OESTERREICHISCHE NATIONALBANK
EUROSYSTEM

OeNB Bulletin 2026/5
28 pages

Evenhanded surveillance? Geopolitical alignment and IMF appraisal tone

Does the IMF judge its member countries evenhandedly or do political allies get friendlier assessments? We measure the tone of IMF country reports and find no evidence to suggest that countries more closely aligned with the United States or the EU are given preferential treatment. Instead, the IMF's tone tracks economic fundamentals. When it comes to geopolitics, IMF surveillance thus appears evenhanded. Beyond geopolitics, countries scoring higher on a broad economic freedom index receive warmer IMF appraisals.

Authors

Martin Feldkircher
Oesterreichische Nationalbank,
International Economics
martin.feldkircher@oenb.at
and
Vienna School of International Studies
martin.feldkircher@da-vienna.ac.at

Theo Götze
Vienna School of International Studies
theo.goetze@da-vienna.at

Publication

13 November 2026

JEL classification
F33, F55, C23

Keywords
IMF surveillance, geopolitical alignment, economic ideology, ideal point distance, sentiment analysis, evenhandedness



IMF appraisal tone shows no geopolitical bias

IMF country reports do not adopt a friendlier tone for countries more closely aligned with the USA or the EU. This holds across four sentiment measures, from word lists to language models, and multiple robustness checks.



The economy speaks louder than politics

Inflation, economic growth and the current account shape what IMF staff write. As our model is able to pick up these effects, it would have the statistical power to identify any political bias as well. It does not detect any, however.



Economic freedom prompts warmer tone

Three out of four sentiment measures indicate a more favorable tone for countries that score higher on a broad economic freedom index that covers not only free-market policies but also the rule of law and price stability.

The views expressed are those of the authors and do not necessarily reflect those of the Oesterreichische Nationalbank or the Eurosystem.

Abstract[1]

In this paper, we assess whether country surveillance by the International Monetary Fund (IMF) is evenhanded, as the IMF intends it to be. More specifically, we test for two forms of potential bias, namely geopolitical and ideological bias. For this purpose, we utilise IMF Article IV staff appraisals and estimate, in a panel setting, whether the tone of these staff appraisals can be systematically related to either a measure that is based on joint voting in the United Nations General Assembly (a proxy for geopolitical alignment) or to variables that relate to free markets, free exchange rates, no capital controls and fiscal austerity (economic attributes historically associated with the IMF's economic ideology). We find no evidence that geopolitical alignment shapes the tone of staff appraisals. Neither a country's alignment with the United States nor its alignment with the European Union is systematically related to the tone of corresponding IMF Article IV staff appraisals. This holds for four sentiment measures, ranging from word lists to language model-based scores, and across a wide range of robustness checks. Macroeconomic fundamentals, by contrast, are reflected in appraisal tone, and our empirical design has the power to detect alignment effects of a similar size. With regard to economic ideology, a standard economic freedom index is associated with a more favourable tone for most sentiment measures. Since the index also captures the quality of institutions, this circumstance need not reflect an ideological preference.

1 Introduction

The International Monetary Fund (IMF) regularly surveys its member countries' economic policies through annual Article IV consultations. In effect, these represent an economic checkup conducted for nearly all members every year. Article IV consultation reports are far from being a pure formality. IMF surveillance disclosures move financial markets (Fratzscher and Reynaud, 2011). Their effect on sovereign debt prices is much larger in emerging market economies than in high-income economies, and investors value the IMF as a source of credible information (Breen and Doak, 2023). Since credit rating agencies explicitly reference IMF Article IV assessments in their sovereign rating methodologies, the same effect extends to the cost and availability of sovereign financing. This also matters for central banks, as sovereign yields anchor the domestic risk-free curve and the transmission of monetary policy. Precisely because its verdicts carry such consequences, the IMF's own framework requires that surveillance be evenhanded (IMF Strategy, Policy, and Review Department, 2012), not only in the substance of policy advice but also in the tone, candour and framing of staff assessments (IMF, 2016).

Whether the IMF is biased in practice has attracted sustained scholarly attention. Powerful member states, above all the United States, have an incentive to use the IMF's resources to reward the countries whose support they value. Most of the related literature finds that the reward goes to geopolitically aligned countries. Alignment with the United States, measured through United Nations General Assembly (UNGA) voting proximity, systematically improves treatment across multiple IMF outputs (see Steinwand and Stone,

[1] The authors are grateful to Fabio Rumler (OeNB) and an anonymous referee for helpful comments and valuable suggestions, to Márton Kardos (Aarhus University) for data support and to participants of an internal discussion seminar at WU (Vienna University of Economics and Business) and of the Austrian Statistical Days 2026 organised by the Austrian Statistical Society (ÖSG) in Vienna for helpful discussions.

2008, for an early review of related literature, or Momani, 2004, for a case study of IMF-Egypt loan negotiations documenting US intervention). In addition, countries that hold a temporary seat on the UN Security Council are more likely to receive IMF programmes and face fewer programme conditions (Dreher et al., 2009; Dreher et al., 2015).[2] More recent work traces shareholder influence on the IMF's formal governance. On the IMF Executive Board, the five largest shareholders tend to act as a collective principal, with the United States as the focal actor. Their interventions track their bilateral trade and aid ties with the country under discussion, while their UNGA voting affinity is not a robust predictor (Forster et al., 2025). In a similar vein, Copelovitch and Powers (2021) caution that evidence of, specifically, US dominance is sensitive to how alignment is measured, as the conventional UNGA measure may partly capture shared preferences among several large shareholders.[3] To date, however, this literature has focused almost exclusively on lending decisions, conditionality design (see e.g. Clark, 2022) and forecasts. Whether the same alignment effect extends to the tone of surveillance, the channel through which the IMF communicates its assessment to markets and the public every year, remains untested.

In this paper, we ask whether the IMF's surveillance is evenhanded. More specifically, we test whether staff appraisals, the evaluative summaries at the close of each Article IV consultation, are more favourable for some countries than for others, conditional on the countries' macroeconomic fundamentals. We consider two distinct forms of potential bias. The first is *geopolitical alignment* with the IMF's major shareholders. We focus on the United States, which holds by far the largest IMF voting share, and on the European Union, which by convention holds the position of IMF Managing Director. The second is adherence to the IMF's own *economic ideology*, as summarised in the Washington Consensus of liberalisation, deregulation and privatisation. The international financial institutions diffused this policy paradigm through the conditions attached to their lending, and its core elements have persisted in their lending practice (Babb, 2013). Section 4.1 discusses how both forms of bias could reach the tone of staff appraisals.

Against this background, we test these arguments using a corpus of Article IV staff appraisals covering 164 countries from 2001 to 2024.[4] We measure appraisal tone with four sentiment measures. They range from simple dictionary-based measures (Loughran-McDonald (LM), see Loughran and McDonald, 2011, and Syuzhet, see Jockers, 2015) to a contextualised but seed-phrase-anchored embedding approach (concept vector projection, see Lyngbaek et al., 2025) and a purely transformer-based approach (FinBERT, see Araci, 2019). Geopolitical alignment is measured as the ideal point distance (IPD) to the United States and to the European Union, derived from UNGA voting (Bailey et al., 2017).[5]

Our main results can be summarised as follows. First, we find no evidence that geopolitical alignment shapes the IMF's appraisal tone. Conditional on country and year fixed effects and macroeconomic fun-

[2] Other documented forms of bias, which are not analysed in this study, include IMF growth forecasts being tilted towards optimism when the IMF's lending exposure to a country is high (Aldenhoff, 2007) and surveillance assessments that read more favourably for countries with large outstanding IMF loans (Fratzscher and Reynaud, 2011). Parallel patterns have been documented at the World Bank, where US alignment predicts fewer and softer policy conditions (Clark and Dolan, 2021; Andersen et al., 2006) or more favourable debt sustainability ratings (Lang and Presbitero, 2018).

[3] Similar caution applies to the choice of outcome measure itself. At the World Bank, replacing conventional *counts* of loan conditions with a text-based measure of condition *stringency* eliminates the standard finding that powerful principals secure softer conditionality (Winter et al., 2025).

[4] A related master's thesis by one of the co-authors (Götze, 2026), using a similar design but a shortened dataset, reports a null result for the alignment-tone relationship.

[5] Kirsamer and Riha (2025) apply a compound geopolitical index, combining the Bailey et al. (2017) ideal point distance with two further political indices. They identify three distinct geopolitical blocs, namely a US-aligned cluster, a Chinese-aligned cluster and a South American/Southern African cluster.

damentals, neither the distance of the appraised country to the United States nor its distance to the EU is significantly related to appraisal tone. This holds for all four sentiment measures and across a wide range of robustness checks, including alternative reference points for the EU such as the Managing Director's home country. Second, this null result is informative and does not merely reflect a lack of statistical power. Since alignment varies mostly across countries and only little within them (section 4.2), a fixed effects design has limited variation to work with. Still, our specification picks up the effects of macroeconomic fundamentals on tone, and it has the power to detect an alignment effect of similar size. Moreover, a comparison of persistently more aligned with less aligned countries, relative to the United States or the EU, does not show favourable treatment of the more aligned countries, either. Third, turning to economic ideology, a standard economic freedom index is associated with a more favourable appraisal tone for three of the four sentiment measures. Since the index also covers the rule of law and sound monetary policy, part of this association may reflect stronger institutions rather than ideology.

The remainder of this paper is structured as follows. Section 2 describes the data. Section 3 presents the sentiment measurement approach and descriptive evidence. Section 4 discusses the channels through which bias could arise, introduces the empirical model we employ and reports the regression results, organised around the two forms of potential bias: geopolitical alignment and economic ideology. Section 5 concludes.

2 Data

In this section, we describe the data, our main focus variable sentiment, which is derived based on text data, as well as the macroeconomic control variables.

2.1 Corpus

Our dependent variable, sentiment, is extracted from Article IV staff appraisals. For that purpose, we started manually downloading Article IV consultations[6]. This extraction yielded considerable data gaps, especially for the years 2014 and 2015. We hence complemented the manually collected sample with a supplementary web-scraping routine. The IMF's legacy archive hosts country reports under a predictable URL pattern (https://www.imf.org/external/pubs/ft/scr/{year}/cr{YY}{NNN}.pdf). We queried this pattern systematically across report numbers for each target year, verified that the returned file was indeed a PDF and retained only those whose first pages matched an Article IV consultation header. The resulting reports were then matched to a country and year and checked against our manually collected sample to discard duplicates. This procedure added 484 additional country-year staff appraisals to the corpus, concentrated in 2001–2006 and, in particular, 2012–2015, closing most of the earlier gap.[7]

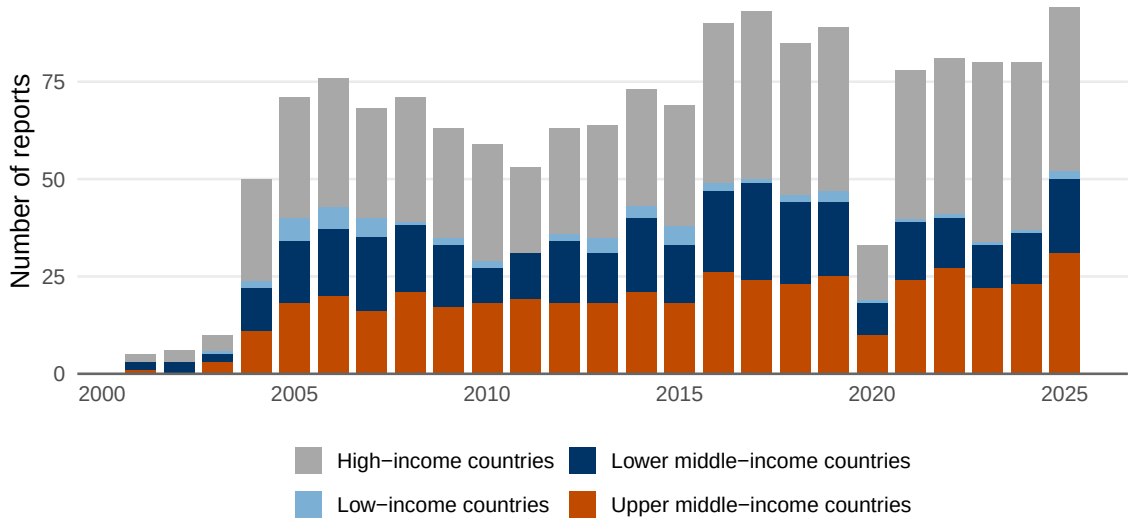
Consultations that were combined with a programme request or review were removed by hand, so that the corpus is restricted to *surveillance documents*. This yielded a final corpus of 1,604 Article IV consultation reports published by the IMF between 2001 and 2025 and covering 167 countries. The estimation samples are smaller once lagged macroeconomic controls are merged and the noisy 2025 ideal point estimates are

[6] The website we retrieved the articles from is the IMF's Publications search page (imf.org, country reports filtered by type Article IV).

[7] The archive's legacy URL scheme only covers older report vintages. An equivalent attempt for 2020 returned no results, since more recent reports are hosted under a different scheme not amenable to systematic enumeration.

excluded (see below). Chart 1 shows the distribution of reports over time and income classification taken from the World Bank.

Chart 1
Number of IMF Article IV reports over time



Note: Report counts by year and World Bank income classification (Atlas method).
Source: IMF eLibrary; authors' calculations.

The chart shows that time coverage of reports is low at the beginning of the sample and then again in 2020 due to the COVID-19 pandemic. These dynamics partly reflect changes in the IMF's transparency framework over the sample period: Publication moved from a largely voluntary norm towards a clearer default expectation with explicit timelines for release following Executive Board discussion (IMF Strategy, Policy, and Review Department, 2010; IMF, 2013). The IMF also documents delays for a subset of member countries in completing Article IV consultations (IMF Strategy, Policy, and Review Department, 2025). The sharp increase of published reports that started in 2004 was triggered by the decision to publish reports by default (Mihalyi and Mate, 2018).

2.2 Staff appraisal sections

From the reports, we extract a specific section, the so-called staff appraisal. This section – typically the final substantive section before the statistical annex – contains the IMF staff's integrated evaluative assessment of the member economy's macroeconomic situation and policies. It is extracted from the full PDF reports using a combination of regular-expression matching on section headers and manual validation of boundary detection.

The staff appraisal is preferable to the full consultation report as the text unit for two reasons. First, it contains the evaluative content: The descriptive sections of Article IV reports – background, economic outlook, sectoral analysis – are factual summaries, not assessments. Second, it is more consistent in length and structure across countries and years, reducing noise from variation in report scope.

Box 1 shows the staff appraisal for Panama from March 2023 to give a concrete example. For illustrative purposes, we will later also demonstrate how sentiment is estimated using this example text.

Box 1: Example of an IMF staff appraisal – Panama, March 2023

64. Post-pandemic policies should focus on rebuilding buffers and ensuring the convergence of Panamanian income levels with those in advanced countries continues.

65. The government's plans to further reduce the overall NPFS deficit to 3 percent in 2023 and 1½ percent of GDP in 2025 are appropriate. To sustainably reduce the fiscal deficit while preserving social spending and creating room for more education spending, tax revenue will need to increase. Improving tax and customs collection efficiency and broadening the tax base — by reducing exemptions, deductions, and tax expenditures — will be key.

66. Exiting the FATF grey list remains a priority. The authorities need to expediently address the remaining deficiencies in Panama's AML/CFT regulatory framework identified by the FATF.

67. With no lender of last resort and deposit insurance, it is imperative that the banking system remains well-capitalized and liquid. Current capital-adequacy and liquidity indicators in the banking system are well above regulatory minima, and continued intensive supervision and monitoring and expanding the macroprudential policy toolkit will help mitigate future asset quality and liquidity risks. Fintech companies should be subject to AML/CFT regulation and supervision. The ongoing FSAP will further inform the performance and vulnerabilities in Panama's financial system.

68. Going forward, it is likely that construction will not provide the same boost to growth as it has in the past two decades. For convergence to continue, other sectors will need to take over, and bringing human capital and governance standards in line with those in advanced countries would help.

69. It is recommended that the next Article IV consultation takes place on the standard 12-month cycle.

Source: IMF Article IV Consultation Report, Panama, March 2023 (staff appraisal, paragraphs 64–69).

The appraisal presented in box 1 is the shortest in our sample. Across the estimation sample, staff appraisals have a mean length of 819 words, ranging from 267 to 2,505 words.

2.3 Control variables: ideal point distance and macroenvironment

Our main focal variable is political alignment, which we measure using ideal point distance (IPD) estimates from Bailey et al. (2017). In a nutshell, IPD summarises each country's voting behaviour at the UNGA – covering all member states from 1946 to the present – as a single position on a common scale, the same way a series of yes/no survey answers can be boiled down to one overall score. Countries that vote similarly end up close together on this scale, whereas countries that vote very differently end up far apart. IPD is then simply the distance between a country's position and that of a reference country. In the main analysis, we focus on IPD to the United States, but also include distance to the European Union. In both cases, higher values indicate greater geopolitical distance.

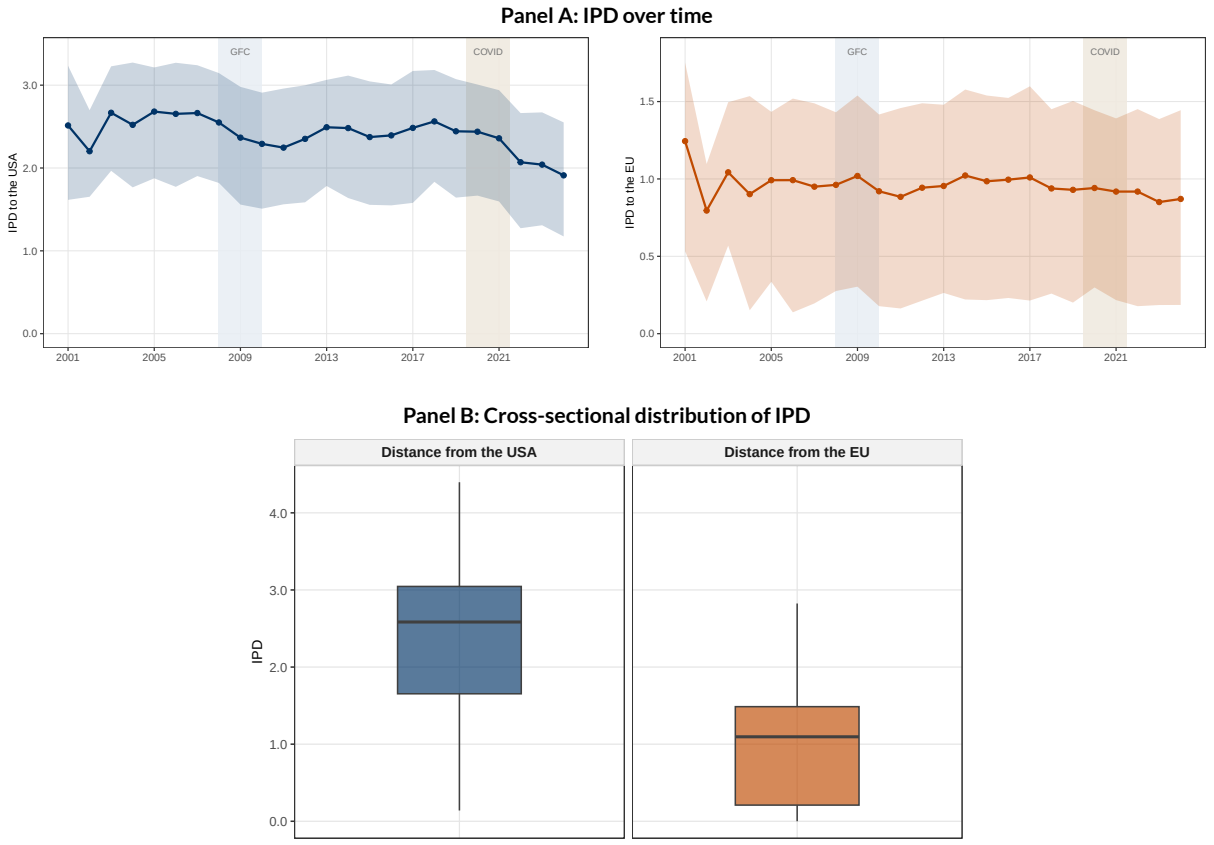
The United States is a UN member state with its own position on this scale, so distance to it is straightforward. The European Union is not: It does not cast its own vote at the UNGA, so it has no ideal point of its own. We therefore construct a proxy EU position as the simple average of member states' individual

ideal points. Rather than letting membership track the EU's actual, changing composition over the sample – which would conflate the alignment signal with a second, unrelated source of variation (*how many*, and *which*, countries enter the average each year) – we fix the member set to the 14 countries that belonged to the EU for the entire 2001–2024 sample window with no accession or departure: Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Denmark, Ireland, Greece, Portugal, Spain, Austria, Finland and Sweden.[8] IPD_EU is then the distance between a country's own position and this fixed membership average, constructed the same way as the US distance. Section 4 additionally reports three alternatives, namely (1) a version of the same fixed member set weighted by GDP at purchasing power parity (PPP), (2) the distance to the Managing Director's home country and (3) the distance to the average of the four non-US G5 members (United Kingdom, Germany, France, Japan).

Chart 2 shows the cross-country average of IPD over time to the USA and the EU (top panel), together with the distribution of IPD values across countries and years (bottom panel).

Chart 2

Ideal point distance to the USA and the EU: evolution over time and cross-country distribution



Note: Panel A: cross-country mean (solid line) and interquartile range (shaded area) of ideal point distance (IPD) to the USA (left) and the EU (right) by year; shaded vertical bars indicate the global financial crisis (2008–2010) and the COVID-19 pandemic (2020–2021). Panel B: distribution of IPD to the USA (left) and the EU (right) across pooled country-year observations; boxes show the interquartile range, whiskers extend to 1.5 times the interquartile range.
Source: Bailey et al. (2017); authors' calculations.

The chart shows that IPD is rather stable across time. For the USA, there is a slight downward trend visible, whereas IPD to the EU stays remarkably sideways. That said, there is a significant jump in IPD to the

[8] In the related literature, alignment is often measured relative to individual members of the G5, the IMF's five largest shareholders (see, e.g. Forster et al., 2025). Our choice of an EU reference point follows from the channel we test, the Managing Director, who by convention is European.

USA in 2025 – not shown in the chart – which can be attributed to the increasingly two-dimensional behaviour of UNGA voting.[9] We therefore restrict the regression sample to publication years through 2024.

Comparing the two reference countries, the distance to the USA is higher than the distance to the EU, but the dispersion is also higher for USA distances – consistent with the EU position itself being an average across many countries' ideal points, and therefore mechanically less extreme than any single reference country's position. Although the IPD is widely used in the literature, two limitations are worth noting. First, UNGA voting may reflect strategic signalling or issue-specific coalitions that do not fully capture a country's broader geopolitical stance. Second, the UNGA agenda is concentrated in particular issue areas – Israel-Palestine, human rights, arms control – that may not capture the economic dimensions of US–country relationships most relevant for IMF surveillance.[10] These limitations will be partially addressed by the econometric model, a panel fixed effects design, which identifies the alignment effect from within-country changes over time and reduces concerns about cross-sectional confounding.

Three main macroeconomic control variables are included in the baseline specifications, drawn from the World Bank's World Development Indicators (WDI). GDP growth controls for the economic performance channel: In years of stronger growth, IMF staff have less basis for critical assessments. Inflation controls for macroeconomic stress: Higher inflation creates conditions for prescriptive language on fiscal and monetary policy adjustment. The current account balance as a share of GDP controls for external vulnerability: Countries with stronger external positions receive, on average, more favourable assessments (this describes the estimated sample-wide relationship, not a claim about any single country). All three variables are entered at a one-year lag relative to the report publication year. Table 1 summarises every variable used across the main specification, the extensions and the robustness checks (section 4), together with its source, coverage and summary statistics.

[9] A small number of countries – most notably China and Russia – now appear closer to the US and EU on the standard one-dimensional scale largely because of overlapping positions on a second, distinct axis (e.g. Ukraine-related and human rights votes) rather than genuine convergence in overall alignment. Collapsing this onto one dimension distorts 2025 distances in particular.

[10] The IPD is the most widely used metric to measure geopolitical alignment. Alternative approaches consider joint memberships in treaties (Lupu, 2013; Lupu, 2016) and joint memberships in international organisations (Voeten, 2021).

Table 1

Data sources and summary statistics – all explanatory and control variables used in the main, extensions and robustness specifications

Variable	Source	N	Mean	SD	Minimum	Maximum
Political distance						
IPD (US distance)	Bailey et al. (2017)	1,510	2.40	0.82	0.14	4.40
IPD (EU distance, fixed 14-member set)	Bailey et al. (2017)	1,510	0.95	0.68	0.00	2.82
IPD (MD home country, robustness)	Bailey et al. (2017)	1,510	1.16	0.75	0.00	3.32
IPD (G4 average, robustness)	Bailey et al. (2017)	1,510	1.12	0.73	0.00	3.07
Macroeconomic controls (lagged)						
GDP growth (annual %)	World Bank WDI (NY.GDP.MKTP.KD.ZG)	1,510	3.46	5.34	−32.91	86.83
Inflation (annual %)	World Bank WDI (FP.CPI.TOTL.ZG)	1,501	5.00	14.83	−8.79	521.15
Current account balance (% of GDP)	World Bank WDI (BN.CAB.XOKA.GD.ZS)	1,494	−0.34	15.88	−60.72	314.91
Economic ideology (extensions)						
Foreign exchange regime (1 = peg)	Ilzetzki et al. (2019), Coarse sheet	1,468	0.47	0.50	0	1
Capital account restriction (1 = dual market)	Ilzetzki et al. (2019), Unified sheet	1,482	0.09	0.29	0	1
Gross government debt (% of GDP)	IMF World Economic Outlook	1,501	53.59	40.95	0.00	543.40
EFW index (0–10)	Fraser Institute EFW	1,382	6.93	0.97	3.18	8.77
Robustness checks						
UNSC member (temporary, 1 = yes)	UN Security Council elected member records (authors' compilation)	1,510	0.06	0.24	0	1
Sentiment measures (outcomes)						
LM polarity (per word)	Loughran and McDonald (2011)	1,510	−0.015	0.023	−0.110	0.057
Syuzhet score (per word)	Jockers (2015)	1,510	0.031	0.009	0.003	0.074
CVP economic sentiment	Lyngbaek et al. (2025), own seed phrases	1,510	0.070	0.046	−0.078	0.223
FinBERT net score	ProsusAI/finbert (Araci, 2019)	1,510	0.53	0.26	−0.40	0.93

Note: SD = standard deviation. WDI = World Development Indicators. N = country-year observations with valid data on the canonical estimation dataset (N = 1,510; full corpus 1,604 appraisals, 167 countries, 2001–2025, restricted to publication years through 2024). Macroeconomic controls are matched at a one-year lag (t−1). The ideal point distance (IPD) measures, UN Security Council (UNSC) membership and FinBERT are defined at the publication year t. MD home country = the country of the Managing Director in office for most of the year. G4 = United Kingdom, Germany, France, Japan. LM = Loughran and McDonald. CVP = concept vector projection. Foreign exchange regime and capital account restriction from Ilzetzki et al. (2019) (classification through 2016, held at each country's last observed value, 2017–2024). EFW = Fraser Institute Economic Freedom of the World index (0–10, higher = more market-oriented; discussed in section 4.3). FinBERT net score = probability positive minus probability negative from a pretrained financial text sentiment classifier (ProsusAI/finbert), one of the four sentiment measures (section 3). Group means of all four sentiment measures are reported in table 2. The regression sample after merging lagged macroeconomic controls comprises 1,485 observations in the full (table 3) specification. Extension and robustness variables are estimated on their own available samples (see tables 5–7).
Source: Authors' compilation based on the sources listed.

3 Sentiment estimation

In this section, we describe the dependent variable, the tone of the IMF's Article IV staff appraisals.

3.1 Measuring sentiment

We measure tone with four sentiment measures. They can be ordered by how much the researcher defines in advance what counts as favourable tone. At one end are two word lists, which fix the tone of each word before any text is scored. Next is concept vector projection (CVP), which reads words in context but anchors tone to example sentences that we formulate. At the other end is FinBERT, a language model whose notion of tone is learned from financial text rather than set by us.

The first word list is the dictionary of Loughran and McDonald (2011), calibrated on financial texts. We define Loughran and McDonald (LM) polarity as the difference between positive and negative word counts relative to total words:

$$\text{LM polarity} = \frac{\text{Positive words} - \text{negative words}}{\text{Total words}} \tag{1}$$

Its negative word list captures the evaluative language of financial commentary, such as *deteriorate*, *vulnerable* and *unsustainable*. The second word list is the general-purpose Syuzhet lexicon (Jockers, 2015), which assigns valence scores to individual words based on the NRC emotion lexicon (Mohammad and Turney, 2013). We divide the Syuzhet score of an appraisal by its number of words.

The two word lists have complementary weaknesses. General-purpose lexica misclassify much of the vocabulary of economic policy. In corporate annual reports, almost three-quarters of the words that a general dictionary classifies as negative are not negative in a financial context (Loughran and McDonald, 2011). The same holds in our corpus, where Syuzhet scores *tax*, *government* and *regulatory* as negative (table A3 in the annex). The LM dictionary avoids this problem, but it was built for corporate filings rather than for assessments of economic policy. Moreover, both word lists ignore the context in which a word appears.

Our third measure takes this context into account while keeping tone anchored in the language of staff appraisals. It is based on CVP, proposed by Lyngbaek et al. (2025) and applied to central bank communication by Feldkircher et al. (2026). CVP defines tone as a direction in the embedding space of a transformer model.[11] For that purpose, we formulate about 30 example sentences in the style of a staff appraisal, half of them describing favourable and half unfavourable economic conditions (table A2 in the annex). The concept vector is the difference between the average embeddings of the two groups. We split each appraisal into short text windows, project each window onto the concept vector and average the window scores, weighted by window length. The examples cover a wide range of topics, from structural reforms over growth, inflation, the external position, public debt and the financial sector to governance and institutions. CVP thus measures how favourably an appraisal describes a country's economic and policy situation in a broad sense.

Our fourth measure, FinBERT, leaves the definition of tone to a language model. It fine-tunes Google's BERT (Devlin et al., 2018) on financial text (Araci, 2019) and, like CVP, reads each word in its context. We use the ProsusAI/finbert checkpoint, which classifies text as positive, negative or neutral, and take the probability of a positive minus that of a negative classification.[12]

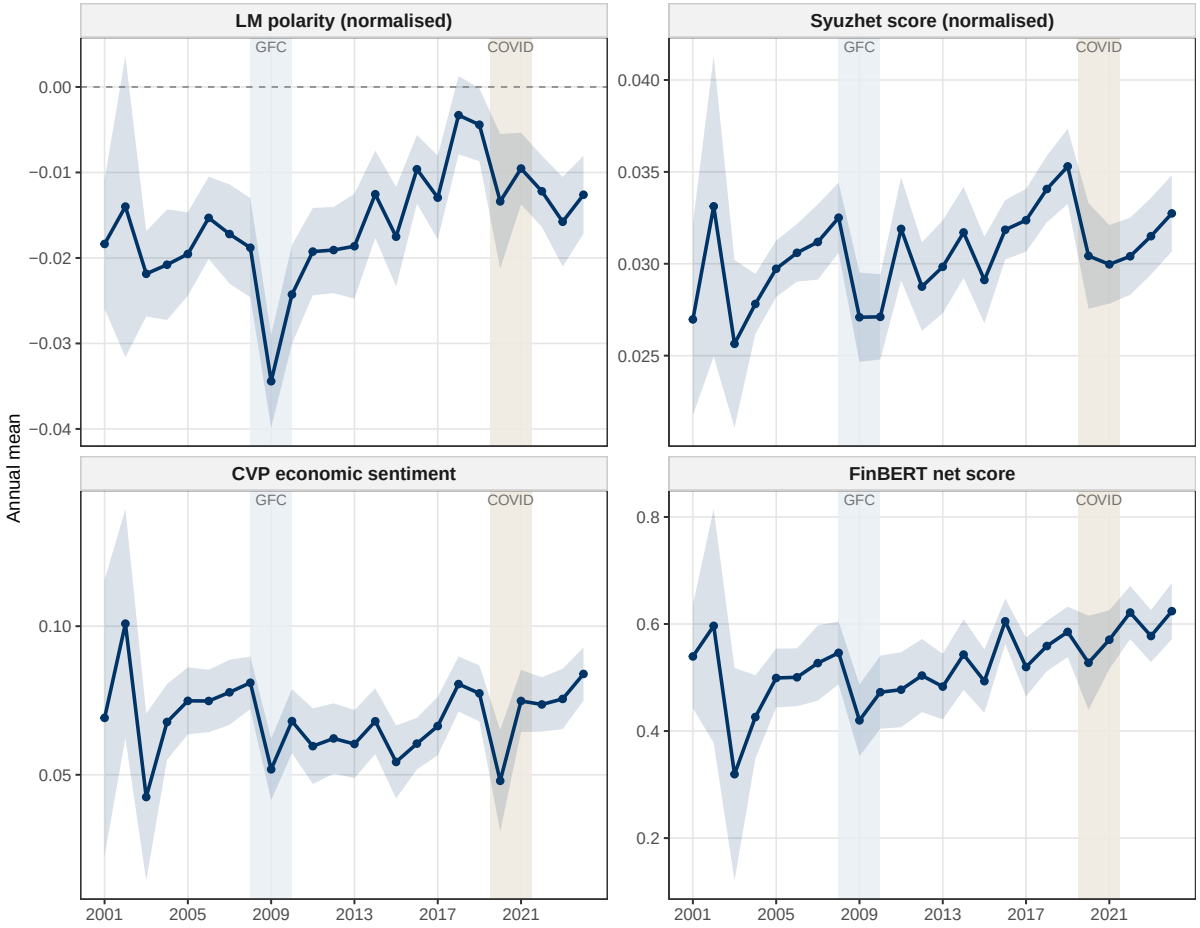
[11] An embedding is a numerical vector that represents a token, i.e. a mapping of text into a vector space produced by a transformer model. Such models differ in how much surrounding text, the context length, they can take into account. Following Feldkircher et al. (2026), we use the nvidia/llama-embed-nemotron-8b model. Its context length of 32,768 tokens is large enough to embed an entire appraisal at once, so that every token can attend to all others.

[12] We do not use generative large language models (LLMs) to score tone. Prompting an LLM to rate a text on a numerical scale adds a choice that the researcher has to make in advance, namely the rating scale itself. Jo et al. (2026) show that changing a single element

None of the four measures is superior on all counts. The word lists are transparent and easy to replicate, but they ignore context. The CVP measure is anchored in seed phrases and can hence be tailored to the text and application at hand. By the same token, this can be viewed as a weakness, since the researcher has to choose the example sentences a priori. FinBERT is the measure least controlled by the researcher, and the financial text it was trained on may differ from the language of policy assessments. We therefore show all measures side by side instead of picking one. That said, the profession may nowadays regard context-based measures as preferable. To bring all measures to the same scale, we z-standardise them. Coefficients are thus expressed in standard deviations of tone and are comparable across measures.

Chart 3 shows the four measures over time. Ignoring the years before 2004, for which only a handful of reports are available, all four fall during the global financial crisis and again in 2020, at the onset of the COVID-19 pandemic. That the measures track these episodes lends credibility to them as measures of appraisal tone.

Chart 3
Mean sentiment by year, four measures, 2001–2024



Note: Annual means of each measure in its own units. FinBERT net score = probability positive − probability negative. LM = Loughran and McDonald. CVP = concept vector projection. Shaded bands: global financial crisis (2008–2010) and COVID-19 pandemic (2020–2021). Shaded ribbons: 95% confidence intervals.
Source: IMF eLibrary Article IV consultation reports; authors' calculations.

The chart also shows that the measures move together but differ in their patterns. Their pairwise cor-

of this scale, i.e. its format, range or sign, shifts scores by about one tenth of the scale's range on average.

relations in the estimation sample range from 0.27 to 0.47.

Table 2 shows the sample composition by income group and the average tone of each group in standard deviations. High-income countries account for almost half of the reports.

Table 2
Sample composition and mean sentiment by income group

Income group	Countries	Reports	Share	LM	Syuzhet	CVP	FinBERT
High income	57	702	46.5%	−0.01	−0.03	0.24	0.07
Upper middle income	46	423	28.0%	0.02	−0.04	−0.13	−0.12
Lower middle income	48	333	22.1%	0.04	0.14	−0.21	0.09
Low income	16	52	3.4%	−0.26	−0.23	−0.82	−0.52
Total	164	1,510	100.0%	0.00	0.00	0.00	0.00

Note: LM = Loughran and McDonald. CVP = concept vector projection. Sentiment columns in standard deviations. Each report's score is standardised over all reports in the table and then averaged by group. Income groups follow the World Bank Atlas method in the report year, so a country may appear in more than one group. The country total counts distinct countries. Estimation sample, 1,510 reports, 2001–2024.
Source: Authors' calculations.

On all four measures, appraisals of low-income countries are the least favourable, by between a quarter and four-fifths of a standard deviation. Differences among the other groups are smaller, and their order depends on the measure. Crucially, lower-income countries are also more politically distant from the United States on average, and the correlation between income classification and IPD is substantial.

4 Regression results

This section tests two distinct forms of potential bias in Article IV staff appraisal tone, geopolitical alignment and economic ideology. We first discuss the channels through which each could reach appraisal tone and then turn to the empirical model.

4.1 Channels and hypotheses

Politically powerful countries seek influence in international institutions and thereby shape their institutional and economic stance. For the IMF, there are at least two such candidate members. The first is the United States, which holds by far the largest voting share (around 16%). The second is the European Union, which by convention always holds the position of Managing Director. We call this a *governance channel*.[13] It could manifest itself in two ways. First, geopolitical proximity to either of these members could influence staff appraisal tone. Second, the economic ideology largely favoured by these same stakeholders could shape it. For lending, the two are documented to be linked. The market liberalisation preferences that the largest shareholders express on the Executive Board are associated with more market-liberalising conditions in subsequent programmes (Forster et al., 2025).

The governance channel operates in practice through *institutional pressure* on IMF staff. The IMF's own 2011 internal evaluation found that staff "felt pressure to align their conclusions with IMF policies and positions" (Roy and Ramos, 2012). Such pressure need not be exercised directly. A former IMF economist interviewed by Clark (2022) describes the United States as rarely hands-on in designing programmes. In-

[13] A different channel, which we do not test, runs through individual staff rather than member-country interests. Mission chiefs bring their own professional and ideological priors to policy content (Amgalan et al., 2025), although this bias loses its explanatory power when the borrower is a close US ally (Lang et al., 2025). Staff who encounter an issue such as climate change on a country assignment also carry that attention with them when they are rotated or promoted and can thereby shift the IMF's position without directives from management or member states (Clark and Zucker, 2024).

stead, staff are simply expected not to clash with the US Treasury. In the case of the EU, the Managing Director does not draft surveillance reports directly but sets the IMF's overall priorities and analytical guidance. Staff have incentives to internalise this guidance. Hence, EU proximity could shape appraisal tone if a European Managing Director's priorities tilted the IMF's stance towards EU-aligned countries.

The staff appraisal, an evaluative narrative rather than a mechanical calculation, leaves staff considerable room for discretion.[14] Our empirical exercise documents whether the predicted alignment-tone and ideology-tone relationships hold under a demanding specification, rather than adjudicating precisely how the governance channel's influence is transmitted.

4.2 Empirical model

Our empirical workhorse is a two-way fixed effect (TWFE) panel, estimated separately for each sentiment measure $S \in \{\text{LM, Syuzhet, CVP, FinBERT}\}$ and separately for each of the two alignment measures:

$$\text{M1:} \quad S_{it} = \alpha_i + \lambda_t + \beta\, \text{IPD_US}_{it} + \gamma' \mathbf{X}_{i,t-1} + \varepsilon_{it} \tag{2a}$$

$$\text{M2:} \quad S_{it} = \alpha_i + \lambda_t + \beta\, \text{IPD_EU}_{it} + \gamma' \mathbf{X}_{i,t-1} + \varepsilon_{it} \tag{2b}$$

where i indexes countries and t years, α_i and λ_t denote country and year fixed effects, and ε_{it} is an idiosyncratic error clustered at the country level. Our focal variables are the ideal point distances to the United States (IPD_US) to test for the US voting channel and the ideal point distance to the European Union (IPD_EU) testing for the EU Managing Director nationality channel. A negative coefficient would indicate that countries more aligned with the USA or the European Union receive a more positive tone in the staff appraisal. We enter the two distances in separate models because they are highly correlated. EU members and the United States often vote alike in the UNGA, so a joint model would split one common alignment signal into two offsetting coefficients. This also follows the convention in the literature of measuring alignment with individual major shareholders. We report the joint model as a robustness check (section 4.5).

We include three macroeconomic controls, gathered in the matrix $\mathbf{X}_{i,t-1}$, namely GDP growth, inflation and the current account balance. All of them enter with a one-year lag, which accounts for the time between the consultation and the publication of the Article IV report. As a side benefit, this lag also mitigates endogeneity concerns.[15] Note two features of our design. First, the panel is unbalanced. Second, because country fixed effects absorb all time-invariant country characteristics and year fixed effects absorb common global shocks, our specification is a stringent test of whether the alignment variables can explain any remaining variation in the sentiment scores. Moreover, year fixed effects absorb common movements over time, such as the slight downward drift in average distance to the United States (chart 2) or shifts in the average tone of all reports in a given year. A specification without year effects would conflate these with the within-country relationship of interest. In later robustness exercises, we will extend the main set-up given

[14] Such discretion in nominally technocratic assessments is what allows political interests to bias outcomes (Lang and Presbitero, 2018). In lending, management and the Executive Board tend to rubber-stamp staff proposals and step in mainly when the vital interests of major shareholders are at stake (Reinsberg et al., 2022).

[15] When all country-level controls enter with a two-year lag instead, the alignment coefficients remain insignificant for all four measures, in the main specification as well as in the extensions of section 4.3. The economic freedom index also remains positive and significant for most sentiment measures. Results are available upon request.

in equation 2 by adding further variables of interest.

Table 3 reports the results for the four sentiment measures. All outcomes are z-standardised on the estimation sample. Each coefficient thus gives the change in tone, in standard deviations, associated with a one-unit change in the regressor. IPD is measured in ideal point units and the macroeconomic controls in percentage points.

Table 3
Main results – two-way fixed effects (country + year); standard errors clustered by country

Variable	LM	Syuzhet	CVP	FinBERT
Panel A: M1, distance to the United States				
IPD_US	−0.076	−0.105	−0.052	0.104
	(0.161)	(0.191)	(0.150)	(0.186)
Inflation	−0.0045***	−0.0034**	−0.0036	−0.0021†
	(0.0008)	(0.0011)	(0.0022)	(0.0012)
GDP growth	0.0077	0.0091	0.0225*	0.0066
	(0.0064)	(0.0066)	(0.0113)	(0.0083)
Current account balance/GDP	0.0072**	0.0039*	0.0076**	0.0021
	(0.0022)	(0.0018)	(0.0028)	(0.0021)
R^2 / within R^2	0.285 / 0.015	0.311 / 0.008	0.513 / 0.037	0.268 / 0.003
Panel B: M2, distance to the European Union				
IPD_EU	0.044	−0.181	−0.045	−0.024
	(0.174)	(0.204)	(0.164)	(0.170)
Inflation	−0.0043***	−0.0036**	−0.0036	−0.0023*
	(0.0008)	(0.0011)	(0.0022)	(0.0012)
GDP growth	0.0080	0.0088	0.0225*	0.0063
	(0.0063)	(0.0066)	(0.0113)	(0.0083)
Current account balance/GDP	0.0073***	0.0040*	0.0077**	0.0020
	(0.0022)	(0.0017)	(0.0028)	(0.0021)
R^2 / within R^2	0.285 / 0.014	0.311 / 0.009	0.513 / 0.037	0.268 / 0.003
Country + year fixed effects	Yes	Yes	Yes	Yes
N	1,485	1,485	1,485	1,485

Note: LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. Dependent variables are z-standardised, so coefficients are in standard deviations of tone per unit of the regressor (IPD in ideal point units, macroeconomic controls in percentage points). Macroeconomic controls lagged one year. All columns use the identical sample. IPD_EU is the fixed 14-member EU average (section 2.3). Standard errors clustered by country in parentheses. ***p < 0.001, **p < 0.01, *p < 0.05, †p < 0.1.
Source: Authors' calculations.

Our results show the following. First, the IPD_US coefficient is small and statistically insignificant for all four sentiment measures, and its sign differs across them (panel A of table 3). The same holds for IPD_EU (panel B). We therefore find no evidence that within-country movements in UNGA voting alignment with either the United States or the European Union translate into systematically different appraisal tone.

Macroeconomic fundamentals, by contrast, are reflected in appraisal tone. Higher lagged inflation is associated with more negative appraisals for LM, Syuzhet and FinBERT, and the coefficient is significant for all three. A stronger current account balance is associated with more favourable tone for LM, Syuzhet and CVP. Lagged GDP growth enters with the expected positive sign for all measures and is significant for CVP. Each measure thus responds to at least one macroeconomic fundamental in the expected direction, which lends credibility to the scores as measures of appraisal tone. It also shows that the specification has the power to detect systematic tone differences where they exist.

How large an alignment effect can these estimates rule out? The answer depends on how much IPD actually moves within countries. Political alignment is a slow-moving, largely country-specific characteristic. More specifically, 92% of the variance of IPD_US is between countries. Once country and year fixed effects

are removed, its standard deviation shrinks from 0.81 to 0.15 ideal point units. Nevertheless, the within-country variation is not concentrated in a few countries. Only about 6% of countries have a single report or a practically constant US distance, and the median country's US distance varies by about 0.6 ideal point units across its reports. A one-unit change in IPD, about seven times the typical within-country movement, is hence far outside the variation that identifies the two-way fixed effects estimates. Following Mummolo and Peterson (2018), we instead interpret magnitudes in terms of a *typical identifying move*, i.e. a change in a regressor by one standard deviation of its variation left after removing country and year fixed effects. Table 4 reports the implied change in tone for each regressor of M1 and M2, together with its 95% confidence interval and the minimum effect that the design detects with 80% power.

Table 4
Effect sizes – change in tone (SD) for a typical within-country move

Variable	Typical move	LM	Syuzhet	CVP	FinBERT
		Effect [95% CI], MDE in second row			
IPD_US	0.15	−0.01 [−0.06, 0.04] 0.07	−0.02 [−0.07, 0.04] 0.08	−0.01 [−0.05, 0.04] 0.06	0.02 [−0.04, 0.07] 0.08
IPD_EU	0.14	0.01 [−0.04, 0.05] 0.07	−0.02 [−0.08, 0.03] 0.08	−0.01 [−0.05, 0.04] 0.06	0.00 [−0.05, 0.04] 0.07
Inflation	12.7	−0.06 [−0.08, −0.04] 0.03	−0.04 [−0.07, −0.01] 0.04	−0.05 [−0.10, 0.01] 0.08	−0.03 [−0.06, 0.00] 0.04
GDP growth	4.1	0.03 [−0.02, 0.08] 0.07	0.04 [−0.02, 0.09] 0.07	0.09 [0.00, 0.18] 0.13	0.03 [−0.04, 0.09] 0.09
Current account balance/GDP	10.5	0.08 [0.03, 0.12] 0.06	0.04 [0.00, 0.08] 0.05	0.08 [0.02, 0.14] 0.08	0.02 [−0.02, 0.07] 0.06

Note: SD = standard deviation. CI = confidence interval. LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. IPD_US and macroeconomic controls from M1, IPD_EU from M2 (table 3). Typical move = standard deviation of the regressor after removing country and year fixed effects (Mummolo and Peterson, 2018). Effect = coefficient × typical move, in standard deviations of tone. MDE (second row) = minimum detectable effect at 5% significance and 80% power, $(t_{0.975} + t_{0.80}) \times$ standard error × typical move.
Source: Authors' calculations.

For all four measures, a typical within-country move towards the United States changes tone by at most 0.02 standard deviations. The confidence intervals exclude favourable effects for closer countries larger than about 0.07 standard deviations. By contrast, the same specification picks up tone changes of a similar size for typical moves in the macroeconomic fundamentals. For example, a typical increase in the current account balance improves LM and CVP tone by about 0.08 standard deviations. Moreover, the design would detect an alignment effect of 0.06 to 0.08 standard deviations with 80% power. In other words, the data could have revealed an alignment effect of the same order as the effect of macroeconomic fundamentals, but they do not. The same holds for the distance to the European Union.

4.3 Economic ideology: exchange rate policy, capital account openness, debt and economic freedom

Beyond geopolitical alignment, we test a second, distinct form of potential bias: whether the IMF's own historical association with a free-market, Washington Consensus-style economic ideology – of which capital account liberalisation was a canonical element (Kaya and Reay, 2019) – is itself reflected in appraisal tone. Under this framing, countries operating open, unified foreign exchange markets, carrying less government debt and scoring higher on a standard economic freedom index might be expected to receive more favourable appraisal language. We test four dimensions of a country's macroeconomic and institutional environment that proxy for adherence to this ideology, adding them one at a time to the main specification

(table 3, M1) – an exchange rate regime dummy, a capital account restriction dummy, government debt (% of GDP) and the Fraser Institute Economic Freedom of the World (EFW) index – each estimated on its own available sample under the identical two-way fixed effects specification and control set, and then jointly in a fifth, full specification with all four together. The results are shown in table 5.

Table 5

Extensions – exchange rate regime, capital account openness, government debt and economic freedom (two-way fixed effects)

Variable	LM	Syuzhet	CVP	FinBERT
+ Foreign exchange regime dummy				
IPD_US	−0.114	−0.131	−0.101	0.077
	(0.164)	(0.196)	(0.155)	(0.201)
Foreign exchange regime (1 = peg)	0.121	0.027	−0.095	0.099
	(0.144)	(0.224)	(0.166)	(0.156)
Observations	1,443	1,443	1,443	1,443
+ Capital account restriction				
IPD_US	−0.084	−0.080	−0.025	0.127
	(0.158)	(0.189)	(0.148)	(0.186)
Capital account restriction	−0.284†	−0.159	−0.245	−0.392†
	(0.152)	(0.212)	(0.230)	(0.213)
Observations	1,457	1,457	1,457	1,457
+ Government debt ratio				
IPD_US	−0.124	−0.121	−0.108	0.072
	(0.159)	(0.196)	(0.150)	(0.190)
Gross government debt (% of GDP)	−0.0005	−0.0002	−0.0027**	−0.0025***
	(0.0011)	(0.0013)	(0.0009)	(0.0007)
Observations	1,476	1,476	1,476	1,476
+ Economic freedom (EFW)				
IPD_US	0.025	−0.045	−0.132	−0.007
	(0.198)	(0.189)	(0.160)	(0.159)
EFW index (0–10)	0.073	0.229*	0.414**	0.380*
	(0.141)	(0.108)	(0.126)	(0.167)
Observations	1,370	1,370	1,370	1,370
Full specification (all four together)				
IPD_US	−0.106	−0.045	−0.214	−0.046
	(0.209)	(0.205)	(0.161)	(0.169)
Foreign exchange regime (1 = peg)	0.051	−0.006	−0.139	0.056
	(0.140)	(0.228)	(0.158)	(0.153)
Capital account restriction	−0.208	−0.194	−0.288*	−0.511*
	(0.223)	(0.197)	(0.134)	(0.223)
Gross government debt (% of GDP)	−0.0002	0.0002	−0.0023***	−0.0021*
	(0.0009)	(0.0011)	(0.0007)	(0.0008)
EFW index (0–10)	0.115	0.312*	0.461***	0.350*
	(0.139)	(0.121)	(0.127)	(0.165)
Observations	1,302	1,302	1,302	1,302

Note: LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. Dependent variables z-standardised as in table 3. Each block adds the named variable(s) to M1 (table 3). Macroeconomic controls and fixed effects included but not shown. Foreign exchange regime and capital account restriction (1 = dual or parallel market) from Ilzetzki et al. (2019), debt from IMF World Economic Outlook, EFW = Fraser Institute Economic Freedom of the World index. Standard errors clustered by country in parentheses. ***p < 0.001, **p < 0.01, *p < 0.05, †p < 0.1.
Source: Authors' calculations.

First, we examine whether the decision of a country regarding its exchange rate regime has an impact on staff appraisals' tone. For that purpose, we utilise the de facto measure of Ilzetzki et al. (2019), by including a dummy variable that is one for a peg and zero otherwise.[16] The coefficient is small and statistically

[16] Ilzetzki et al. (2019) sort country-years into six coarse regime categories. The dummy equals one for the first category (pegs, currency boards and narrow bands) and zero for crawling pegs, managed floats and free floats (categories 2–4). The remaining two categories, freely falling exchange rates and dual markets without parallel-market data, are set to missing, since they do not reflect a deliberate regime choice. These country-years drop out of the foreign exchange regime block and of the full specification. The

insignificant for all four sentiment measures, so appraisal tone does not appear to be shaped by the choice of exchange rate regime.

The next measure we consider is the capital account restriction proxy, also constructed from the Ilzetzki et al. (2019) data. This dummy variable equals one when a country operates a dual or parallel exchange market rather than a single unified one, which is reflective of having some sort of capital controls in place. The coefficient is negative for all four measures. In the block that adds only this variable to M1, it is significant at the 10% level for LM and FinBERT. In the full specification with all four extension variables, it is significant for CVP and FinBERT. This points to somewhat less favourable appraisal language in years in which a country maintains a dual or parallel exchange market. As Ilzetzki et al. (2019) note explicitly, this indicator is a conservative, asymmetric proxy for capital account restrictions: A sustained gap between the official and parallel-market rate cannot persist where capital moves freely, so a dual market implies binding restrictions, but a unified market does not imply their absence – countries can maintain extensive capital controls while still keeping a single official rate. The variable should therefore be read as identifying restrictions severe enough to produce a black-market premium, a lower bound on capital account restriction rather than a comprehensive measure. The mixed result may also reflect a shift in the IMF's own position. The IMF now acknowledges the potential utility of capital controls (Kentikelenis et al., 2016). Hence, maintaining them need not carry a penalty in appraisal tone.

Next, we examine the effect of government debt. For the two word lists, the coefficient is close to zero and insignificant, both in the block that adds only debt and in the full specification. For FinBERT and CVP, by contrast, higher debt is associated with significantly less favourable tone in both specifications. Debt ratio coverage is close to complete (1,476 of the 1,485 observations of the main specification), so the results cannot be attributed to a restricted subsample. A higher debt ratio is a weaker economic fundamental rather than an ideological marker, and staff appraisals naturally discuss it. The two measures that read words in context thus pick up this discussion, while the word lists do not.

Economic freedom indices are the standard way to operationalise free-market policies in the literature on the international financial institutions (Heinzel et al., 2021). The fourth variable, the Fraser Institute Economic Freedom of the World (EFW) index – an established composite measure of a country's overall market orientation, motivated here as a "free-market" proxy by Graafland and Verbruggen (2022) – is positive and significant for Syuzhet, FinBERT and CVP, both in the block that adds only the EFW index and in the full specification. It is insignificant only for LM. Taken at face value, more economically free countries receive more favourable tone within-country, holding political alignment and macroeconomic conditions fixed. One caveat applies. The EFW index also covers the rule of law and sound monetary policy (Graafland and Verbruggen, 2022). A more favourable appraisal of a country with stronger institutions need not reflect ideology.

Across all five extension specifications, the coefficient on US distance remains statistically insignificant for all four sentiment measures, so the absence of an alignment effect in table 3 does not hinge on omitting these country characteristics.

Summing up, the economic freedom index is the one candidate for ideological bias that is related to ap-

classification ends in 2016. For later years, we carry forward each country's 2016 classification. The same applies to the capital account restriction dummy below.

praisal tone across measures. It is associated with more favourable tone for three of the four measures, although part of this association may reflect stronger institutions rather than ideology. Capital account restrictions go along with less favourable tone for some measures and specifications. Foreign exchange regime choice is not related to tone, and government debt matters only for the two context-based measures. In the next section, we examine cross-country variation in a complementary exercise.

4.4 Between-country alignment

Two-way fixed effects identify the alignment-tone relationship purely from within-country variation over time – year-to-year movements in a country's UNGA voting position. Because country fixed effects absorb all time-invariant differences between countries, TWFE is silent on a related but distinct question: Do persistently more US-aligned countries receive systematically different tone than persistently more distant ones, on average, over the whole sample? We address it with a separate cross-sectional regression of each country's time-averaged sentiment score on its time-averaged covariates (one observation per country, $N = 164$), with heteroskedasticity-robust standard errors.[17] This exercise is informative because IPD_US varies overwhelmingly across countries rather than within them (92% of its total variance is between-country, compared with 96% for IPD_EU, 77–92% for the exchange rate variables above and 18–55% for the macroeconomic controls). A cross-sectional specification uses exactly this variation, at the cost of being unable to net out unobserved, time-invariant country characteristics beyond what is explicitly controlled for. Table 6 reports the results.

[17] In a robustness exercise, we also tried to decompose the within and between variation using the model of Allison (2009) and implemented in the R package panelr. However, that model was not capable of including year fixed effects, so no meaningful decomposition was possible.

Table 6
Between-country alignment – cross-sectional OLS on country means

Variable	LM	Syuzhet	CVP	FinBERT
Panel A: M1, distance to the United States				
IPD_US	0.046	0.116†	−0.083	0.044
	(0.068)	(0.065)	(0.092)	(0.064)
Inflation	0.0028	−0.0047	−0.0124†	0.0004
	(0.0041)	(0.0062)	(0.0072)	(0.0059)
GDP growth	−0.0082	−0.0167	−0.0313	−0.0577†
	(0.0329)	(0.0287)	(0.0391)	(0.0335)
Current account balance/GDP	0.0094†	0.0120**	0.0144	0.0101†
	(0.0052)	(0.0044)	(0.0093)	(0.0052)
Constant	−0.101	−0.173	0.275	0.061
	(0.117)	(0.123)	(0.207)	(0.123)
R^2	0.048	0.074	0.111	0.108
Panel B: M2, distance to the European Union				
IPD_EU	0.074	0.194*	−0.041	0.079
	(0.076)	(0.077)	(0.094)	(0.079)
Inflation	0.0024	−0.0060	−0.0135†	−0.0002
	(0.0041)	(0.0064)	(0.0075)	(0.0060)
GDP growth	−0.0085	−0.0177	−0.0342	−0.0583†
	(0.0322)	(0.0277)	(0.0377)	(0.0327)
Current account balance/GDP	0.0094†	0.0121**	0.0144	0.0102†
	(0.0053)	(0.0044)	(0.0093)	(0.0052)
Constant	−0.060	−0.077	0.126	0.093
	(0.087)	(0.087)	(0.124)	(0.097)
R^2	0.050	0.090	0.106	0.112
Observations	164	164	164	164

Note: LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. Dependent variables are z-standardised on the document-level estimation sample before averaging, so coefficients are in standard deviations of tone. Cross-sectional ordinary least squares (OLS), one observation per country (each variable averaged over the country's appraisals in the estimation sample). Heteroskedasticity-robust standard errors in parentheses. ***p < 0.001, **p < 0.01, *p < 0.05, †p < 0.1.
Source: Authors' calculations.

For US alignment, the between-country coefficient is insignificant for LM, FinBERT and CVP. For Syuzhet, it is positive and significant at the 10% level. Countries that are on average more distant from the United States thus receive, if anything, slightly more favourable Syuzhet tone. This is the opposite of what US favouritism would imply. The same pattern holds for EU alignment. The Syuzhet coefficient is positive and significant, again the opposite of an EU favouritism reading, while the other three measures show no association. Across countries, the macroeconomic controls explain little of the differences in tone. The current account balance is positively associated with tone for LM, Syuzhet and FinBERT. Inflation is negatively associated with CVP tone, and GDP growth enters with a negative sign for FinBERT. The two exchange rate variables from section 4.3, entered one at a time into the same cross-sectional regression, show no consistent pattern either. The foreign exchange regime dummy is negative and significant for CVP only, and the capital account restriction proxy is negative and significant at the 10% level for LM only. In the CVP specification with the foreign exchange regime dummy, the US coefficient turns negative and significant at the 10% level. This result does not carry over to the other measures or specifications.

The two exercises use very different parts of the variation in alignment, and their precision differs accordingly. The cross-section exploits the large differences between countries. The standard deviation of country-average IPD_US is about 0.76 ideal point units, five times the 0.15 left within countries after removing the fixed effects. However, with 164 observations and without country fixed effects, the cross-section pins down the coefficient much less precisely. For the four measures, the 95% confidence intervals

for the tone difference associated with a one standard deviation difference in country-average US distance are about 0.2 to 0.3 standard deviations wide. The within-country estimates (table 4) rule out much smaller effects, but only for the comparatively small movements in alignment that countries actually undergo over our sample. Summing up, neither dimension of the data shows a consistent alignment effect.

4.5 Additional robustness checks

We probe whether the absence of an alignment effect is an artefact of particular modelling choices, examining the IPD_US coefficient from five further angles: an inference check, an omitted-confounder control, two sample period restrictions and the joint model with IPD_EU. Table 7 summarises these checks. The same checks for IPD_EU leave its coefficient insignificant for all four measures (results available upon request). We then turn to alternative reference points for the EU channel.

Table 7
Robustness checks – US alignment coefficient

Check	LM	Syuzhet	CVP	FinBERT	N
Two-way clustering (country + year)	−0.076	−0.105	−0.052	0.104	1,485
UN Security Council control	−0.080	−0.113	−0.053	0.096	1,485
Sample restricted to 2004–2024	−0.080	−0.082	−0.041	0.129	1,464
Sample restricted to 2001–2023	−0.147	−0.129	−0.032	0.116	1,405
Joint with IPD_EU	−0.238	0.043	−0.047	0.267	1,485

Note: LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. Each row reports the IPD_US coefficient from a separate two-way fixed effects specification based on M1 (table 3); standard errors clustered by country unless stated otherwise. Dependent variables z-standardised. UNSC = temporary UN Security Council membership dummy. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.
Source: Authors' calculations.

First, inference: Clustering standard errors two-way by country and year leaves the US coefficient insignificant for all four measures, so the null result does not depend on the choice of clustering dimension. Second, we include a dummy variable that reflects temporary membership of the UN Security Council (UNSC). Dreher et al. (2015) show that countries that are members of the Council get preferential IMF treatment in terms of fewer loan conditions. For our data, the UNSC dummy itself is insignificant for all four measures, and the US coefficient is essentially unchanged and remains insignificant. We thus find no indication that temporary Security Council members receive a more benevolent tone in their appraisals.

Third, we re-estimate the main specification restricted to publication years 2004–2024, dropping 2001–2003 (5, 6 and 10 documents, respectively) to address the concern that these sparsest early years could disproportionately affect the result. The US coefficient is essentially unchanged and insignificant for all four measures. Fourth, we drop the most recent publication year, 2024, instead. The US coefficient remains insignificant for all four measures.

Fifth, we enter both distances jointly. For LM, Syuzhet and FinBERT, the US and EU coefficients take opposite signs, and for LM and FinBERT the US coefficient becomes larger in absolute value. All coefficients remain insignificant. This offsetting pattern reflects how closely the two distances move together rather than two separate channels.

Finally, we replace the EU-14 average in M2 by three alternative reference points, each entered on its own. The first weights the same 14 members by their share of group PPP-GDP instead of equally. The second is the distance to the home country of the Managing Director in office for most of the year (Germany

2001–2003, Spain 2004–2007, France 2008–2019, Bulgaria 2020–2024). This is the most direct operationalisation of the channel that motivates the EU term. The third follows the convention in the literature of measuring alignment with the IMF's five largest shareholders (G5). Together with IPD_US, the distance to the average ideal point of the other four (United Kingdom, Germany, France, Japan) covers the G5. Table 8 reports the three alternatives next to the EU-14 average.

Table 8
EU channel – alternative reference points for the second distance measure

Distance to reference point	LM	Syuzhet	CVP	FinBERT
EU-14 average (M2)	0.044	−0.181	−0.045	−0.024
	(0.174)	(0.204)	(0.164)	(0.170)
EU-14, PPP-GDP-weighted	−0.028	−0.251	−0.128	−0.039
	(0.180)	(0.209)	(0.166)	(0.164)
MD home country	−0.086	−0.161	0.019	0.105
	(0.156)	(0.176)	(0.157)	(0.168)
G4	−0.026	−0.256	−0.152	−0.055
	(0.182)	(0.208)	(0.161)	(0.155)
Observations	1,485	1,485	1,485	1,485

Note: LM = Loughran and McDonald. CVP = concept vector projection. PPP = purchasing power parity. Each row is a separate two-way fixed effects specification as M2 in table 3, with the EU-14 average replaced as indicated. MD = home country of the Managing Director in office for most of the year. G4 = average ideal point of the United Kingdom, Germany, France and Japan. Dependent variables z-standardised. Standard errors clustered by country in parentheses. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$.
Source: Authors' calculations.

None of the alternatives changes the picture. The distance to each reference point is insignificant for all four sentiment measures.

Overall, none of the five checks produces a US alignment coefficient that is significant at conventional levels, and none of the four reference points for the EU channel is related to tone.

5 Conclusions

This paper tests two distinct forms of potential bias in IMF Article IV staff appraisals (geopolitical alignment and economic ideology), drawing on a corpus of around 1,500 consultations covering 164 countries from 2001 to 2024. Our main finding is a null result. We find no evidence that geopolitical alignment shapes the tone of IMF staff appraisals.

For alignment with the United States, our two-way fixed effects estimates are insignificant for all four sentiment measures employed, i.e. two word lists, a transformer-based classifier and an embedding-based measure. Moreover, none of our robustness checks yields a coefficient that is significant at conventional levels. Since alignment varies mostly across countries and only little within countries, a null result in a fixed effects design could simply reflect a lack of statistical power. However, our estimates do not support this interpretation. The same specification picks up the effects of macroeconomic fundamentals on tone, and it has the power to detect an alignment effect of similar size. Exploiting the much larger variation across countries, a cross-sectional comparison of persistently more aligned with less aligned countries, relative to the United States, does not show favourable treatment of the more aligned countries, either. We also test for alignment with the European Union, as the IMF Managing Director is, by convention, European. Here, too, we find no association. This holds whether alignment is measured relative to the average of 14 EU member countries, relative to the IMF Managing Director's home country or relative to the other G5

members.

For economic ideology, our second test, the picture is different. A standard economic freedom index is associated with a more favourable tone for three of the four measures, both alone and jointly with the other proxies. However, the index also covers the rule of law and sound monetary policy. A more favourable appraisal of economically free countries may thus reflect stronger institutions rather than an ideological preference. A country's de facto exchange rate regime is found to be unrelated to tone, while capital account restrictions and higher public debt correlate with a less favourable tone for some measures.

How do these findings fit a literature that documents political influence on IMF lending? One interpretation is that the governance channel operates where the principals pay attention. Member states may monitor surveillance less closely than costlier, binding instruments such as loan conditionality (Clark and Zucker, 2024). Evidence from the IMF Executive Board is consistent with this claim. The interventions of the largest shareholders track their bilateral economic ties rather than their voting affinity in the UN General Assembly (UNGA). In lending, too, UNGA voting alignment with the United States does not predict the number of programme conditions (Reinsberg et al., 2022). In addition, these ties are less robust predictors in Article IV discussions than in discussions of lending programmes (Forster et al., 2025). A country's ideal point distance to the United States is also unrelated to whether its Article IV reports address climate change. Instead, whether they do depends on the career paths of the IMF staff involved (Clark and Zucker, 2024). Moreover, Article IV surveillance covers almost all member countries every year and is subject to an explicit evenhandedness framework, which makes differential treatment more visible and more costly. Officials in member countries, in turn, appear to judge the IMF's impartiality by how well its policy advice fits their own economic policy paradigm rather than by their alignment with the United States (Heinzel et al., 2021). Our results are consistent with these findings, but they cannot discriminate between them.

What do our results imply for those who use Article IV reports? Central banks, finance ministries and market participants read IMF staff appraisals as an assessment of a country's economic situation. Our results suggest that they can read the tone of these appraisals without worrying that changes in a country's geopolitical alignment might cause a shift in tone. For all four measures, any more favourable treatment of US-aligned countries would have to be of, at most, the same order as the change in tone that accompanies a typical within-country change in macroeconomic fundamentals. With regard to this dimension, Article IV surveillance appears evenhanded.

That said, two limitations constrain our interpretation. First, ideal point distance captures geopolitical alignment only coarsely through UNGA voting and may miss the bilateral economic and diplomatic relationships that are most relevant for IMF staff behaviour. Second, all four sentiment measures capture aggregate tone rather than the specific policy vocabulary through which editorial pressures may operate. The absence of an effect on aggregate tone therefore does not rule out political influence on the content of staff appraisals. A natural next step is hence to examine whether political alignment shapes word choice in staff appraisals, conditional on the macroeconomic environment.

References

Aldenhoff, F.-O. 2007. Are economic forecasts of the International Monetary Fund politically biased? A public choice analysis. Review of International Organizations 2. 239–260. DOI: 10.1007/s11558-006-9010-x.

Allison, P. D. 2009. Fixed Effects Regression Models. Thousand Oaks, CA: SAGE Publications.

Amgalan, J., J. Cohen-Setton and A. Korinek. 2025. Reading the Fund: A Systematic Analysis of IMF Fiscal Advice Using Large Language Models. IEO Background Paper BP/25-01/05. 3 November. Independent Evaluation Office of the International Monetary Fund.

Andersen, T. B., H. Hansen and T. Markussen. 2006. US politics and World Bank IDA lending. Journal of Development Studies 42(5). 772–794. DOI: 10.1080/00220380600741946.

Araci, D. 2019. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. arXiv preprint arXiv:1908.10063.

Babb, S. 2013. The Washington Consensus as transnational policy paradigm: Its origins, trajectory and likely successor. Review of International Political Economy 20(2). 268–297. DOI: 10.1080/09692290.2011.640435.

Bailey, M., A. Strezhnev and E. Voeten. 2017. Estimating dynamic state preferences from United Nations voting data. Journal of Conflict Resolution 61(2). 430–456.

Breen, M. and E. Doak. 2023. The IMF as a global monitor: surveillance, information, and financial markets. Review of International Political Economy. DOI: 10.1080/09692290.2021.2004441.

Clark, R. 2022. Bargain down or shop around? Outside options and IMF conditionality. The Journal of Politics 84(3). 1791–1805. DOI: 10.1086/719269.

Clark, R. and L. R. Dolan. 2021. Pleasing the principal: US influence in World Bank policymaking. American Journal of Political Science 65(1). 36–51.

Clark, R. and N. Zucker. 2024. Climate cascades: IOs and the prioritization of climate action. American Journal of Political Science 68(4). 1299–1314. DOI: 10.1111/ajps.12793.

Copelovitch, M. and R. Powers. 2021. Do We Really Know What We Think We Know About the Politics of IMF Lending? Measuring and Reassessing US Influence in Global Financial Governance. Paper prepared for the 2021 International Political Economy Society Conference. University of Colorado Boulder. 19 October.

Devlin, J., M.-W. Chang, K. Lee and K. Toutanova. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

Dreher, A., J.-E. Sturm and J. R. Vreeland. 2015. Politics and IMF Conditionality. Journal of Conflict Resolution 59(1). 120–148. DOI: 10.1177/0022002713499723.

Dreher, A., J.-E. Sturm and J. R. Vreeland. 2009. Global horse trading: IMF loans for votes in the United Nations Security Council. European Economic Review 53(7). 742–757. DOI: 10.1016/j.euroecorev.2009.03.002.

Feldkircher, M., M. Kardos and K. L. Nielbo. 2026. From Tone to Trajectory: Continuous Sentiment and the Shape of Monetary Policy Communication. Working Paper 281. Oesterreichische Nationalbank.

Forster, T., D. Honig and A. Kentikelenis. 2025. Formal governance matters: when, how, and why states act on the IMF Executive Board. Review of International Political Economy 32(3). 818–846. DOI: 10.1080/09692290.2024.2441136.

Fratzscher, M. and J. Reynaud. 2011. IMF surveillance and financial markets – a political economy analysis. European Journal of Political Economy 27(3). 405–422.

Götze, T. 2026. Evenhanded Surveillance? Geopolitical Alignment and Sentiment in IMF Article IV-Consultations. MA thesis. Vienna School of International Studies (Master of Digital International Affairs).

Graafland, J. and H. Verbruggen. 2022. Free-Market, Perfect Market and Welfare State Perspectives on "Good" Markets: an Empirical Test. Applied Research in Quality of Life 17. 1113–1136. DOI: 10.1007/s11482-021-09946-2.

Heinzel, M., J. Richter, P.-O. Busch, H. Feil, J. Herold and A. Liese. 2021. Birds of a feather? The determinants of impartiality perceptions of the IMF and the World Bank. Review of International Political Economy 28(5). 1249–1273. DOI: 10.1080/09692290.2020.1749711.

Ilzetzki, E., C. M. Reinhart and K. S. Rogoff. 2019. Exchange Arrangements Entering the Twenty-First Century: Which Anchor will Hold? Quarterly Journal of Economics 134(2). 599–646. DOI: 10.1093/qje/qjy033.

IMF. 2013. Guidance Note on the Fund's Transparency Policy. Policy Paper 2013.96. International Monetary Fund. DOI: 10.5089/9781498341042.007.

IMF. 2016. Evenhandedness of Fund Surveillance – Principles and Mechanism for Addressing Concerns. Policy Paper 16.9. International Monetary Fund. DOI: 10.5089/9781498346047.007.

IMF Strategy, Policy, and Review Department. 2010. Guidance Note on the Fund's Transparency Policy. Policy Paper. Approved by Reza Moghadam. 12 March; updated 23 December. International Monetary Fund.

IMF Strategy, Policy, and Review Department. 2012. Guidance Note for Surveillance Under Article IV Consultations. Policy Paper. Approved by Siddharth Tiwari. 10 October. International Monetary Fund.

IMF Strategy, Policy, and Review Department. 2025. List of IMF Member Countries with Delays in Completion of Article IV Consultations or Mandatory Financial Stability Assessments over 18 Months. Policy Paper 2025.026. International Monetary Fund. DOI: 10.5089/9798229019415.007.

Jo, W., G. Kim and M. S. Hee. 2026. Change the Number, Change the Answer: Assessing Rating Scale Sensitivity in LLM-Based Text Scoring. arXiv preprint arXiv:2610.22187.

Jockers, M. L. 2015. Revealing Sentiment and Plot Arcs with the Syuzhet Package. http://www.matthewjockers.net/2015/02/02/syuzhet/. Blog post. matthewjockers.net. 2 February.

Kaya, A. and M. Reay. 2019. How did the Washington consensus move within the IMF? Fragmented change from the 1980s to the aftermath of the 2008 crisis. Review of International Political Economy 26(3). 384–409. DOI: 10.1080/09692290.2018.1511447.

Kentikelenis, A. E., T. H. Stubbs and L. P. King. 2016. IMF conditionality and development policy space, 1985–2014. Review of International Political Economy 23(4). 543–582. DOI: 10.1080/09692290.2016.1174953.

Kirsamer, N. and K. Riha. 2025. Mapping geopolitical fragmentation: a bilateral index for economic analysis. OeNB Bulletin Q4/2025.

Lang, V. and A. F. Presbitero. 2018. Room for discretion? Biased decision-making in international financial institutions. Journal of Development Economics 130. 1–16.

Lang, V., L. Wellner and A. Kentikelenis. 2025. Biased bureaucrats and the policies of international organizations. American Journal of Political Science 69. 1486–1504. DOI: 10.1111/ajps.12921.

Loughran, T. and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Journal of Finance 66(1). 35–65.

Lupu, Y. 2013. The Informative Power of Treaty Commitment: Using the Spatial Model to Address Selection Effects. American Journal of Political Science 57(4). 912–925. DOI: 10.1111/ajps.12033.

Lupu, Y. 2016. Why Do States Join Some Universal Treaties but Not Others? An Analysis of Treaty Commitment Preferences. Journal of Conflict Resolution 60(7). 1219–1250. DOI: 10.1177/0022002714560344.

Lyngbaek, L., P. Feldkamp, Y. Bizzoni, K. L. Nielbo and K. Enevoldsen. 2025. Continuous Sentiment Scores for Literary and Multilingual Contexts. Anthology of Computers and the Humanities 3. Ed. by T. Arnold, M. Fantoli and R. Ros. 480–497. DOI: 10.63744/nVu1Zq5gRkuD.

Mihalyi, D. and A. Mate. 2018. Text-mining IMF country reports – an original dataset. SSRN Electronic Journal. DOI: 10.2139/ssrn.3268934.

Mohammad, S. M. and P. D. Turney. 2013. Crowdsourcing a word-emotion association lexicon. Computational Intelligence 29(3). 436–465.

Momani, B. 2004. American politicization of the International Monetary Fund. Review of International Political Economy 11(5). 880–904.

Mummolo, J. and E. Peterson. 2018. Improving the interpretation of fixed effects regression results. Political Science Research and Methods 6(4). 829–835. DOI: 10.1017/psrm.2017.44.

Reinsberg, B., T. Stubbs and A. Kentikelenis. 2022. Unimplementable by design? Understanding (non-)compliance with International Monetary Fund policy conditionality. Governance 35. 689–715. DOI: 10.1111/gove.12605.

Roy, R. and R. A. Ramos. 2012. IMF Article IV Reports: An Analysis of Policy Recommendations. Working Paper 86. Brasilia. February. https://hdl.handle.net/10419/71819. International Policy Centre for Inclusive Growth (IPC-IG).

Steinwand, M. C. and R. W. Stone. 2008. The International Monetary Fund: a review of the recent evidence. Review of International Organizations 3(2). 123–149.

Voeten, E. 2021. Ideology and International Institutions. Princeton, NJ: Princeton University Press.

Winter, J., B. Cormier, T. Kramarz and M. S. Manger. 2025. Reassessing World Bank conditionality: beyond count measures. Political Science Research and Methods. 1–21. DOI: 10.1017/psrm.2025.19.

Annex

A1 US alignment: coefficient consistency across specifications

Table A1

US alignment (IPD_US) across all specifications

Specification	LM	Syuzhet	CVP	FinBERT
TWFE main (M1)	−0.076	−0.105	−0.052	0.104
Between-country (cross-sectional)	0.046	0.116†	−0.083	0.044
TWFE + UNSC membership control	−0.080	−0.113	−0.053	0.096
TWFE + foreign exchange regime	−0.114	−0.131	−0.101	0.077
TWFE + capital account restriction	−0.084	−0.080	−0.025	0.127
TWFE + government debt	−0.124	−0.121	−0.108	0.072
TWFE + EFW index	0.025	−0.045	−0.132	−0.007
TWFE + all four extension variables	−0.106	−0.045	−0.214	−0.046

Note: LM = Loughran and McDonald. CVP = concept vector projection. IPD = ideal point distance. UNSC = UN Security Council. EFW = Fraser Institute Economic Freedom of the World index. Each row reports the IPD_US coefficient from a separate model. Dependent variables are z-standardised. TWFE (two-way fixed effects) rows: specifications with the lagged macroeconomic controls (M1 in table 3), standard errors clustered by country. Between-country row: cross-sectional ordinary least squares (OLS) on country means, heteroskedasticity-robust standard errors. ***p < 0.001, **p < 0.01, *p < 0.05, †p < 0.1.
Source: Authors' calculations.

A2 CVP example sentences

Table A2
Example sentences defining the CVP concept vector

Favourable	Unfavourable
Recent reforms are steps in the right direction and should be sustained.	Progress on structural reforms has fallen well short of what is needed.
The economy is expanding at a solid pace and the near-term outlook is favorable.	The economy is in recession and the near-term outlook is challenging.
Growth has been robust and broad-based, underpinned by strong domestic demand.	Growth has stalled and economic activity remains weak.
Macroeconomic fundamentals are sound and the economy has proved resilient to shocks.	A widening current account deficit and large macroeconomic imbalances are weighing on the outlook.
The external position is strong, supported by a current account surplus and ample reserves.	Public debt is on an unsustainable path and fiscal space has been exhausted.
Public finances are on a sound footing and public debt is low.	Inflation is high and rising, eroding households' purchasing power.
Inflation is low and stable and financial conditions are supportive of growth.	The external position has deteriorated sharply and international reserves are critically low.
Strong buffers and policy frameworks underpin the economy's resilience to shocks.	Risks to the outlook are firmly tilted to the downside.
The banking system is well capitalized, liquid, and profitable.	The economy remains highly vulnerable to external shocks and a sudden stop in capital inflows.
External vulnerabilities are limited and reserves provide an ample cushion.	Elevated public debt and large gross financing needs are a major source of vulnerability.
Debt is assessed as sustainable and the risk of debt distress is low.	The banking sector is fragile, with high non-performing loans and thin capital buffers.
Strong institutions and credible policy frameworks anchor macroeconomic stability.	International reserves have fallen well below adequate levels.
The central bank is independent, transparent, and has a strong track record.	Weak governance and pervasive corruption are holding back investment and growth.
Fiscal transparency is high and public financial management is sound.	Institutional capacity is limited and policy implementation has been uneven.
	Fiscal transparency is poor and large quasi-fiscal operations obscure the true fiscal position.

Note: The concept vector is the difference between the average embeddings of the favourable and the unfavourable sentences (section 3). One further favourable sentence ("Risks to the outlook are broadly balanced") was dropped after a leave-one-out check, as it was closer to the unfavourable than to the favourable pole.
Source: Authors' compilation.

A3 Illustrative word classification: LM and Syuzhet

Table A3

Illustrative sentences and sentiment classification – LM and Syuzhet

Overall score: net = −1 (pos = 3, neg = 4)		Overall score: 9.35	
LM positive	LM negative	Syuzhet positive	Syuzhet negative
boost	deficit (×2)	advanced (×2)	tax (×4)
efficiency	deficiencies	well (×2)	deficit (×2)
improving	imperative	adequacy	regulatory (×2)
		appropriate	deficiencies
		asset	government
		boost	pandemic
		continue	risks
		efficiency	
		ensuring	
		focus	
		forward	
		growth	
		improving	
		increase	
		performance	
		priority	
		provide	
		rebuilding	
		recommended	

Note: LM = Loughran and McDonald. Illustrative sentences drawn from the Article IV corpus. LM classification follows Loughran and McDonald (2011). Polarity word lists show the words from the LM dictionary identified in each sentence and their valence (±1).
Source: IMF eLibrary Article IV consultation reports; authors' calculations.

© Oesterreichische Nationalbank, 2026. All rights reserved.
Address: Otto-Wagner-Platz 3, 1090 Vienna
PO Box 61, 1011 Vienna, Austria
Website: www.oenb.at

May be reproduced for noncommercial, educational and scientific purposes provided that the source is acknowledged.

Opinions expressed by the authors do not necessarily reflect the official viewpoint of the Oesterreichische Nationalbank or the Eurosystem.

Data protection information www.oenb.at/en/dataprotection
ISSN 2960-5415 (online)