation/exchange rate) and non-compatible objectives (e.g., employment or growth) are both explicit and co-equal
C. Multiple Balanced Objectives (MBO)		
0.40	Multiple Balanced Objectives	Multiple objectives, including price stability (inflation/exchange rate), but with no clear prioritization or hierarchy
0.30	Compatible Objective Primary	Compatible objectives (e.g., financial stability) are primary; price stability (inflation/exchange rate) is explicitly secondary
0.20	Employment/Growth Primary	Non-compatible objectives (e.g., employment or growth) are primary; price stability (inflation/exchange rate) is explicitly secondary
0.10	No Clear Mandate	No clear mandate or objectives are stated
0.00	Non-Price Focus	Objectives are stated, but price stability (inflation or exchange rate) is not mentioned at all

Note: This classification scheme is derived from a rule-based string matching algorithm that examines central bank mandate texts for specific keywords and hierarchical relationships. "Price stability" encompasses both inflation control and exchange rate stability objectives. Scores from 1.00 to 0.70 represent mandates where price stability is dominant. Scores of 0.50 and 0.40 represent mandates that explicitly include employment alongside price stability. Scores from 0.30 to 0.00 represent mandates with multiple balanced objectives or those that do not prioritize price stability.

Table 2: Comparative Coding of "Objectives" in Central Bank Independence Indices

Index	Coding of Objectives	Exact Scores / Categories	Source
Cukierman, Webb & Neyapti (1992)	Six-tier coding of statutory objectives; prominence of price stability and conflict resolution.	**1.00** = Price stability is the major or only objective in the charter, and the central bank has the final word in case of conflict with other government objectives. **0.80** = Price stability is the only objective. **0.60** = Price stability is one goal, with other compatible objectives, such as a stable banking system. **0.40** = Price stability is one goal, with potentially conflicting objectives, such as full employment. **0.20** = No objectives stated in the bank charter. **0.00** = Stated objectives do not include price stability.	Cukierman et al. (1992), Table 1
Grilli, Masciandaro & Tabellini (1991)	Political independence includes a binary 'final objectives' item: whether monetary stability is a statutory requirement.	**1** = Monetary stability explicitly mandated. **0** = Not mandated.	GMT (1991), Economic Policy
Romelli (2024)	Five-tier scale, refining CWN by distinguishing non-conflicting vs. conflicting additional objectives.	**1.00** = Price stability single/primary objective. **0.75** = Price stability with non-conflicting objectives, no priority. **0.50** = Price stability + other (possibly conflicting) goals incl. financial stability, no priority. **0.25** = Price stability + growth/development, no priority. **0.00** = Price stability not included.	Romelli (2024)
Dincer & Eichengreen (2006, 2014); Dincer, Eichengreen and Martinez (2024)	Adopt CWN objectives coding in cross-country legal datasets. Objectives dimension is taken directly from CWN's six categories.	Same six categories and scores as Cukierman et al. (1992): 1.00 to 0.00.	Dincer & Eichengreen (2006, 2014)
QoG / Garriga dataset (2016, 2025 update)	Reproduces CWN objectives coding verbatim in global panel.	Identical six categories and scores as Cukierman et al. (1992).	Garriga (2016, 2025)

Table 3: Summary Statistics by Classification Implementation Method, 1900-2024

Classification Method	Mean (1)	Std. Dev. (2)	Min. (3)	Max. (4)	Num. Obs. (5)
Researcher-coded	0.493	0.325	0.0	1.0	7,729
String-based	0.425	0.343	0.0	1.0	7,729
LLM-coded	0.564	0.261	0.1	0.9	7,729

Notes: Three implementation approaches apply the 11-point classification system from Table 1 to central bank mandates of 120 countries. Researcher-coded represents traditional expert manual classification. String-based employs rule-based keyword matching. LLM-assisted coding utilizes large language model analysis with the BART (Bidirectional and Auto-Regressive Transformer). The observations are at the country-year level.

Table 4: Central Bank Mandate Classifications by Time Period, 1900-2024

	Classification Method		
Mandate Category	Researcher-coded	String-based	LLM-coded
Panel A: 1900-2024			
Price Stability Dominance (PSD)	0.285	0.481	0.467
Multiple Balanced Objectives (MBO)	0.188	0.078	0.382
Dual Employment Inclusive (DEI)	0.27	0.032	0.111
Total Observations	*7729*	*7729*	*7729*
Total Countries	*120*	*120*	*120*
Panel B: 1900-1970			
Price Stability Dominance (PSD)	0.07	0.258	0.399
Multiple Balanced Objectives (MBO)	0.094	0.083	0.383
Dual Employment Inclusive (DEI)	0.229	0.036	0.108
Total Observations	*2249*	*2249*	*2249*
Total Countries	*65*	*65*	*65*
Panel C: 1970-1980			
Price Stability Dominance (PSD)	0.042	0.436	0.425
Multiple Balanced Objectives (MBO)	0.307	0.072	0.418
Dual Employment Inclusive (DEI)	0.315	0.042	0.115
Total Observations	*791*	*791*	*791*
Total Countries	*76*	*76*	*76*
Panel D: 1980-2000			
Price Stability Dominance (PSD)	0.19	0.481	0.419
Multiple Balanced Objectives (MBO)	0.304	0.081	0.449
Dual Employment Inclusive (DEI)	0.338	0.033	0.111
Total Observations	*1983*	*1983*	*1983*
Total Countries	*118*	*118*	*118*
Panel E: 2000-2024			
Price Stability Dominance (PSD)	0.571	0.662	0.561
Multiple Balanced Objectives (MBO)	0.157	0.075	0.328
Dual Employment Inclusive (DEI)	0.244	0.027	0.111
Total Observations	*2964*	*2964*	*2964*
Total Countries	*119*	*119*	*119*

Notes: The table shows the distribution of country-year observations classified by central bank mandate type under three methodologies across different time periods. The categories are: Price Stability Dominant (PSD): mandate score ≥ 0.70; Multiple/Balanced Objectives (MBO): mandate score 0.30-0.69; and Dual/Employment-Inclusive (DEI): mandate score < 0.30. The three classification approaches are: String-based (rule-based keyword matching), LLM (large language model analysis using BART (Bidirectional and Auto-Regressive Transformers), and Researcher-Coded (expert manual coding). Proportions are shown as decimals. Total Observations represent country-year pairs with available mandate data, while Total Countries represent unique countries with mandate data in each period.

Table 5: Descriptive Statistics of Determinant Variables of Central Bank Mandate Categories, 1970-2021

	Mean Values			p-values		
	PSD	DEI	MBO	PSD vs DEI	PSD vs MBO	DEI vs MBO
Panel A: Researcher-coded						
GDP per Capita (log, t-1)	9.013	8.854	8.670	0.016	0.016	0.000
CB Independence	0.631	0.392	0.416	0.002	0.000	0.000
Past Inflation	4.734	7.983	64.466	0.010	0.000	0.006
Trade Openness (log, t-1)	4.341	4.151	4.319	0.000	0.000	0.486
Financial Depth (log)	3.951	3.779	3.513	0.000	0.000	0.000
Financial Openness	0.968	0.413	0.205	0.025	0.000	0.000
Observations	1,114	823	429			
Countries	63	33	21			
Panel B: String-based						
GDP per Capita (log, t-1)	8.981	8.098	7.556	0.000	0.000	0.000
CB Independence	0.540	0.680	0.378	0.000	0.000	0.000
Past Inflation	20.898	6.553	6.846	0.848	0.023	0.023
Trade Openness (log, t-1)	4.352	3.928	3.769	0.015	0.000	0.000
Financial Depth (log)	3.922	3.629	3.347	0.002	0.000	0.000
Financial Openness	0.702	1.256	-0.612	0.000	0.000	0.000
Observations	1,528	59	195			
Countries	74	4	10			
Panel C: LLM-coded						
GDP per Capita (log, t-1)	8.867	9.694	8.797	0.000	0.000	0.209
CB Independence	0.536	0.492	0.433	0.000	0.003	0.000
Past Inflation	14.322	4.082	24.064	0.055	0.001	0.371
Trade Openness (log, t-1)	4.184	4.060	4.414	0.000	0.002	0.000
Financial Depth (log)	3.817	4.119	3.753	0.000	0.000	0.076
Financial Openness	0.505	1.246	0.686	0.000	0.000	0.008
Observations	1,406	287	794			
Countries	65	17	38			
Grand Total	2,558			Countries: 91		

Notes: This table reports mean values of macroeconomic determinants by central bank mandate category and statistical tests of differences between categories across three classification methods. The analysis uses 2,558 country-year observations from 91 countries over the period 1970-2021 where at least one classification method provides valid mandate categorizations after excluding "No Clear Mandate" classifications. Columns (1)–(3) report mean values by mandate category: PSD = Price Stability Dominance; DEI = Dual Employment Inclusive; MBO = Multiple Balanced Objectives. Columns (4)-(6) report p-values from two-sample t-tests comparing mandate categories. The variables shown are: GDP per Capita (log, t-1), CB Independence (LVAU measure), Past Inflation (lagged one year), Trade Openness (log, t-1), Financial Depth (log of credit to private sector/GDP), and Financial Openness (Chinn-Ito-Kaopen index). The three classification approaches are: Researcher-coded (expert manual classification using Dincer categories), String-based (rule-based keyword matching), and LLM-coded (large language model analysis using BART). Sample construction required: (1) central bank independence data (LVAU measure), (2) non-missing lagged GDP per capita, lagged trade openness, past inflation, financial depth, and financial openness, and (3) valid classifications from at least one method. The grand total row shows the consistent sample of 2,558 observations used across all three classification methods.

Table 6: Determinants of Having a Price Dominant Objective (PSD) for a Central Bank - Panel Data Analysis

	Dependent variable: PSD (=1)								
	Researcher-Coded			String-Based			LLM-Coded		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP per Capita (log, t-1)	0.120***	0.121***	0.037*	0.119***	0.124***	0.160***	0.041**	0.088***	0.040
	(0.016)	(0.017)	(0.020)	(0.020)	(0.022)	(0.025)	(0.020)	(0.021)	(0.024)
CB Independence	1.863***	1.863***	1.765***	1.218***	1.221***	1.245***	0.789***	0.821***	0.810***
	(0.046)	(0.047)	(0.048)	(0.058)	(0.058)	(0.061)	(0.057)	(0.057)	(0.059)
Past Inflation	−0.00005**	−0.00005**	−0.00004*	0.0001***	0.0001***	0.0001***	0.00002	0.00002	0.00001
	(0.00002)	(0.00002)	(0.00002)	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Trade Openness (log, t-1)		−0.004	−0.024		−0.015	−0.007		−0.140***	−0.149***
		(0.021)	(0.021)		(0.027)	(0.027)		(0.026)	(0.026)
Financial Depth (log)			0.099***			−0.051***			0.076***
			(0.013)			(0.016)			(0.016)
Financial Openness			0.012**			0.006			−0.021***
			(0.006)			(0.007)			(0.007)
Constant	−1.396***	−1.391***	−0.891***	−0.903***	−0.882***	−1.049***	0.064	0.256	0.420**
	(0.135)	(0.138)	(0.150)	(0.170)	(0.174)	(0.190)	(0.166)	(0.169)	(0.185)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558
R^2	0.843	0.843	0.847	0.747	0.747	0.748	0.764	0.767	0.770
Adjusted R^2	0.837	0.837	0.841	0.737	0.737	0.738	0.755	0.758	0.761

Notes: The table shows the regression coefficients from a regression estimated on a county-time panel dataset. The dependent variable equals 1 if the central bank has the mandate type specified in the title, zero otherwise. All regressions include country fixed effects. Columns (1), (4) and (7) show coefficients a model including core macroeconomic variables. The model in Columns (2), (5) and (8) adds legal origin dummies (socialist origin omitted). The model in Columns (3), (6) and (9) adds government effectiveness. The three classification approaches are indicated in the column headers: String-based (rule-based keyword matching), LLM-coded (large language model analysis using Bidirectional and Auto-Regressive Transformer architecture), and Researcher (expert manual coding). Robust standard errors in parentheses. Analysis dataset prepared with 2,558 observations from 91 countries, 1970-2021. Countries with fewer than 30 years of observations are excluded to ensure sufficient within-country variation for reliable fixed effects estimation.

Table 7: Determinants of Having a Dual-Employment Objective (DEI) for a Central Bank - Panel Data Analysis

	Dependent variable: DEI (=1)								
	Researcher-Coded			String-Based			LLM-Coded		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP per Capita (log, t-1)	0.032*	0.036**	0.073***	−0.003	−0.004	−0.042***	0.031***	0.035***	0.054***
	(0.016)	(0.018)	(0.021)	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)	(0.013)
CB Independence	−0.217***	−0.214***	−0.164***	0.147***	0.147***	0.077***	−0.166***	−0.162***	−0.174***
	(0.048)	(0.048)	(0.050)	(0.018)	(0.018)	(0.018)	(0.029)	(0.029)	(0.030)
Past Inflation	−0.00000	−0.00000	−0.00001	−0.0001***	−0.0001***	−0.0001***	−0.00001	−0.00001	−0.00000
	(0.00002)	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Trade Openness (log, t-1)		−0.013	−0.004		0.002	−0.008		−0.015	−0.012
		(0.022)	(0.022)		(0.008)	(0.008)		(0.013)	(0.013)
Financial Depth (log)			−0.041***			0.034***			−0.036***
			(0.013)			(0.005)			(0.008)
Financial Openness			−0.010*			0.021***			0.018***
			(0.006)			(0.002)			(0.004)
Constant	−0.099	−0.081	−0.322**	−0.085	−0.087	0.208***	−0.127	−0.107	−0.126
	(0.139)	(0.142)	(0.156)	(0.053)	(0.054)	(0.057)	(0.085)	(0.087)	(0.095)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558
R^2	0.813	0.813	0.814	0.738	0.738	0.755	0.846	0.846	0.849
Adjusted R^2	0.806	0.806	0.807	0.728	0.728	0.746	0.840	0.840	0.843

Notes: The table shows the regression coefficients from a regression estimated on a county-time panel dataset. The dependent variable equals 1 if the central bank has the mandate type specified in the title, zero otherwise. All regressions include country fixed effects. Columns (1), (4) and (7) show coefficients a model including core macroeconomic variables. The model in Columns (2), (5) and (8) adds legal origin dummies (socialist origin omitted). The model in Columns (3), (6) and (9) adds government effectiveness. The three classification approaches are indicated in the column headers: String-based (rule-based keyword matching), LLM-coded (large language model analysis using Bidirectional and Auto-Regressive Transformer architecture), and Researcher (expert manual coding). Robust standard errors in parentheses. Analysis dataset prepared with 2,558 observations from 91 countries, 1970-2021. Countries with fewer than 30 years of observations are excluded to ensure sufficient within-country variation for reliable fixed effects estimation.

Table 8: Determinants of Having Multiple Balanced Objectives (MBO) for a Central Bank - Panel Data Analysis

	Dependent variable: MBO (=1)								
	Researcher-Coded			String-Based			LLM-Coded		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP per Capita (log, t-1)	−0.095***	−0.119***	−0.046***	−0.069***	−0.088***	−0.085***	−0.021	−0.071***	−0.062***
	(0.014)	(0.015)	(0.018)	(0.010)	(0.011)	(0.013)	(0.019)	(0.021)	(0.024)
CB Independence	−0.520***	−0.536***	−0.466***	0.155***	0.142***	0.152***	−0.445***	−0.480***	−0.499***
	(0.041)	(0.041)	(0.042)	(0.030)	(0.030)	(0.031)	(0.055)	(0.055)	(0.057)
Past Inflation	0.0001***	0.0001***	0.0001***	−0.00000	−0.00000	−0.00000	−0.00001	−0.00001	−0.00000
	(0.00002)	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00003)	(0.00003)	(0.00003)
Trade Openness (log, t-1)		0.070***	0.087***		0.056***	0.057***		0.150***	0.151***
		(0.019)	(0.019)		(0.014)	(0.014)		(0.025)	(0.025)
Financial Depth (log)			−0.093***			−0.0004			−0.024
			(0.011)			(0.008)			(0.015)
Financial Openness			−0.001			−0.005			0.018***
			(0.005)			(0.004)			(0.007)
Constant	1.182***	1.085***	0.691***	0.458***	0.380***	0.345***	0.508***	0.301*	0.329*
	(0.120)	(0.122)	(0.133)	(0.087)	(0.088)	(0.097)	(0.161)	(0.163)	(0.179)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558	2,558
R^2	0.782	0.783	0.789	0.775	0.776	0.776	0.745	0.749	0.750
Adjusted R^2	0.774	0.775	0.781	0.766	0.768	0.768	0.735	0.739	0.740

Notes: The table shows the regression coefficients from a regression estimated on a county-time panel dataset. The dependent variable equals 1 if the central bank has the mandate type specified in the title, zero otherwise. All regressions include country fixed effects. Columns (1), (4) and (7) show coefficients a model including core macroeconomic variables. The model in Columns (2), (5) and (8) adds legal origin dummies (socialist origin omitted). The model in Columns (3), (6) and (9) adds government effectiveness. The three classification approaches are indicated in the column headers: String-based (rule-based keyword matching), LLM-coded (large language model analysis using Bidirectional and Auto-Regressive Transformer architecture), and Researcher (expert manual coding). Robust standard errors in parentheses. Analysis dataset prepared with 2,558 observations from 91 countries, 1970-2021. Countries with fewer than 30 years of observations are excluded to ensure sufficient within-country variation for reliable fixed effects estimation.

References

Andres Alonso-Robisco and José Manuel Carbó. Analysis of cbdc narrative of central banks using large language models. *Finance Research Letters*, 58:104373, 2023. doi: 10.1016/j.frl.2023.104373.

Christoph Bertsch, Isaiah Hull, Robin L. Lumsdaine, and Xin Zhang. Central bank mandates and monetary policy stances: Through the lens of Federal Reserve speeches. *Journal of Econometrics*, 249:105948, May 2025. ISSN 03044076. doi: 10.1016/j.jeconom.2025.105948. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407625000028.

Martin T. Bohl, Dimitrios Kanelis, and Pierre L. Siklos. Central bank mandates and monetary policy tone: Evidence from ecb and fed speeches. *Journal of International Money and Finance*, 2023.

Youngjin Chae and Thomas Davidson. Large language models for text classification: From zero-shot learning to nstruction-tuning. *Sociological Methods & Research*, 2025.

Thiago Christiano Silva, Kei Moriya, and Romain Veyrune. From Text to Quantified Insights. *IMF Working Papers*, 2025(109):1, 6 2025. ISSN 1018-5941. doi: 10.5089/9798229013802.001. URL http://dx.doi.org/10.5089/9798229013802.001.

R. Clarida, J. Galí, and M. Gertler. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37(4):1661–1707, 1999.

Alex Cukierman, Steven B. Web, and Bilin Neyapti. Measuring the Independence of Central Banks and Its Effect on Policy Outcomes. *The World Bank Economic Review*, 6(3):353–398, 1992. ISSN 0258-6770, 1564-698X. doi: 10.1093/wber/6.3.353. URL https://academic.oup.com/wber/article-lookup/doi/10.1093/wber/6.3.353.

Simon Dikau and Ulrich Volz. Central bank mandates, sustainability objectives and the promotion of green finance. *Ecological Economics*, 184:107022, June 2021. ISSN 09218009. doi:

10.1016/j.ecolecon.2021.107022. URL https://linkinghub.elsevier.com/retrieve/pii/S092180092100080X.

N. Dincer and B. Eichengreen. Central bank transparency: Where, why and with what effects? In J.-P. Touffut, editor, *Central Banks as Economic Institutions*. Edward Elgar, Cheltenham, 2008a.

N. Dincer and B. Eichengreen. Central bank transparency: Causes, consequences and updates. *Theoretical Inquiries in Law*, 11:75–123, 2010a.

N. Nergiz Dincer and Barry Eichengreen. Central bank transparency: Where, why, and to what effect? NBER Working Paper 13003, National Bureau of Economic Research, 2008b.

N. Nergiz Dincer and Barry Eichengreen. Central bank transparency: Causes, consequences and updates. *Theoretical Inquiries in Law*, 11(1):75–123, 2010b.

N Nergiz Dincer and Barry Eichengreen. Central bank transparency and independence: Updates and new measures. *34th issue (March 2014) of the International Journal of Central Banking*, 2018.

Barry Eichengreen and Andreas Kakridis, editors. *The spread of the modern central bank and global cooperation: 1919-1939*. Studies in macroeconomic history. Cambridge University Press, Cambridge, United Kingdom ; New York, NY, 2024. ISBN 978-1-00-936754-7.

Ricardo Duque Gabriel, Ioannis Lazopoulos, and Diana Lima. Financial stability mandates and central bank policy. *Journal of Central Banking Theory and Practice*, 11(3):5–31, 2022.

Ana Carolina Garriga. Central bank independence in the world: A new data set. *International Interactions*, 42(5):849–868, 2016. ISSN 0305-0629. Publisher: Taylor & Francis.

Ana Carolina Garriga and Cesar M. Rodriguez. Balancing Act or Policy Pitfall? The Effects

of Central Bank Dual Mandates. *SSRN Electronic Journal*, 2025. ISSN 1556-5068. doi: 10.2139/ssrn.5401879. URL https://www.ssrn.com/abstract=5401879.

Vittorio Grilli, Donato Masciandaro, and Guido Tabellini. Political and monetary institutions and public financial policies in the industrial countries. *Economic policy*, 6(13): 341–392, 1991. ISSN 1468-0327. Publisher: Oxford University Press Oxford, UK.

Michael Heseltine and Bernhard Clemm von Hohenberg. Large language models as a substitute for human experts in annotating political ext. *Research & Politics*, 11(2), 2024.

Emily Jackson and Simon Baily. The bank of england's new net zero mandate: Discourse analysis and participant interviews. *Environmental Politics*, 2023.

Wenzhe Li. What determine monetary policy mandates in central bank legislation? by legal tradition or economic considerations? *SSRN Working Paper*, 2024. Available at SSRN: https://ssrn.com/abstract=4938363.

F.J. Rodriguez-Rodriguez, L.A. Hierro, and A.J. Garzon. Fed and ecb reaction functions during quantitative easing: Three phases of monetary policy, both conventional and unconventional. *Journal of Policy Modeling*, 46(5):928–945, 2024. ISSN 0161-8938. doi: https://doi.org/10.1016/j.jpolmod.2024.03.003. URL https://www.sciencedirect.com/science/article/pii/S0161893824000292.

Davide Romelli. Trends in central bank independence: a de-jure perspective. *SSRN Electronic Journal*, 2024. ISSN 1556-5068. doi: 10.2139/ssrn.4716704. URL https://www.ssrn.com/abstract=4716704.

Alexandre Tombini. Fulfilling central bank mandates in times of high uncertainty. Speech at the Regional Central Bank Governance Forum, April 2025. BIS Chief Representative for the Americas. Available at: https://www.bis.org/speeches/sp250404.htm.

C. E. Walsh. *Monetary theory and policy*. MIT Press, 2003.

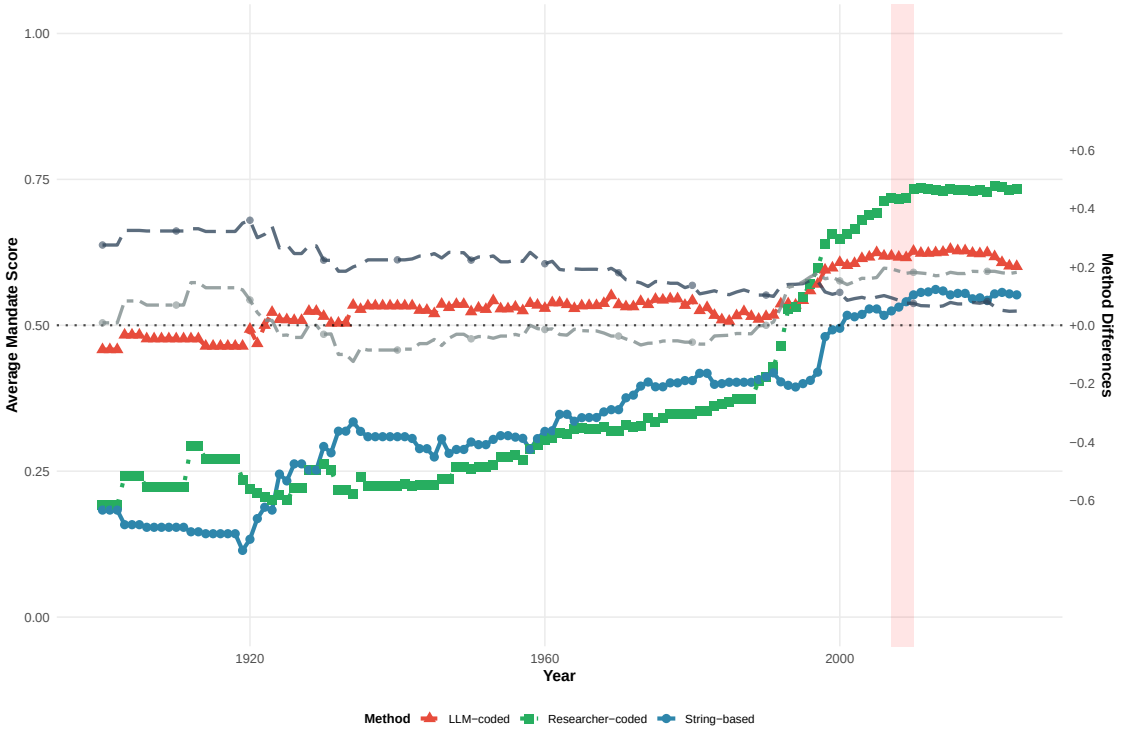
Xiaofei Wang et al. Text classification via large language models. *arXiv preprint arXiv:2305.08377*, 2023.

M. Woodford. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, 2003.

A Appendix

A.1 Additional Figures and Tables

Figure A1: Implementation Stability and Method Differences Over Time



Notes: Dual-axis figure shows mandate scores (left axis, colored lines) and cross-method differences (right axis, black lines). Primary series display average scores by implementation method using conventions from Figure 1. Secondary series show differences relative to string-based classification: LLM minus String (solid black line), Researcher minus String (dashed black line). Shaded region marks Global Financial Crisis (2007-2010).

Table A1: Keyword Dictionaries for String-Based Classification

Category	Keywords
Price Stability	price stability, stable prices, inflation, monetary stability, purchasing power, value of the currency, stability of prices, value of money, stable value, currency value, currency stability, exchange rate stability, stable exchange rate, exchange rate target, fixed exchange rate, pegged exchange rate, monetary value, domestic price stability, sound currency, sound money, convertibility, gold standard, metallic standard, specie payments, parity, monetary standard, circulation, stable exchange, note issue, currency board, currency reform, monetary reform, exchange parity, maintain the value
Employment	employment, labor market, job, full employment, unemployment, maximum employment, labor force, jobs, employment creation, job growth
Financial Stability	financial stability, banking stability, stable financial system, financial system stability, banking system stability, systemic risk, financial resilience, banking system, lender of last resort, orderly markets, financial panic, banking crisis, monetary panic
Other Objectives	economic growth, development, balance of payments, credit, sustainability, economic welfare, prosperity, economic development, output, gdp, production, commercial interests, trade, commerce, industrial development, national interests, fiscal agent, government finance, public finance, war finance, treasury operations
Sole	sole, only, exclusively, single, main objective, primary goal, unique, exclusive, singular, exclusive focus, principal aim, chief concern, paramount duty, first duty
Primary	primary, main, principal, predominant, overriding, foremost, without prejudice to, first, fundamental, chief, paramount, key, central, essential, priority, chiefly, primarily, first and foremost, above all, principally, first consideration
Secondary	secondary, subordinate, subsidiary, also, in addition, another, other, additional, further, supporting, ancillary, auxiliary, lesser, minor
Final Authority	final authority, independence, autonomous, overriding, independent, authority, autonomous, sovereignty, discretion, ultimate authority, sole discretion, at its discretion, full power, prerogative, exclusive right, complete authority
Co-equal	equally, balanced, dual, both, together with, alongside, as well as, coequal, co-equal, equal weight, equivalent, joint, concurrent, parallel, twin

Notes: This table presents the comprehensive keyword dictionaries used in our string-based classification algorithm. Keywords cover both modern monetary policy terminology and historical language from pre-1940 central banking documents to capture institutional evolution across different monetary regimes. The algorithm searches for these terms within mandate texts and applies hierarchical logic to determine appropriate classification scores from our 11-point system.

Table A2: LLM Classification Implementation Details

Parameter	Specification
Model Architecture	Facebook BART-Large-MNLI
Classification Type	Zero-shot Classification
Number of Labels	11 (corresponding to classification scores)
Input Processing	Contextually enriched prompts
Device Configuration	CUDA when available, CPU fallback
Token Length	Up to 1024 tokens per input
Temperature	Default (1.0)
Output Format	Single highest-scoring label

Notes: Implementation details for BART-based mandate classification. The model employs zero-shot classification across 11 candidate labels corresponding to our classification framework. Each label incorporates detailed descriptions to guide model interpretation, while text inputs are processed using contextual prompts that explicitly specify the central bank mandate classification task.

Table A3: Classification Logic Flow for String-Based Method

Score	Classification Logic
1.00	(Sole AND Price Stability AND NOT Employment AND NOT Financial Stability AND NOT Other) OR (Gold Standard/Currency Board AND NOT Employment AND NOT Other)
0.90	Price Stability AND (Sole OR Primary) AND Final Authority AND (Employment OR Financial Stability OR Other)
0.80	(Primary AND Price Stability AND Financial Stability AND NOT Employment AND NOT Other) OR (Price Stability AND Financial Stability AND Historical Banking Terms AND NOT Employment)
0.70	Primary AND Price Stability AND (Employment OR Other) AND Secondary AND NOT Co-equal
0.60	Price Stability AND Financial Stability AND NOT Employment AND NOT Primary AND NOT Secondary
0.50	Price Stability AND Employment AND (Co-equal OR (NOT Primary AND NOT Secondary))
0.40	Price Stability AND (Employment OR Financial Stability OR Other) AND NOT Primary AND NOT Secondary
0.30	Compatible objectives primary with price stability secondary, or historical banking system priority
0.20	Employment/Growth primary with price stability secondary
0.10	No clear mandate or empty text
0.00	Objectives stated but no price stability mentioned

Notes: Logical flow for string-based classification algorithm. The algorithm applies these rules sequentially, with early conditions taking precedence. Boolean operators (AND, OR, NOT) combine keyword presence indicators. Historical terms accommodate pre-1940 monetary policy language including gold standard and currency board arrangements.