IEO

Independent Evaluation Office
of the International Monetary Fund

ONLINE APPENDIX

IEO Evaluation of the IMF Advice on Fiscal Policy: Reading the Fund: A Systematic Analysis of IMF Fiscal Advice Using Large Language Models

Contents Page

ABBREVIATIONS

AE	Advanced Economy
CDS	Credit Default Swap
EMMIE	Emerging Market and Middle-Income Economy
FB	Fiscal Balance
FE	Fixed Effect
GDP	Gross Domestic Product
GFN	Gross Financing Need
LIC	Low-Income Country
VE	Vulnerability Exercise
WEO	World Economic Outlook

I. Robustness to Alternative Measures of Cyclical Conditions and Fiscal Vulnerability

This section tests whether our main findings—regarding the countercyclicality of IMF fiscal advice and its sensitivity to fiscal risks—remain robust when alternative measures of key explanatory variables are used. We first consider alternative measures of cyclical conditions to test whether our results on the countercyclicality of IMF advice are robust. We then examine alternative measures of fiscal sustainability and vulnerability to assess the robustness of our findings on the sensitivity of IMF advice to fiscal risks. Finally, we examine whether the responsiveness of IMF advice differs systematically across countries facing different degrees of fiscal risk, drawing on the Fund's composite Vulnerability Exercise (VE) assessments.

A. Output Stabilization

(i) Alternative Measures of Cyclical Conditions

To test the robustness of our findings on the countercyclicality of IMF fiscal advice, we consider several alternative indicators of countries' cyclical positions beyond our baseline output gap measure. The baseline combines World Economic Outlook (WEO)—used whenever they were available at the time of each Article IV consultation—with a Hamilton-filter estimates that we compute from the corresponding vintage of the WEO real GDP series when no WEO output gap was reported. Because the real GDP series is available contemporaneously, the Hamilton-filter estimates are constructed using the vintage closest to the publication date of each Article IV report. This approach maximizes sample coverage and gives priority to the Fund's contemporaneous assessments. That being said, no single measure of the output gap is perfect. To assess the robustness of our results, we test alternative ways of capturing cyclical conditions. Unlike the baseline, which blends WEO and Hamilton-filter gaps for wider coverage, the alternative measures rely on a single method throughout, trading coverage for methodological consistency. The first alternative uses WEO output gaps only, relying exclusively on the contemporaneous staff estimates available at the time of each Article IV report. This approach preserves the judgmental information embedded in Fund assessments but results in a smaller sample, especially for low-income countries (LICs). The second uses the Hamilton-filter output gap. This ensures methodological comparability and remains feasible even in data-poor environments, though it is subject to the usual limitations of statistical filters. The third indicator is the deviation of real growth from its five-year historical average, as in the IMF's Vulnerability Exercise and Kilic Celik and others (2023). This backward-looking measure captures whether current real GDP growth is above or below its recent trend, capturing growth momentum rather than the degree of slack. The fourth indicator is the deviation of actual growth from the five-year-ahead WEO projection, defined as the difference between the current real GDP growth rate and the rate projected five years into the future. Because IMF forecasts generally assume output gaps close within the forecast horizon, this forward-looking measure serves as a practical proxy for potential growth, as also used in Kose and others (2023).

(ii) Main Results

Table AI.1 presents the regression results obtained when replacing the baseline output gap with each of the four alternative cyclical indicators. Across all specifications, the results consistently confirm that IMF fiscal advice is countercyclical.

Table AI.1. Robustness: Alternative Measures of Cyclical Conditions

Variables	(1)	(2)	(3)	(4)
Output gap (WEO only)	11.685***			
	(2.324)			
Output gap (Hamilton filter)		4.882***		
		(0.971)		
Growth deviation from 5-year average			11.442***	
			(2.413)	
Growth deviation from 5-year ahead WEO projection				9.433***
				(1.897)
FB gap	2.821***	1.880**	0.852	1.106
	(0.975)	(0.730)	(0.720)	(0.726)
Debt-to-GDP	1.232***	1.331***	1.114***	1.317***
	(0.210)	(0.179)	(0.179)	(0.182)
τ_1	0.437***	0.145	0.593***	0.475***
	(0.122)	(0.122)	(0.099)	(0.102)
τ_2	1.278***	1.037***	1.484***	1.373***
	(0.129)	(0.126)	(0.107)	(0.109)
N	1767	2734	2732	2606
Country FE	No	No	No	No
Year FE	No	No	No	No
Pseudo R^2 (McFadden)	0.409	0.136	0.142	0.171
Pseudo R^2 (Ugba & Gertheiss)	0.725	0.301	0.312	0.369
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

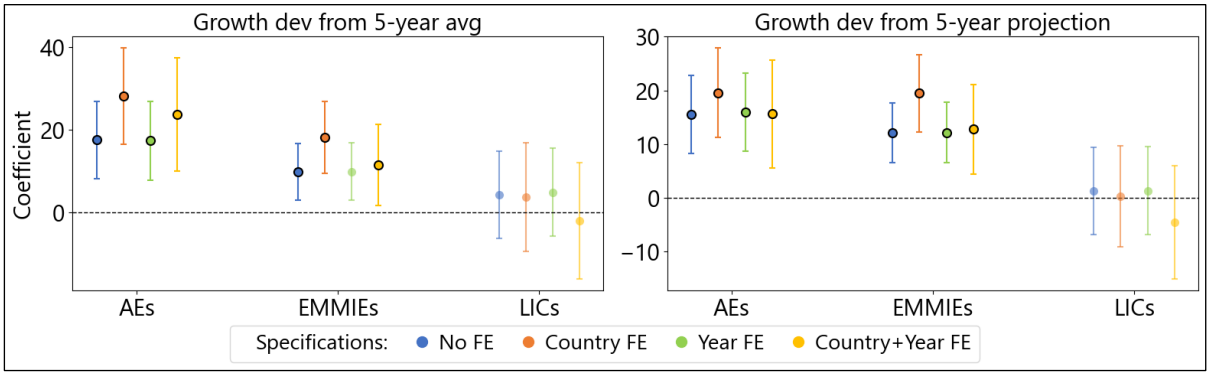
Using the WEO-only output gap (column 1) yields a large and statistically significant positive coefficient (11.69), indicating that as economic slack diminishes—that is, as output rises above potential—the Fund is more likely to recommend fiscal tightening. The Hamilton filter-based output gap (column 2) produces a smaller but still significant positive coefficient (4.88), supporting the same interpretation. The deviation of real GDP growth from its past five-year historical average (column 3) also shows a strong, positive, and statistically significant effect (11.44), and the deviation of real GDP growth from the five-year-ahead WEO projection (column 4) likewise yields a positive and statistically significant effect (9.43).

In all specifications, the fiscal balance gap and the debt-to-GDP ratio remain strong and statistically significant predictors of tighter fiscal advice, with coefficients generally ranging from 0.85 to 2.82 and from 0.88 to 1.33, respectively. These findings underscore that IMF fiscal recommendations reflect a combination of macroeconomic stabilization motives and fiscal sustainability considerations.

(iii) Results by Income Group

Building on the robustness exercises above, we next examine whether the countercyclical nature of IMF fiscal advice holds across income groups when using the alternative cyclical indicators. Figure AI.1 reports results obtained when the deviation of real growth from its five-year average and the deviation from the five-year-ahead WEO projection are interacted with income group dummies.

Figure AI.1. Robustness: Income Group Heterogeneity with Alternative Measures of Cyclical Conditions



Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: All regressions use contemporaneous data. Vertical lines indicate 95 percent confidence intervals. Detailed estimates corresponding to these results are reported in Tables AV.15. and AV.16. in Section V.

For AEs, the interaction terms are consistently large and statistically significant across specifications, indicating that fiscal advice leans strongly against the cycle regardless of the cyclical measure used. For EMMIEs, countercyclical recommendations emerge in several specifications, particularly when country fixed effects are included, suggesting a more conditional responsiveness to the cycle that may reflect heterogeneity in market access or institutional capacity. For LICs, the coefficients on the cyclical interaction terms are generally positive but not statistically significant, implying weaker or less systematic countercyclical advice. Overall, these results show that while IMF advice is broadly countercyclical across all cyclical indicators, its strength varies with countries' income levels and economic resilience.

B. Fiscal Sustainability

(i) Alternative Measures of Fiscal Sustainability

The previous set of robustness checks examined whether our main findings on the countercyclicality of IMF fiscal advice hold when using alternative measures of cyclical conditions. We now test whether these results remain stable when fiscal sustainability indicators are defined more broadly. Our baseline analysis measures fiscal sustainability with two commonly used indicators—the fiscal balance gap and the debt-to-GDP ratio—which capture only a narrow dimension of fiscal space. As emphasized in IMF (2018), fiscal space is inherently multi-

dimensional, reflecting a government's capacity to raise spending or reduce taxes without undermining debt sustainability or market access. We therefore expand the analysis to include additional fiscal vulnerability indicators, drawing on the cross-country fiscal space dataset by Kose and others (2022) supplemented with variables compiled independently. These indicators capture four main dimensions of fiscal risk.

The first dimension corresponds to our baseline solvency metrics but employs alternative formulations from Kose and others (2022) that assess long-term debt sustainability. Debt-related measures include general government gross debt-to-GDP ratio, debt as a share of average tax revenues, and debt relative to a 10-year moving average of GDP, which scale debt either by fiscal capacity or by a smoothed measure of the output base. On the fiscal balance side, we employ sustainability gap indicators that compare the actual or projected primary balance to the level required to stabilize debt under various macroeconomic scenarios (historical, current, stressed, and benign).

The second dimension captures liquidity and market access pressures, which reflect short-term financing risks and investor confidence. This dimension includes several complementary indicators drawn from the Kose and others (2022) dataset and additional sources. Indicators include gross financing needs (as a share of GDP), the effective nominal interest rate on government debt, debt service as a share of general government revenue, and the share of central government debt maturing within 12 months. To capture market sentiment and external financing conditions, we also include foreign-currency long-term sovereign credit ratings (1–21 scale, higher = stronger) and five-year sovereign credit default swap (CDS) spreads (in basis points). Together these proxies describe refinancing pressures and perceived creditworthiness—key determinants of IMF advice when liquidity risks rise.

The third dimension concerns balance sheet vulnerabilities related to debt composition, rollover risk, and external exposure. Indicators include the share of general government debt denominated in foreign currency, the share of debt securities held by nonresidents, and the total share of government debt held by nonresidents, all of which proxy for sensitivity to exchange rate and external funding shocks. We also consider the share of concessional external debt in total public external debt, which mitigates refinancing risk by providing more stable and predictable funding, and the average maturity of sovereign debt, which reflects the rollover profile of public liabilities. Together, these indicators provide a comprehensive view of how the composition and maturity structure of debt affect fiscal vulnerability and the potential for liquidity pressures to translate into solvency risks.

The fourth dimension encompasses external and private sector debt risks, which can generate contingent liabilities for the public sector and heighten macro-financial vulnerabilities. This set includes broad measures of external leverage—such as total external debt as a share of GDP and external debt in foreign currency as a share of total external debt—as well as indicators of private sector exposure, including private external debt and domestic credit to the private sector (both as shares of GDP). To capture near-term refinancing pressures, we also include short-term

external debt as a share of total external debt, and several ratios comparing external debt to international reserves, both including and excluding gold holdings. These variables together capture the risk that external or private sector imbalances could spill over to the sovereign balance sheet, prompting the Fund to recommend a more conservative fiscal stance to mitigate systemic vulnerabilities.

The final robustness exercise takes a different approach from the preceding analyses. Rather than substituting or augmenting specific explanatory variables, we directly exploit the composite fiscal vulnerability indicators produced by the IMF's Vulnerability Exercise (VE) to examine how the relationship between macro-fiscal conditions and IMF fiscal advice varies with countries' overall level of fiscal fragility. Conceptually, this exercise is similar to the earlier heterogeneity analysis by income group, but here the sample is divided according to measured fiscal vulnerability rather than income level. The VE provides a model-based, cross-country framework for assessing near-term fiscal crisis risk across the Fund's membership. It produces a fiscal risk index estimating the probability of a fiscal crisis over a one- to two-year horizon, using a machine-learning model trained on historical data for 188 countries. A fiscal crisis is defined broadly to include sovereign defaults or restructurings, recourse to exceptional financing (for example, IMF lending above 100 percent of quota), implicit defaults through arrears or very high inflation, and sharp losses of market confidence leading to a sudden stop or large increases in spreads. The VE assessment combines over 100 indicators spanning fiscal, real, external, and institutional dimensions. Country-year observations are ranked by their estimated fiscal risk and categorized into three tiers: low vulnerability (below the 50th percentile), medium vulnerability (50th–80th percentile), and high vulnerability (above the 80th percentile). We interact these vulnerability tiers with the main explanatory variables from the baseline specification to test whether the responsiveness of IMF advice to cyclical and solvency conditions differs systematically across countries with varying levels of fiscal risk.

(ii) Results: Long-Run Sustainability

Replacing the baseline solvency indicators with Kose and others (2022) measures confirms that higher debt—whether scaled by GDP, trend GDP, or tax revenue—is associated with a greater likelihood of fiscal tightening advice.[1] The coefficients are positive and statistically significant. Sustainability gap measures yield similarly positive results, especially under stressed scenarios, where the fiscal adjustment required to stabilize debt is greatest. This suggests IMF advice is particularly sensitive to sustainability concerns in adverse environments, whereas under benign conditions the response is weaker. Overall, these findings bolster the interpretation that the IMF tailors its fiscal recommendations not only to observable debt levels but also to forward-looking assessments of debt stabilization needs. This responsiveness is especially pronounced when projected debt dynamics appear most precarious.

[1] Detailed estimates corresponding to these results are reported in Table AV.17. in Section V.

(iii) Results: Liquidity and Market Access

Gross Financing Needs (GFNs), calculated as the sum of the fiscal deficit and maturing debt over the following 12 months, serve as a key measure of rollover risk. Large GFNs indicate heavier refinancing burdens and greater vulnerability to interest rate or market shocks. The average interest rate on public debt, also from the IMF's Debt Sustainability Analysis (DSA) dataset, reflects the effective cost of borrowing faced by the public sector; higher values may signal investor concerns about liquidity or solvency, particularly in high-debt contexts. The debt service burden as percent of general government revenue, obtained from the IMF's internal WEO database, combines interest payments and principal repayments and captures near-term fiscal pressures associated with debt servicing. The share of central government debt maturing within 12 months as a percent of GDP provides a direct measure of short-term refinancing risk. In addition, sovereign CDS spreads, taken from Bloomberg and J.P. Morgan, serve as a real-time market indicator of perceived sovereign risk and rollover vulnerability, particularly relevant during episodes of financial stress. Finally, we include sovereign credit ratings as a more stable but comprehensive proxy for a country's creditworthiness. Using annual averages of long-term foreign-currency ratings from Moody's, Standard & Poor's, and Fitch (sourced from Bloomberg), we capture a broader set of risk factors, including institutional quality, macroeconomic fundamentals, and fiscal performance. Ratings are averaged across agencies to construct a composite index.

The results from our analysis confirm that market-based indicators of liquidity risk and creditworthiness are strong correlates of IMF fiscal advice.[2] Countries facing tighter liquidity constraints or perceived to be at greater risk of losing market access tend to receive tighter fiscal advice, even after controlling for solvency-related variables. The results also underscore the multidimensional nature of fiscal risk, with both flow and stock vulnerabilities playing an important role in shaping the Fund's recommendations. GFNs are positively and significantly associated with fiscal tightening advice. The estimated coefficient implies that countries facing larger near-term funding pressures tend to be advised to tighten fiscal policy, consistent with the idea that heavy rollover burdens heighten refinancing risk and warrant preemptive adjustment. This finding aligns with IMF guidance that highlights GFN thresholds as early warning indicators in Debt Sustainability Frameworks (DSFs). The effective nominal interest rate on public debt is also positively and significantly related to fiscal tightening advice. A higher average interest rate may signal greater risk premia or refinancing costs and, thus, heighten the urgency for fiscal correction. The large magnitude of the coefficient further suggests that high borrowing costs are viewed by the IMF as a key constraint on fiscal space and a channel through which market signals inform advice. Sovereign credit ratings and CDS spread, the two market perception variables, behave as expected. Lower credit ratings are associated with significantly tighter fiscal advice, and the coefficient is both large and highly significant. Similarly, wider CDS spreads are correlated with tighter advice. These results suggest that the IMF pays close attention to market

[2] Detailed estimates corresponding to these results are reported in Table AV.18. in Section V.

perceptions of sovereign risk when formulating fiscal recommendations. This is consistent with the Fund's role in helping countries maintain or restore market access and manage vulnerabilities associated with adverse shifts in investor sentiment.

The only unexpected result concerns the debt service and the debt maturing variables, which enter the regression with a negative and significant coefficient. In principle, higher debt service and a higher share of debt maturing over the next 12 months should indicate greater liquidity pressure and hence be associated with tighter advice. One possible explanation for this anomaly is omitted variable bias: countries with higher debt service levels may also have stronger fundamentals or greater debt-carrying capacity (e.g., higher income, broader revenue base, more stable investor base), allowing the IMF to adopt a less conservative stance despite elevated repayment obligations. Alternatively, the Fund may expect that these countries can manage high debt service burdens through other means (e.g., liability management operations or donor support), weakening the link between this variable and fiscal advice.

(iv) Results: Balance Sheet Vulnerabilities

Among these five indicators, only the total share of debt held by nonresidents is significantly associated with tighter IMF fiscal stance advice.[3] This suggests that greater reliance on external creditors raises concerns about rollover risk and market confidence, prompting the IMF to favor a more conservative stance. Other indicators show limited importance. Neither foreign currency-denominated government debt nor the subset of debt securities held by nonresidents show significant effects. This suggests that the IMF may be responding more to aggregate external exposure than to the form or currency denomination of that debt (see next section). Similarly, concessional financing terms and average maturity of sovereign debt do not appear to influence fiscal stance advice in a systematic way. These findings imply that while balance sheet vulnerabilities are conceptually important, the IMF's fiscal advice appears particularly attuned to indicators that signal potential loss of market access or external investor confidence, rather than those related to rollover structure or concessionality per se.

(v) Results: External and Private Sector Debt

The results offer nuanced insights into how these external and private-sector indicators relate to IMF fiscal advice.[4] Currency mismatch risk emerges as particularly important. External debt denominated in foreign currency is positively and significantly associated with tighter fiscal recommendations, suggesting that high levels of foreign currency debt heighten concerns about balance sheet vulnerabilities. When a depreciation could sharply raise debt-servicing costs, IMF staff appear more likely to favor a more conservative near-term fiscal stance. In contrast, higher levels of total external debt as a share of GDP are associated with looser fiscal advice, a counterintuitive result that may reflect an omitted variable bias, such as differences in

[3] Detailed estimates corresponding to these results are reported in Table AV.19. in Section V.

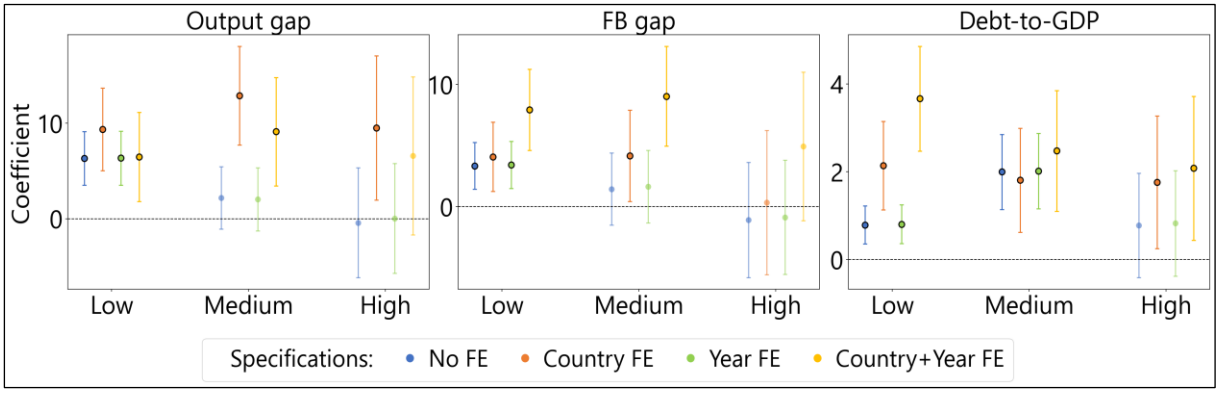
[4] Detailed estimates corresponding to these results are reported in Table AV.20. in Section V.

institutional quality or market access. Countries with stronger financial credibility may sustain higher external debt without triggering alarm, thereby diluting the role of gross debt as a red flag in Fund advice. Private sector indicators also show an unexpected pattern. Both private external debt and domestic credit to the private sector are negatively and significantly associated with tightening advice, suggesting that deeper or more developed financial systems are interpreted as signs of resilience and capacity to absorb shocks, reducing the perceived need for fiscal tightening. The relationship between short-term external debt and near-term fiscal advice is also inverse to conventional concerns about rollover risk. The share of short-term external debt in total external debt is negatively and significantly associated with tightening advice. One possible explanation is that countries with significant short-term borrowing may also have more active liquidity management frameworks or central bank backstops that mitigate this risk in the eyes of Fund staff. Finally, liquidity buffer indicators, such as external debt-to-reserve ratios, show no significant relationship with IMF's near-term fiscal stance advice. This suggests that these ratios, while commonly used in crisis prediction models, may not strongly influence IMF near-term fiscal recommendations.

(vi) Heterogeneity by Fiscal Vulnerability

The results shown in Figure AI.2 confirm that the IMF's fiscal advice response varies in important ways depending on the degree of underlying fiscal vulnerabilities. Interacting fiscal-vulnerability ratings with key explanatory variables reveals systematic differences in IMF advice. Countries facing higher risks are more likely to receive recommendations for fiscal tightening overall, consistent with the expectation that underlying vulnerabilities shape the Fund's advice even before observable cyclical conditions are taken into account.

Figure AI.2. Fiscal Vulnerability Heterogeneity in Determinants of Recommended Near-Term Fiscal Stance



Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Total effect of the interaction term is reported. All regressions use contemporaneous data. Vertical lines indicate 95 percent confidence intervals. Detailed estimates corresponding to these results are reported in Table AV.21. in Section V.

Staff recommendations are more clearly countercyclical when fiscal vulnerabilities are lower. For country-years with low fiscal risk, a positive output gap is strongly associated with tighter fiscal advice. This relationship weakens as vulnerability increases. This relationship remains positive but smaller for medium-risk countries and becomes weaker and statistically insignificant when vulnerabilities are high. This pattern suggests that in more fragile contexts, cyclical signals play a smaller role in shaping advice, which likely reflects the overriding need to stabilize debt dynamics or rebuild fiscal buffers.

The relationship between fiscal advice and the fiscal balance gap is strongest when fiscal vulnerabilities are low. For countries with low fiscal vulnerability, a larger required adjustment to stabilize the debt-to-GDP ratio is strongly associated with tighter fiscal advice. In contrast, for countries classified as facing high fiscal vulnerabilities, the coefficients are either negative or statistically insignificant. This likely reflects the reality that highly vulnerable countries rarely have the flexibility to deviate from debt-stabilizing fiscal paths.

Debt levels remain a robust predictor of fiscal tightening advice across most vulnerability categories. For medium-risk countries, the association between debt and tightening recommendations is particularly strong and consistent, with coefficients ranging from 2.0 to 2.5 across specifications. For low-risk cases, debt levels also significantly influence fiscal advice. By contrast, the association between debt and fiscal recommendations weakens considerably for High-vulnerability country-years and becomes statistically insignificant in several specifications. This may reflect an omitted variable problem since countries facing high fiscal vulnerabilities typically do not accumulate very large debt stocks because market access is constrained and borrowing space is limited. As a result, observed debt levels in these settings may understate underlying risks, which reduces the explanatory power of the debt-to-GDP ratio in driving staff advice.

II. LLM-Based Dataset Construction and Diagnostics

This section documents the main descriptive features and internal diagnostics of the dataset constructed from IMF Article IV consultation reports. It first summarizes the coverage of the sample across income groups and area departments, highlighting the frequency of surveillance over 1998–2023. It then examines the persistence of IMF policy advice across consultations, providing a descriptive sense of how often fiscal, monetary, and macroprudential recommendations change over time. The final two subsections assess the performance of the large language model (LLM) classification, reporting the model's self-assessed certainty scores and testing the sensitivity of results to alternative certainty thresholds. Together, these descriptive and diagnostic exercises provide a comprehensive overview of the underlying dataset used in the empirical analysis.

A. Coverage of Article IV Reports

Table AII.1 summarizes the coverage of Article IV consultation reports included in the dataset. The sample comprises 3,115 reports across 193 countries from 1998–2023, averaging roughly one consultation every 1.6 years per country. AEs are covered most frequently, with an average

interval of 1.3 years between consultations, followed by EMMIEs at 1.6 years LICs at 2 years. Coverage is well distributed across IMF area departments, with the European Department accounting for the largest share of reports (826), reflecting the relatively high concentration of AEs in that region.

Table AII.1. Coverage of Article IV Reports by Income Group and Area Department, 1998–2023

	Number of AIVs	Number of Countries	Number of AIVs Per Country	Average Frequency (Years between AIV)
Income Groupings				
Advanced Economies (AEs)	779	39	20	1.3
Emerging Market and Middle-Income Economies (EMMIEs)	1565	96	16.3	1.6
Low-Income Countries (LICs)	771	58	13.3	2.0
IMF Departments				
European Dept (EUR)	826	45	18.4	1.4
Western Hemisphere Dept (WHD)	579	35	16.5	1.6
Asia-Pacific Dept (APD)	633	36	17.6	1.5
African Dept (AFR)	603	45	13.4	1.9
Middle East & Central Asia Dept (MCD)	469	31	15.1	1.7
Total	3115	193	16.1	1.6

Sources: IMF Article IV Reports; IEO staff calculations.
Note: Average frequency is calculated as the number of years in the sample (1998–2023, i.e., 26 years) divided by the number of Article IV reports per country.

B. Persistence of IMF Policy Advice

Table AII.2 reports how often the recommended policy stance changes between consecutive Article IV consultations. On average, near-term fiscal advice changes in about one-third of cases (32 percent), whereas medium-term fiscal guidance is more stable, with changes recorded in about 8 percent of observations. Monetary and macroprudential advice show comparable frequencies of change, though most of the shifts in the macroprudential category reflect classification transitions—for example, between "Unclear" and "Tighten"—rather than substantive reversals in policy direction (96% of that 34% is a shift between Unclear/Neutral and Tighten).

Table AII.2. Frequency of Changes in Policy Advice Between Consecutive Article IV Reports
(In percent)

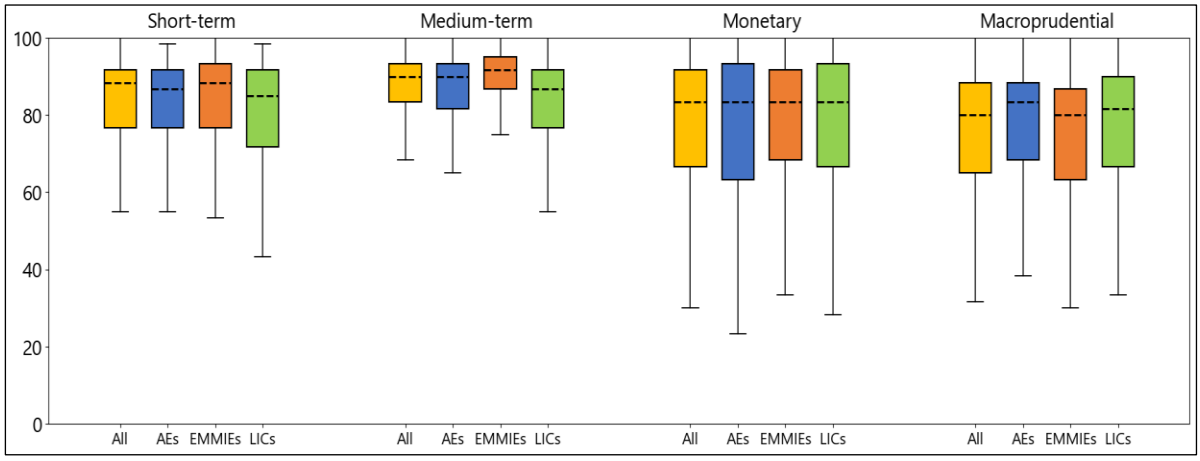
	All	AE	EMMIE	LIC
Fiscal near-term stance	32	39	30	30
Fiscal medium-term stance	8	12	5	10
Monetary stance	29	26	31	26
Macroprudential stance	34	33	35	33

Sources: IMF Article IV Reports; IEO staff calculations.

C. Model Certainty and Classification Confidence

Figure AII.1 displays the distribution of LLM classification certainty scores across four policy areas—near-term fiscal stance, medium-term fiscal stance, monetary policy, and macroprudential policy—for all countries and by income group. The certainty scores represent the model's self-reported confidence (on a 0–100 scale) in the accuracy of its classification, generated after reasoning through each case. We explicitly instruct the model in the prompt to report a certainty score for each classification. The boxplots show the median (black dashed line), interquartile range, and overall dispersion of certainty values. Certainty levels are generally high across income groups, with medians in the mid-80s to 90s range. Differences across groups are modest: AEs and EMMIEs tend to show slightly higher median certainty in some areas, while LICs display somewhat greater variability.

Figure AII.1. Certainty Weights Across Policy Areas and Income Groups



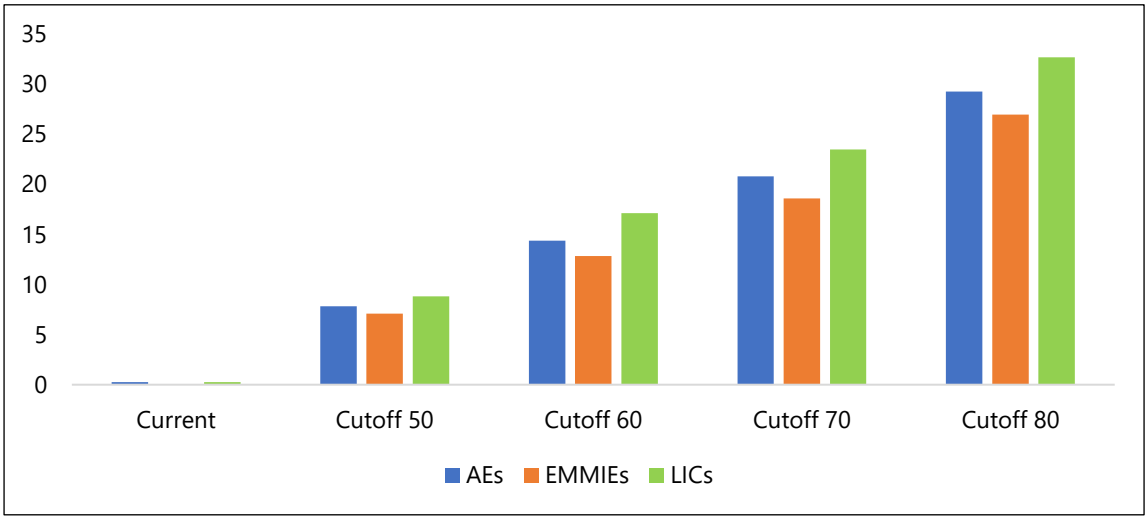
Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Boxplots show the distribution of LLM classification certainty weighs across four policy areas (short-term, medium-term, monetary, and macroprudential) for all countries and by income group (AE, EMMIE, LIC). The black dashed line within each box represents the median certainty, while the boxes and whiskers indicate the interquartile range and overall dispersion of certainty scores.

D. Sensitivity to Certainty Cutoffs

Figure AII.2 illustrates how the proportion of country-year Staff Appraisals classified as Unclear evolves under different certainty cutoffs applied to the LLM-based classification for the short-term fiscal stance. In this exercise, any classification with a final certainty score below the cutoff threshold (50, 60, 70, or 80 percent) is reclassified as Unclear. The bars show the percentage of Unclear cases relative to the total number of country-year observations within each income group. While the baseline (current) classification yields almost no Unclear cases, the share rises steadily as the cutoff increases, reaching around one-third of observations for LICs at the 80 percent threshold.

Figure AII.2. The Percentage of Unclear Short-Term Fiscal Stance

(In percent)



Source: IEO staff calculations.
Note: Percentages indicate the share of Unclear classifications relative to the total number of country-year observations within each income group.

III. Data Description and Baseline Robustness Checks

This section provides descriptive statistics, explains key variable constructions, and presents baseline robustness tests for the empirical analysis. It begins by summarizing how the frequency of fiscal tightening advice varies with countries' macroeconomic and fiscal conditions, offering a descriptive perspective on the dataset. It then reports summary statistics for the main explanatory variables and details how the fiscal balance gap and other covariates are constructed and scaled for estimation. The final subsections assess the robustness of the baseline regression, first to the use of alternative WEO data vintages and then to the choice of large language model (LLM) used in classifying the fiscal stance. Together, these exercises establish the empirical foundations of the analysis and confirm that the main results are stable across data sources and model specifications.

A. Distribution of Fiscal Advice by Macroeconomic Conditions

Table AIII.1 cross-classifies the share of IMF fiscal advice recommending tightening by countries' cyclical position and level of fiscal vulnerability. The table shows the proportion of tightening recommendations across three output-gap categories—positive, near balance, and negative—and three fiscal-risk tiers (low, medium, and high) derived from the Vulnerability Exercise (VE). Overall, the share of tightening advice tends to increase with higher assessed fiscal risk and, to a lesser extent, with stronger cyclical conditions. Patterns are broadly similar across income groups, though the small number of observations in some cells, especially for LICs, calls for caution in interpretation. The table provides a descriptive summary of how the frequency of tightening recommendations varies with both cyclical and fiscal-risk classifications.

Table AIII.1. Distribution of Fiscal Tightening Advice by Macroeconomic and Fiscal Conditions
(Percent of cases recommending tightening)

	Output Gap	Fiscal Vulnerabilities		
		Low Risk	Medium Risk	High Risk
All Countries	Positive	71	79	84
	Close to balance	64	84	78
	Negative	60	74	85
AEs	Positive	70	83	100
	Close to balance	58	100	100
	Negative	58	87	100
EMMIEs	Positive	71	88	89
	Close to balance	70	89	60
	Negative	63	76	91
LICs	Positive	71	71	82
	Close to balance	100	71	88
	Negative	83	59	74

Sources: Vulnerability Exercise (VE) assessment; WEO; IEO staff calculations.
Note: Output gap: "Positive" indicates an output gap above 0.5 percent of GDP; "Close to balance" corresponds to a gap between -0.5 percent and +0.5 percent; and "Negative" refers to a gap below -0.5 percent. Fiscal vulnerabilities are classified as Low, Medium, or High according to the IMF's VE for the fiscal sector (IMF, 2021). The VE is a country-specific assessment of near-term macroeconomic risks, estimating the likelihood of various stress events over a 1-to-2-year horizon. These risks are converted into model-based categories: countries above the 80th percentile are classified as High risk, those below the median as Low risk, and those in between as medium risk.

B. Summary Statistics

Table AIII.2 reports descriptive statistics for the main explanatory variables used in the baseline regressions—the output gap, the fiscal balance (FB) gap, and the debt-to-GDP ratio—covering 189 countries over 1998–2023.

Table AIII.2. Summary Statistics on Key Explanatory Variables
(In percent)

	1st Quartile	Mean	Median	3rd Quartile
Output gap				
All	-1.07	2.72	0.50	6.09
AEs	-1.88	-0.52	-0.57	0.31
EMMIEs	-1.11	2.09	0.46	4.29
LICs	3.87	7.91	8.35	12.25
FB gap				
All	-4.65	-2.23	-1.39	1.33
AEs	-3.64	-1.17	-1.03	1.62
EMMIEs	-4.81	-2.12	-1.34	1.44
LICs	-6.15	-3.74	-2.14	0.85
Debt-to-GDP				
All	29.86	52.28	46.58	68.03
AEs	36.76	61.28	56.87	83.46
EMMIEs	26.59	48.35	44.01	64.50
LICs	30.36	50.41	42.88	61.81

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: The table reports summary statistics for the key explanatory variables used in the baseline model. The dataset covers 189 countries over the period 1998–23. For countries without WEO-provided output gap estimates, we calculated output gaps using the Hamilton filter.

Output gaps combine IMF WEO estimates, when available, with Hamilton-filter estimates when WEO data are missing. Average output gaps are close to zero in AEs, moderately positive in EMMIEs, and substantially higher in LICs. Part of these differences reflects greater reliance on Hamilton-filter estimates in the latter groups: roughly 35 percent of all observations are Hamilton-based—2 percent for AEs, 30 percent for EMMIEs, and 84 percent for LICs. Because Hamilton-filter gaps tend to be smoother and slightly more positive than WEO estimates, this partly explains the higher means observed for EMMIEs and LICs. Although these methodological differences may introduce small artificial level discrepancies across countries, any resulting bias is likely limited (see Section I.A. for robustness tests on the countercyclicality of IMF advice).

The fiscal balance gap measures the difference between the debt-stabilizing fiscal balance and actual fiscal balance. A negative value indicates that the observed fiscal balance is weaker than the level required to stabilize debt. Given the upward trend in public debt over much of the sample period, it is therefore not surprising that fiscal balance gaps are predominantly negative across income groups. This implies that, on average, fiscal deficits would need to be reduced—or surpluses increased—to prevent further rises in debt ratios. Debt ratios are highest in AEs and somewhat lower in EMMIEs and LICs.

C. Construction and Preparation of Explanatory Variables

Definition and Construction of the Fiscal Balance Gap

The fiscal balance gap (FB gap) measures the distance between a country's actual fiscal balance and the fiscal balance required to stabilize the debt-to-GDP ratio at a given level of debt d^*. Intuitively, it captures the size of fiscal adjustment needed to prevent debt from rising further. A negative gap implies that the actual fiscal balance is weaker than the debt-stabilizing level—meaning deficits would need to be reduced or surpluses increased to stabilize debt dynamics.

We build on the calculations performed by Kose and others (2022). Formally, their fiscal sustainability gap is given by:

$$fbsusgap = b - (\frac{-\gamma}{1+\gamma})d^*$$

where $\boldsymbol{b}$ is the overall fiscal balance, $\boldsymbol{\gamma}$ is the nominal output growth rate (a weighted average of the percent change in GDP expressed in local currency and in U.S. dollars at current exchange rates), and $\boldsymbol{d}^*$ is the target debt ratio defined as the historical median value of a country's peer groups of AEs and EMDEs.

For our purposes, we adapt this formulation to obtain a narrower measure that isolates the fiscal adjustment needed to reach the debt-stabilizing balance, rather than to maintain debt at a specific target level $\boldsymbol{d}^*$. We do so because our baseline specification already includes the debt-to-GDP ratio as a separate explanatory variable capturing cross-country differences in debt levels.

To reconstruct the debt-stabilizing balance, we first rearrange Kose and others (2022)'s formulation as follows:

$$(\frac{-\gamma}{1+\gamma}) = \frac{b - fbsusgap}{d^*}$$

and apply this expression to actual country-year debt data, replacing the target debt ratio d^* with the lagged actual debt ratio d_{t-2}. The debt-stabilizing fiscal balance at time t-1 is then:

$$b_{t-1}^{stab} = \left(\frac{-\gamma}{1+\gamma}\right) d_{t-2}$$

The term b_{t-1}^{stab} assumes debt remains constant between t-2 and t-1 (i.e., $d_{t-1} - d_{t-2} = 0$). Finally, we change the sign convention to define the fiscal balance gap as the difference between the debt-stabilizing balance and the actual balance, such that:

$$FB\ gap_{t-1} = b_{t-1}^{stab} - b_{t-1}$$

This sign convention ensures that higher values of the FB gap correspond to greater fiscal adjustment needs, facilitating interpretation of the regression results—larger positive gaps indicate that stronger fiscal consolidation would be required to stabilize debt dynamics.

D. Robustness to Data Vintage

Table AIII.3 reports the results of a robustness check using fixed-vintage data series to assess whether the use of different WEO vintages materially affects the estimated relationships. In this specification, the output gap and debt-to-GDP ratio are taken from the October 2024 WEO vintage and applied consistently across all years in the sample whenever available. When a historical value is missing from that vintage, the corresponding observation is drawn from the most recent earlier vintage.

The results are consistent with the baseline estimates reported in Table 2 of the main paper. The signs, magnitudes, and significance levels of the coefficients on the output gap, fiscal balance gap, and debt-to-GDP ratio remain broadly unchanged across specifications, confirming that the key determinants of IMF fiscal advice are robust to the use of ex-post rather than contemporaneous data. While later-vintage gaps are somewhat smoother—reflecting the incorporation of information unavailable to staff at the time of surveillance—these differences have only minor effects on the estimated relationships. Overall, the findings indicate that the main results are not sensitive to data vintage or real-time measurement issues.

Table AIII.3. Near-Term Stance, Fixed Vintage				
Variables	(1)	(2)	(3)	(4)
Output gap	3.705***	5.468***	3.712***	2.170
	(0.835)	(1.516)	(0.840)	(1.709)
FB gap	1.791**	3.347***	1.905***	7.834***
	(0.708)	(1.024)	(0.712)	(1.274)
Debt-to-GDP	1.160***	1.874***	1.185***	2.765***
	(0.176)	(0.368)	(0.178)	(0.452)
τ_1	0.455***	2.483	0.719**	2.660
	(0.101)	(1.398)	(0.283)	(1.452)
τ_2	1.342***	3.574**	1.613***	3.936***
	(0.108)	(1.399)	(0.286)	(1.454)
N	2741	2575	2741	2575
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.127	0.155	0.133	0.242
Pseudo R^2 (Ugba & Gertheiss)	0.283	0.338	0.295	0.493
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

E. Robustness to Alternative LLM Classifications

Table AIII.4 reports a robustness check assessing whether the results depend on the specific LLM classification used to determine the short-term fiscal stance. The baseline specification defines the stance through a majority-vote across three models—OpenAI's o1, Anthropic's Claude 3.7 Sonnet, and DeepSeek's R1. To test sensitivity, columns (2)–(4) replace the majority-vote classification with the output from each individual model in turn, while column (1) reproduces the baseline estimates from Table 2 of the paper for comparison.

The results show a high degree of consistency across all models. The coefficients on the output gap, fiscal balance gap, and debt-to-GDP ratio remain positive, statistically significant, and similar in magnitude to the baseline, confirming that the core relationships are not driven by the specific LLM used. The coefficient on the output gap remains the largest and most significant variable in every case—ranging from 5.33 to 6.61—indicating that stronger cyclical conditions robustly increase the likelihood of tightening advice. Overall, the findings demonstrate that the results are stable and reproducible across alternative LLM classifications, reinforcing confidence in the robustness of the baseline specification.

Table AIII.4. Near-Term Stance, Different LLM Classifications				
Variables	(1)	(2)	(3)	(4)
Output gap	5.671***	5.333***	5.380***	6.607***
	(0.951)	(0.923)	(0.933)	(1.035)
FB gap	1.752**	1.916***	1.308*	1.794**
	(0.726)	(0.713)	(0.728)	(0.762)
Debt-to-GDP	1.305***	1.355***	1.104***	1.202***
	(0.182)	(0.179)	(0.175)	(0.190)
τ_1	0.362***	0.278***	0.392***	0.573**
	(0.104)	(0.102)	(0.102)	(0.108)
τ_2	1.251***	1.185***	1.340**	1.440***
	(0.110)	(0.108)	(0.109)	(0.115)
N	2720	2722	2673	2709
Country FE	No	No	No	No
Year FE	No	No	No	No
Pseudo R^2 (McFadden)	0.145	0.145	0.140	0.140
Pseudo R^2 (Ugba & Gertheiss)	0.318	0.320	0.309	0.308
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. Column (1) reshows the estimates from Table 2 column (1) of the paper.

IV. Econometric Framework and Model Diagnostics

This section describes the estimation framework used to model IMF near-term fiscal stance advice and presents a series of diagnostic and robustness checks. We first outline the ordered logit specification used in the baseline analysis and explain how country and year effects are incorporated. We then discuss diagnostic tools—focusing on surrogate residuals and the proportional odds assumption—and conclude with robustness tests using alternative estimators, including a partial proportional odds model, an OLS linear specification, and a binary logit model.

A. Ordered Logit Model

The ordered logit model provides a convenient framework for estimating how macroeconomic and fiscal conditions influence the likelihood that IMF staff recommend loosening, maintaining, or tightening the fiscal stance. Let $\mathbf{y} \in \{\mathbf{Loosen}\ (\mathbf{1}), \mathbf{Neutral}\ (\mathbf{2}), \mathbf{Tighten}\ (\mathbf{3})\}$ denote the observed ordinal dependent variable, which reflects an unobserved continuous latent variable $\mathbf{y}^*$ determined as:

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Where $\mathbf{y}^*$ is an $\mathrm{N} \times 1$ vector of latent dependent variables for N observations, $\mathbf{X}$ is an $\mathrm{N} \times \mathrm{K}$ matrix of independent variables, where K is the number of independent variables, $\boldsymbol{\beta}$ is a $\mathrm{K} \times 1$ vector of coefficients, $\boldsymbol{\epsilon}$ is an $\mathrm{N} \times 1$ vector of error terms which follows a logistic distribution. The observed ordinal outcome $\mathbf{y}$ depends on whether $\mathbf{y}^*$ crosses threshold parameters $\boldsymbol{\tau_1}$ and $\boldsymbol{\tau_2}$:

$$y = \begin{cases} \text{Loosen} & \text{if } \mathbf{y}^* \leq \tau_1 \\ \text{Neutral} & \text{if } \tau_1 < \mathbf{y}^* \leq \tau_2 \\ \text{Tigthen} & \text{if } \mathbf{y}^* > \tau_2 \end{cases}$$

The cumulative probability of being in category $\boldsymbol{j}$ (fiscal stance) or higher is:

$$P(y \geq j \mid \mathbf{X}) = \frac{1}{1 + \exp\left(-(\mathbf{X}\boldsymbol{\beta} - \tau_j)\right)}$$

and the probability of being in category $\boldsymbol{j}$ is the difference between adjacent cumulative probabilities:

$$P(y = j \mid \mathbf{X}) = P(y \geq j \mid \mathbf{X}) - P(y \geq j + 1 \mid \mathbf{X})$$

The estimation for $\boldsymbol{n}$ observations and $\boldsymbol{J = 3}$ categories is performed using maximum likelihood estimation (MLE) with the following log-likelihood function:

$$L(\boldsymbol{\beta}, \boldsymbol{\tau}) = \sum_{i=1}^{n} w_i \sum_{j=1}^{J} d_{ij} \log P(y_i = j \mid \mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\tau})$$

where the $\boldsymbol{d_{ij}}$ is an indicator equal to 1 if $\boldsymbol{y_i = j}$, and 0 otherwise, $\boldsymbol{w_i}$ is the weight of observation $\boldsymbol{i}$, and $\mathbf{log}\, \boldsymbol{P}(\boldsymbol{y_i = j} \mid \mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\tau})$ is the probability of observing category $\boldsymbol{j}$ for observation $\boldsymbol{i}$ as described in the equation above. We compute weights $\boldsymbol{w_i}$ based on the certainty values provided by the LLMs when classifying short-term fiscal stance advice.

When we employ individual country and year fixed effects, we simply extend the model with country and/or year dummy variables, which allows us to effectively account for unobserved heterogeneity across countries and years. The latent model then becomes $\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$, where $\mathbf{Z}$ is an $\mathbf{N} \times (\mathbf{M} - \mathbf{1})$ matrix of country dummy variables and $\boldsymbol{\gamma}$ is a $(\mathbf{M} - \mathbf{1}) \times \mathbf{1}$ vector of country fixed effects (coefficients for each country dummy) and/or year fixed effects, where $\mathbf{M}$ is the number of countries.

B. Model Diagnostics: Surrogate Residuals

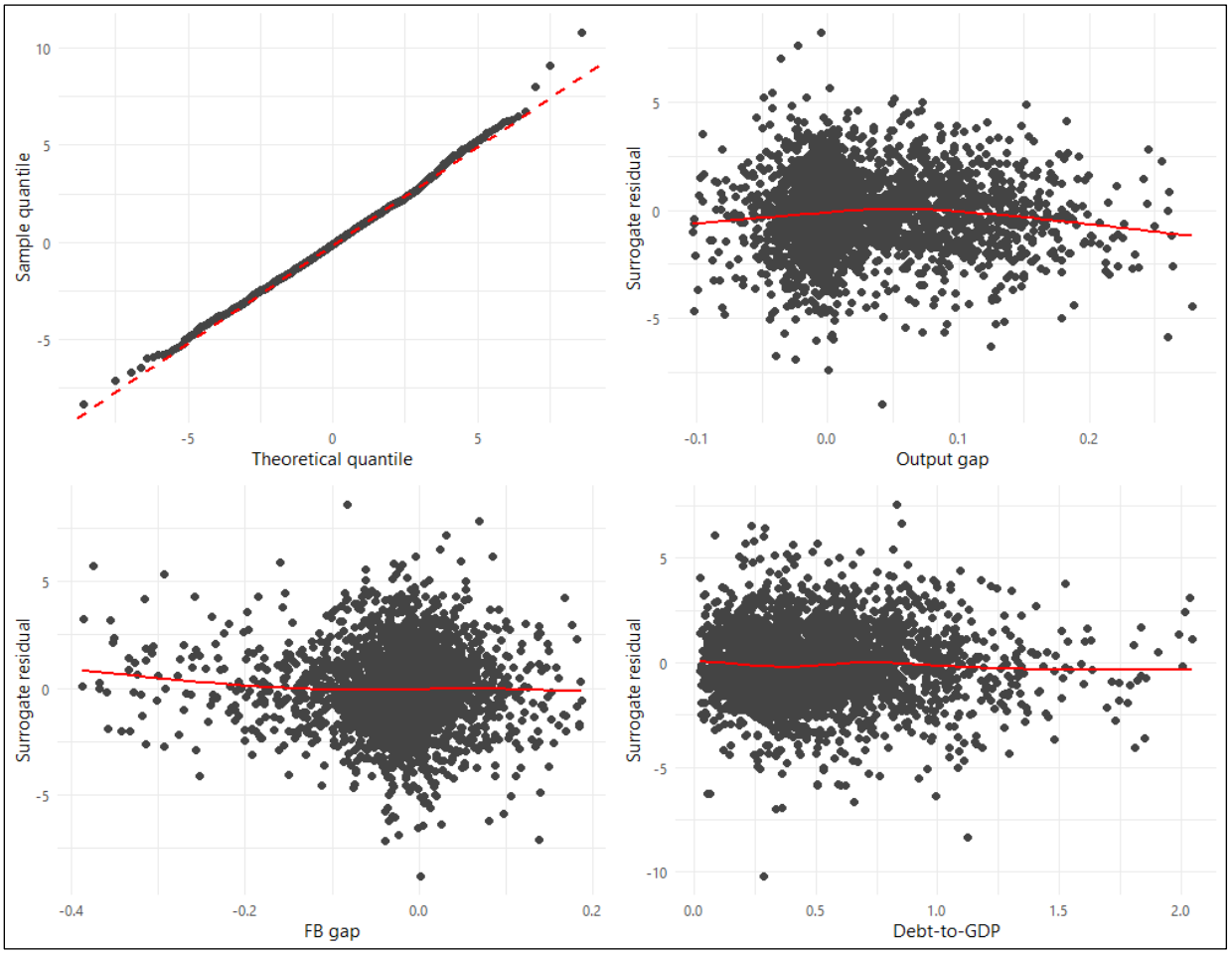
Because the dependent variable is ordinal, conventional residuals are not directly defined. To evaluate model fit, we use surrogate residuals proposed by Liu and Zhang (2017). To define surrogate residuals, let $\boldsymbol{S}$ be a continuous surrogate variable drawn from the truncated distribution of the latent variable $\mathbf{y}^*$, conditional on the observed category (i.e., recommended short-term fiscal stance) $\mathbf{y}$ and the covariates $\mathbf{X}$. Specifically, conditional on $\boldsymbol{y = j}$, the surrogate variable $\boldsymbol{S}$ is sampled from a truncated logistic distribution with location $\mathbf{X}\boldsymbol{\beta}$, constrained within the interval $(\boldsymbol{\tau_{j-1}}, \boldsymbol{\tau_j}]$, where $\boldsymbol{\tau_j}$ are the threshold parameters dividing the ordinal categories. Then the surrogate residuals $\boldsymbol{R_S}$ is:

$$R_S = S - \mathrm{E}(S \mid \mathbf{X})$$

Here $\mathbf{E}(\boldsymbol{S} \mid \mathbf{X})$ is the expected value of the surrogate variable given the covariates, and $\boldsymbol{R_S}$ serves as a continuous analog of a residual. Liu and Zhang (2017) propose that if the fitted model is close to the true model, then $\boldsymbol{R_S}$ has the following properties: (1) zero mean, (2) constant variance, and (3) follows a logistic distribution. Surrogate residuals are valid mostly for model diagnostics.

Figure AIV.1 shows the diagnostic plots. The Q–Q plot in the upper-left panel shows that the surrogate residuals track the 45-degree reference line over the bulk of the distribution, supporting the assumption of logistic errors. The residual–covariate plots suggest that the marginal effect of the output gap diminishes for very large positive values, while the fiscal balance gap exerts a slightly weaker influence at wider deficits. The debt-to-GDP ratio appears to have limited additional effect beyond about 100 percent of GDP, suggesting potential nonlinearity at high debt levels.

Figure AIV.1. Surrogate Residuals Diagnostic Plot for the Baseline Specification



Source: IEO staff calculations.
Note. Top-left: Q-Q plot of the residuals. Top-right: Residual-vs-covariate (Output gap). Bottom-left: Residual-vs-covariate (FB gap). Bottom-right: Residual-vs-covariate (Debt-to-GDP). The model here refers to the baseline results specified in Table 2 column (1) of the paper.

C. Proportional Odds Assumption

A central assumption of the ordered logit model is proportional odds, meaning that explanatory variables have the same effect on the odds of moving up the ordinal scale, regardless of the specific threshold. Formally:

$$\log\left(\frac{P(y \leq j)}{1 - P(y \leq j)}\right) = \tau_j - \mathbf{X}\boldsymbol{\beta}$$

where $\boldsymbol{P(y \leq j)}$ is the probability that the ordinal outcome is less than or equal to category $\boldsymbol{j}$, $\mathbf{X}\boldsymbol{\beta}$ is a linear combination of the predictors and coefficients, and $\boldsymbol{\tau_j}$ are the threshold values that separate the ordinal categories.

Table AIV.1 shows the results from the Brant test for proportional odds assumption. The omnibus test is an overall test that checks whether the proportional odds assumption holds for the model as a whole, rather than for each individual variable. Since the p-value is less than 5 percent significance level, the proportional odds assumption does not hold. To address this, we estimate a partial proportional odds model (Table AIV.2) that relaxes the constraint for variables where the assumption does not hold. The results show that the output gap effect is stronger at the first threshold (moving from Loosen to Neutral/Tighten) than at the second (from Neutral to Tighten), suggesting some variation in its influence across decision points. The fiscal balance gap becomes more influential at the second threshold, while debt-to-GDP maintains a consistent and significant effect across both. Despite minor differences, the qualitative results mirror those of the baseline model, confirming that the main conclusions are robust.

Table AIV.1. Results from Brant Test

Test for	χ^2	df	p-value
Omnibus	22.729	3	4.6e-05***
Output gap	14.761	1	0.0001***
FB gap	0.487	1	0.485
Debt-to-GDP	2.997	1	0.083*

Source: IEO staff calculations.

Table AIV.2. Baseline Regression, Partial Proportional Odds				
Variables	(1)	(2)	(3)	(4)
Output gap τ_1	8.744***	13.245***	8.738***	9.319***
	(1.384)	(2.113)	(1.386)	(2.322)
Output gap τ_2	5.254***	9.836***	5.263***	6.430***
	(0.950)	(1.799)	(0.956)	(2.002)
FB gap τ_1	1.314	2.482*	1.365	8.058***
	(0.950)	(1.332)	(0.951)	(1.536)
FB gap τ_2	1.853**	3.763***	1.992***	8.220***
	(0.737)	(1.082)	(0.743)	(1.341)
Debt-to-GDP τ_1	1.118***	2.168***	1.134***	2.988***
	(0.237)	(0.422)	(0.238)	(0.504)
Debt-to-GDP τ_2	1.340***	2.347***	1.364***	3.287***
	(0.184)	(0.380)	(0.186)	(0.463)
τ_1	0.358***	2.373	0.624**	2.514
	(0.105)	(1.399)	(0.287)	(1.449)
τ_2	1.291***	3.469**	1.567***	3.873***
	(0.132)	(1.404)	(0.99)	(1.455)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

D. Alternative Estimators: Linear and Binary Specifications

For additional robustness, we estimate the baseline model using ordinary least squares (OLS) and a binary logit specification. In the OLS regression (Table AIV.3), the ordinal outcome is treated as a continuous variable (Loosen = 1, Neutral = 3, Tighten = 5). The signs and relative magnitudes of coefficients are consistent with the ordered logit results, reaffirming that stronger cyclical conditions and weaker fiscal positions are associated with tighter fiscal advice.

Table AIV.3. Baseline Regression, OLS				
Variables	(1)	(2)	(3)	(4)
Output gap	3.346***	4.759***	2.776***	2.582***
	(0.453)	(0.869)	(0.531)	(0.821)
FB gap	0.922**	1.199**	1.857***	2.654***
	(0.381)	(0.551)	(0.550)	(0.613)
Debt-to-GDP	0.627***	0.853***	0.590***	0.885***
	(0.083)	(0.163)	(0.105)	(0.178)
Intercept	3.859***			
	(0.054)			
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Adjusted R^2	0.032	0.175	0.142	0.292

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The near-term fiscal stance is coded as 1 for easing, 3 for neutral, and 5 for tightening.

The binary logit model (Table AIV.4) collapses the outcome into a dichotomous variable (Tighten = 1; Loosen/Neutral = 0). The estimated coefficients and fit statistics remain closely aligned with the ordered logit estimates, confirming that the results are not sensitive to the functional form or to the ordinal coding of the dependent variable.

Table AIV.4. Baseline Regression, Binary Logit

Variables	(1)	(2)	(3)	(4)
Output gap	5.194***	9.082***	4.786***	5.747**
	(0.950)	(1.886)	(0.916)	(2.373)
FB gap	1.786**	3.264***	3.390***	7.634***
	(0.740)	(1.242)	(1.036)	(1.914)
Debt-to-GDP	1.339***	2.295***	1.472***	3.198***
	(0.185)	(0.462)	(0.230)	(0.664)
Intercept	0.358***			
	(0.105)			
N	2720	2481	2720	2481
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Squared correlation	0.032	0.195	0.114	0.294

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The squared correlation is the correlation between the model's fitted probabilities and the actual binary outcomes, serving as a pseudo-R^2 measure. For instance, a value of 0.294 indicates (column 4) that about 29 percent of the variation in observed outcomes is linearly associated with the model's predicted probabilities. In columns (2) and (4), the number of observations is lower than in the baseline ordered logit model because, after converting the three-category outcome (Loosen, Neutral, Tighten) into a binary variable (Tighten = 1, otherwise = 0), some fixed effects were dropped due to containing only one type of outcome (all 0s or all 1s).

V. REGRESSION TABLES

Table AV.1. Income Group Heterogeneity in Determinants of Recommended Near-Term Fiscal Stance

Variables	(1)	(2)	(3)	(4)
Output gap × AEs	13.314***	13.316***	13.525***	11.092***
	(3.798)	(4.024)	(3.876)	(4.211)
Output gap × EMMIEs	6.450***	12.858***	6.474***	8.934***
	(1.561)	(2.717)	(1.565)	(2.947)
Output gap × LICs	1.363	7.421**	1.246	3.030
	(1.865)	(3.234)	(1.876)	(3.563)
FB gap × AEs	4.286***	1.178	4.468***	6.403**
	(1.564)	(2.235)	(1.578)	(2.493)
FB gap × EMMIEs	2.878***	7.617***	2.992***	12.558***
	(1.057)	(1.637)	(1.064)	(1.878)
FB gap × LICs	-1.045	0.338	-0.940	3.230
	(1.582)	(2.038)	(1.585)	(2.191)
Debt-to-GDP × AEs	0.721***	0.617	0.763***	2.692***
	(0.267)	(0.649)	(0.269)	(0.735)
Debt-to-GDP × EMMIEs	2.345***	3.120***	2.364***	4.321***
	(0.327)	(0.667)	(0.328)	(0.793)
Debt-to-GDP × LICs	1.651***	2.681***	1.657***	2.356***
	(0.483)	(0.724)	(0.485)	(0.793)
AEs	-0.022	1.367	-0.055	-0.534
	(0.337)	(1.071)	(0.339)	(1.156)
EMMIEs	-0.133	2.135	-0.149	1.198
	(0.325)	(1.621)	(0.326)	(1.668)
τ_1	0.368	0.035	0.662*	0.978
	(0.285)	(0.792)	(0.396)	(0.928)
τ_2	1.279***	1.141	1.580***	2.273**
	(0.288)	(0.793)	(0.398)	(0.930)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.165	0.169	0.170	0.254
Pseudo R^2 (Ugba & Gertheiss)	0.357	0.365	0.367	0.512
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data.

Table AV.2. Changes in the Countercyclicality of Fiscal Advice Over Time			
Variables	(1)	(2)	(3)
Output gap	5.884***	2.675*	3.259*
	(0.965)	(1.533)	(1.790)
FB gap	2.212***	2.556***	1.630**
	(0.725)	(0.731)	(0.753)
Debt-to-GDP	1.440***	1.363***	1.486***
	(0.187)	(0.184)	(0.193)
Year	-0.051***		
	(0.007)		
Post-GFC		-0.768***	
		(0.118)	
Output gap × post-GFC		7.315***	
		(1.191)	
Output gap × 2008–09			7.377***
			(2.267)
Output gap × 2010–12			0.722
			(2.230)
Output gap × 2013–19			7.870***
			(1.889)
Output gap × 2020–23			12.495***
			(2.858)
2008–09			-1.619***
			(0.197)
2010–12			-0.022
			(0.196)
2013–19			-0.705***
			(0.144)
2020–23			-1.546***
			(0.165)
τ_1	104.84***	0.847***	0.871***
	(14.799)	(0.131)	(0.139)
τ_2	105.75***	1.752***	1.820***
	(14.802)	(0.138)	(0.146)
N	2720	2720	2720
Country FE	No	No	No
Year FE	No	No	No
Pseudo R^2 (McFadden)	0.159	0.157	0.185
Pseudo R^2 (Ugba & Gertheiss)	0.346	0.342	0.394
Proportional odds assumption holds	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data. The output gap estimate in the first row (3.259*) of column (3) refers to the output gap interacted with the year bucket for 1998–2007, the reference year bucket.

Table AV.3. Output Gap and Monetary Stance at ZLB				
Variables	(1)	(2)	(3)	(4)
Output gap	3.444**	8.169***	3.538**	7.073***
	(1.370)	(2.335)	(1.378)	(2.497)
Output gap × ZLB	5.784***	12.160***	5.506***	8.581***
	(1.603)	(2.566)	(1.621)	(2.754)
Monetary policy stance	0.364***	0.351***	0.363***	0.401***
	(0.056)	(0.065)	(0.057)	(0.071)
Monetary policy stance × ZLB	0.535***	0.680***	0.539***	0.596***
	(0.071)	(0.086)	(0.072)	(0.094)
ZLB	-1.192***	-1.746***	-1.193***	-3.756***
	(0.309)	(0.367)	(0.315)	(0.687)
FB gap	4.047***	6.696***	4.128***	10.360***
	(0.861)	(1.347)	(0.869)	(1.614)
Debt-to-GDP	1.254***	3.024***	1.275***	3.999***
	(0.210)	(0.480)	(0.212)	(0.559)
τ_1	-0.634***	1.286	-0.077	0.281
	(0.229)	(1.482)	(0.408)	(1.525)
τ_2	0.362	2.578*	0.928**	1.710
	(0.231)	(1.483)	(0.410)	(1.526)
N	2081	1946	2081	1946
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.359	0.235	0.363	0.287
Pseudo R^2 (Ugba & Gertheiss)	0.663	0.481	0.669	0.563
Proportional odds assumption holds	No	No	Yes	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The years 2009–15 and 2020–21 are coded as ZLB years, based on U.S. monetary policy conditions during which the federal funds rate was at or near the effective lower bound. This classification is used as a proxy for global monetary constraints given the central role of U.S. interest rates in shaping global financial conditions. The monetary stance is coded as 1 for easing, 3 for neutral, and 5 for tightening.

Table AV.4. Macroprudential Stance

Variables	(1)	(2)	(3)	(4)
Output gap	5.700***	10.351***	5.717***	6.895***
	(0.951)	(1.798)	(0.956)	(1.999)
FB gap	1.735**	3.462***	1.855**	8.134***
	(0.726)	(1.065)	(0.730)	(1.316)
Debt-to-GDP	1.310***	2.270***	1.332***	3.209***
	(0.182)	(0.376)	(0.183)	(0.459)
Macroprudential policy stance	0.074	0.066	0.068	0.084
	(0.048)	(0.058)	(0.049)	(0.063)
τ_1	0.061	2.151	0.342	2.249
	(0.222)	(1.415)	(0.351)	(1.469)
τ_2	0.950***	3.247**	1.238***	3.530**
	(0.224)	(1.416)	(0.352)	(1.471)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.145	0.163	0.151	0.249
Pseudo R^2 (Ugba & Gertheiss)	0.320	0.353	0.330	0.504
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The stance for macroprudential measures is coded as 1 for easing, 3 for neutral/unclear, and 5 for tightening.

Table AV.5. External Imbalances

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Output gap	4.821***	10.289***	4.842***	6.819***	4.634***	10.277***	4.643***	6.576***
	(0.990)	(1.854)	(0.994)	(2.058)	(1.000)	(1.858)	(1.004)	(2.062)
FB gap	-0.644	1.632	-0.568	6.382***	-0.399	1.650	-0.291	6.779***
	(0.807)	(0.153)	(0.814)	(1.431)	(0.827)	(1.173)	(0.834)	(1.451)
Debt-to-GDP	1.072***	2.271***	1.091***	3.127***	1.078***	2.273***	1.098***	3.174***
	(0.183)	(0.385)	(0.185)	(0.469)	(0.184)	(0.387)	(0.185)	(0.472)
CA deficit as % of GDP	5.635***	5.825***	5.594***	5.604***				
	(0.668)	(1.233)	(0.672)	(1.348)				
CA surplus only as % of GDP					-4.393***	-5.696***	-4.205***	-2.960
					(1.153)	(2.001)	(1.166)	(2.171)
CA deficit only as % of GDP					6.963***	5.941***	7.058***	8.072***
					(1.247)	(1.920)	(1.250)	(2.156)
τ_1	0.366***	2.089	0.657**	2.221	0.298**	2.082	0.588**	2.068
	(0.106)	(1.403)	(0.293)	(1.462)	(0.118)	(1.406)	(0.297)	(1.465)
τ_2	1.275***	3.193**	1.573***	3.512**	1.207***	3.187**	1.504***	3.360**
	(0.112)	(1.404)	(0.296)	(1.463)	(0.124)	(1.407)	(0.300)	(1.466)
N	2686	2524	2686	2524	2686	2524	2686	2524
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.174	0.171	0.179	0.255	0.174	0.171	0.179	0.256
Pseudo R^2 (Ugba & Gertheiss)	0.374	0.368	0.383	0.514	0.375	0.368	0.384	0.515
Proportional odds assumption holds	No	No	No	No	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.6. External Imbalances by Income Group				
Variables	(1)	(2)	(3)	(4)
Output gap × AEs	10.637***	12.772***	10.688***	10.264**
	(3.936)	(4.201)	(4.015)	(4.313)
Output gap × EMMIEs	6.726***	12.880***	6.738***	8.783***
	(1.612)	(2.805)	(1.613)	(3.048)
Output gap × LICs	1.427	7.695**	1.329	2.949
	(1.906)	(3.340)	(1.916)	(3.665)
FB gap × AEs	-1.080	-0.112	-1.116	5.220**
	(1.817)	(2.316)	(1.841)	(2.605)
FB gap × EMMIEs	0.803	5.200***	0.901	10.660***
	(0.194)	(1.807)	(1.204)	(2.039)
FB gap × LICs	-1.247	-0.043	-1.089	1.798
	(1.688)	(2.175)	(1.694)	(2.353)
Debt-to-GDP × AEs	1.039***	0.974	1.074***	2.810***
	(0.291)	(0.676)	(0.293)	(0.759)
Debt-to-GDP × EMMIEs	1.916***	2.860***	1.937***	2.810***
	(0.341)	(0.675)	(0.343)	(0.759)
Debt-to-GDP × LICs	1.665***	2.724***	1.681***	2.224***
	(0.504)	(0.751)	(0.506)	(0.826)
CA deficit as % of GDP × AEs	11.450***	8.387***	11.492***	6.541**
	(1.587)	(2.894)	(1.602)	(3.042)
CA deficit as % of GDP × EMMIEs	4.098***	6.334***	4.037***	5.168***
	(1.001)	(1.775)	(1.005)	(1.923)
CA deficit as % of GDP × LICs	0.222	1.640	0.121	4.104*
	(1.456)	(2.265)	(1.457)	(2.485)
AEs	-0.104	1.340	-0.137	-0.507
	(0.347)	(1.081)	(0.349)	(1.162)
EMMIEs	-0.050	1.875	-0.137	0.987
	(0.329)	(1.629)	(0.349)	(1.672)
τ_1	0.345	0.027	0.676*	0.993
	(0.288)	(0.800)	(0.404)	(0.927)
τ_2	1.286***	1.143	1.624***	2.296**
	(0.291)	(0.802)	(0.406)	(0.929)
N	2686	2524	2686	2524
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.195	0.177	0.199	0.260
Pseudo R^2 (Ugba & Gertheiss)	0.412	0.380	0.420	0.521
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data.

Table AV.7. External Imbalances: Asymmetry				
Variables	(1)	(2)	(3)	(4)
Output gap × AEs	10.603***	12.715***	10.707***	10.035**
	(4.012)	(4.339)	(4.104)	(4.485)
Output gap × EMMIEs	5.919***	13.014***	5.930***	8.823***
	(1.645)	(2.817)	(1.646)	(3.067)
Output gap × LICs	1.313	7.556**	1.205	2.606
	(1.911)	(3.358)	(1.920)	(3.689)
FB gap × AEs	-0.635	0.826	-0.708	6.415**
	(1.867)	(2.381)	(1.892)	(2.670)
FB gap × EMMIEs	1.593	4.935***	1.745	10.572***
	(1.248)	(1.856)	(1.257)	(2.101)
FB gap × LICs	-0.952	0.224	-0.784	2.247
	(1.720)	(2.191)	(1.726)	(2.356)
Debt-to-GDP × AEs	1.072***	1.085	1.116***	3.066***
	(0.292)	(0.683)	(0.295)	(0.769)
Debt-to-GDP × EMMIEs	1.950***	2.820***	1.971***	3.951***
	(0.344)	(0.676)	(0.346)	(0.806)
Debt-to-GDP × LICs	1.702***	2.750***	1.722***	2.191***
	(0.507)	(0.756)	(0.510)	(0.831)
CA surplus only as % of GDP × AEs	-9.471***	-1.772	-9.538***	2.962
	(2.210)	(4.070)	(2.228)	(4.373)
CA surplus only as % of GDP × EMMIEs	-1.018	-7.686***	-0.812	-5.329*
	(1.769)	(2.836)	(1.777)	(3.044)
CA surplus only as % of GDP × LICs	6.012	3.116	6.870	4.044
	(4.395)	(5.544)	(4.453)	(5.591)
CA deficit only as % of GDP × AEs	16.370***	19.310***	16.349***	21.067***
	(4.252)	(6.010)	(4.242)	(6.186)
CA deficit only as % of GDP × EMMIEs	7.047***	5.011*	7.091***	5.114*
	(1.792)	(2.782)	(1.796)	(3.040)
CA deficit only as % of GDP × LICs	2.496	3.630	2.708	8.179**
	(2.064)	(3.084)	(2.090)	(3.499)
AEs	-0.057	1.210	-0.066	-0.762
	(0.384)	(1.097)	(0.386)	(1.173)
EMMIEs	-0.019	2.095	-0.017	1.216
	(0.357)	(1.639)	(0.360)	(1.686)
τ_1	0.148	-0.112	0.474	0.714
	(0.310)	(0.813)	(0.418)	(0.940)
τ_2	1.090***	1.005	1.424***	2.025**
	(0.313)	(0.814)	(0.420)	(0.942)
N	2686	2524	2686	2524
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.197	0.179	0.202	0.264
Pseudo R^2 (Ugba & Gertheiss)	0.416	0.384	0.425	0.528
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data.

Table AV.8. Countercyclicality by Exchange Rate Regime				
Variables	(1)	(2)	(3)	(4)
Output gap × Flexible	11.933***	14.534***	12.257***	8.159**
	(2.357)	(2.478)	(2.376)	(3.566)
Output gap × Intermediate	4.014***	10.325***	3.940**	6.517**
	(1.521)	(2.761)	(1.541)	(2.754)
Output gap × Fixed	0.470	10.053***	0.318	6.535*
	(1.866)	(3.338)	(1.882)	(3.622)
Flexible	-1.082***	-1.356***	-1.109***	-0.573
	(0.149)	(0.339)	(0.150)	(0.373)
Intermediate	-0.501***	-0.754*	-0.499***	-0.484
	(0.167)	(0.386)	(0.168)	(0.411)
FB gap	2.523***	4.353***	2.667***	8.645***
	(0.824)	(1.185)	(0.830)	(1.476)
Debt-to-GDP	1.551***	2.476***	1.582***	3.311***
	(0.202)	(0.433)	(0.204)	(0.520)
τ_1	0.832***	3.418	1.081***	2.065
	(0.152)	(1.444)	(0.325)	(1.496)
τ_2	1.739***	4.531***	1.998***	3.370**
	(0.158)	(1.446)	(0.328)	(1.498)
N	2325	2182	2325	2182
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.261	0.176	0.267	0.262
Pseudo R^2 (Ugba & Gertheiss)	0.523	0.378	0.533	0.525
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IMF AREAER Database; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.9. Countercyclicality by Exchange Rate Regime and Income Group				
Variables	(1)	(2)	(3)	(4)
Output gap × AE × Flexible	23.464***	28.772***	24.190***	22.820***
	(6.094)	(7.213)	(6.227)	(7.602)
Output gap × EMMIE × Flexible	9.879**	7.383	10.366**	-1.022
	(4.012)	(5.380)	(4.036)	(5.661)
Output gap × LIC × Flexible	7.761*	9.391	7.989*	4.016
	(4.333)	(6.496)	(4.359)	(6.997)
Output gap × AE × Intermediate	11.196**	15.775**	11.464**	12.183*
	(5.361)	(7.203)	(5.588)	(7.107)
Output gap × EMMIE × Intermediate	5.592*	9.058**	5.612*	6.226
	(2.867)	(4.602)	(2.867)	(4.904)
Output gap × LIC × Intermediate	1.466	9.170**	1.128	3.223
	(2.404)	(4.123)	(2.445)	(4.509)
Output gap × AE × Fixed	6.008	7.921	5.799	4.998
	(6.924)	(7.712)	(6.957)	(7.820)
Output gap × EMMIE × Fixed	1.867	13.056***	1.835	7.259
	(2.689)	(4.514)	(2.708)	(4.722)
Output gap × LIC × Fixed	0.407	7.262	0.519	9.560
	(4.485)	(6.655)	(4.577)	(7.849)
FB gap	3.008***	4.467***	3.158***	8.656***
	(0.843)	(1.211)	(0.849)	(1.502)
Debt-to-GDP	1.777***	2.529***	1.816***	3.478***
	(0.211)	(0.437)	(0.213)	(0.530)
AE	0.180	0.663	0.323	0.033
	(0.488)	(2.079)	(0.497)	(2.204)
EMMIE	0.293	-0.895	0.323	-1.962
	(0.443)	(2.747)	(0.451)	(2.806)
Flexible	-0.730	-1.357	-0.713	-0.124
	(0.568)	(1.824)	(0.579)	(1.976)
Intermediate	-0.135	-0.916	-0.076	0.274
	(0.485)	(1.738)	(0.494)	(1.878)
τ_1	0.543	1.370	0.812	0.076
	(0.420)	(1.910)	(0.515)	(2.057)
τ_2	1.464***	2.495	1.743***	1.404
	(0.422)	(1.911)	(0.517)	(2.058)
N	2325	2182	2325	2182
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.270	0.181	0.276	0.270
Pseudo R^2 (Ugba & Gertheiss)	0.538	0.388	0.547	0.537
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IMF AREAER Database; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.10. Program Participation				
Variables	(1)	(2)	(3)	(4)
Output gap	5.594***	10.454***	5.600***	6.968***
	(0.960)	(1.798)	(0.965)	(1.996)
FB gap	1.742**	3.485***	1.862**	8.160***
	(0.726)	(1.066)	(0.731)	(1.317)
Debt-to-GDP	1.295***	2.274***	1.317***	3.280***
	(0.182)	(0.379)	(0.183)	(0.463)
Program	0.070	0.025	0.081	-0.231
	(0.128)	(0.204)	(0.129)	(0.226)
τ_1	0.355***	2.386	0.615**	2.606
	(0.105)	(1.402)	(0.287)	(1.451)
τ_2	1.243***	3.482**	1.510***	3.886***
	(0.111)	(1.404)	(0.290)	(1.453)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.145	0.162	0.150	0.249
Pseudo R^2 (Ugba & Gertheiss)	0.319	0.352	0.329	0.504
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); MONA database; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. In the sample, 21 percent of country-year observations correspond to program participation.

Table AV.11. Program Participation Versus Surveillance				
Variables	(1)	(2)	(3)	(4)
Output gap × No program	8.064***	12.389***	8.154***	9.313***
	(1.184)	(1.973)	(1.193)	(2.181)
Output gap × Program	-0.999	4.962*	-1.215	0.641
	(1.819)	(2.811)	(1.828)	(3.067)
FB gap × No program	2.406***	4.070***	2.522***	8.593***
	(0.805)	(1.190)	(0.812)	(1.429)
FB gap × Program	-0.403	0.783	-0.306	5.952**
	(1.834)	(2.355)	(1.848)	(2.521)
Debt-to-GDP × No program	1.205***	2.334***	1.236***	3.481***
	(0.197)	(0.431)	(0.199)	(0.522)
Debt-to-GDP × Program	2.040***	1.995***	2.024***	2.691***
	(0.501)	(0.626)	(0.502)	(0.721)
Program	-0.087	0.344	-0.042	0.334
	(0.300)	(0.392)	(0.301)	(0.428)
τ_1	0.374***	2.170	0.622**	2.380
	(0.113)	(1.405)	(0.291)	(1.461)
τ_2	1.270***	3.268**	1.525***	3.665**
	(0.118)	(1.406)	(0.294)	(1.463)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.151	0.164	0.156	0.261
Pseudo R^2 (Ugba & Gertheiss)	0.330	0.356	0.341	0.508
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); MONA database; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data.

Table AV.12. Fiscal Institutions				
Variables	(1)	(2)	(3)	(4)
Output gap	4.330***	10.590***	4.344***	7.064***
	(0.985)	(1.847)	(0.990)	(2.006)
FB gap	2.184***	3.944***	2.303***	7.761***
	(0.743)	(1.086)	(0.749)	(1.315)
Debt-to-GDP	1.500***	3.051***	1.518***	3.427***
	(0.189)	(0.414)	(0.190)	(0.469)
Fiscal Rules Only	-0.294**	-0.757***	-0.295**	-0.336
	(0.116)	(0.265)	(0.117)	(0.296)
Fiscal Councils Only	-0.477*	-0.389	-0.506*	0.209
	(0.270)	(0.475)	(0.270)	(0.525)
Both Fiscal Rules and Councils	-0.891***	-2.225***	-0.891***	-1.216***
	(0.150)	(0.332)	(0.151)	(0.409)
τ_1	0.580***	3.924	0.851***	3.526
	(0.121)	(1.427)	(0.295)	(1.504)
τ_2	1.481***	5.051***	1.758***	4.815***
	(0.127)	(1.429)	(0.298)	(1.506)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.154	0.181	0.160	0.253
Pseudo R^2 (Ugba & Gertheiss)	0.337	0.386	0.347	0.511
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); Alonso and others (2025a and 2025b); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The fiscal institutions variables are constructed based on the presence of fiscal rules and fiscal councils, based on the IMF Fiscal Rules and Fiscal Councils Datasets (Alonso and others, 2025a; 2025b). For each country-year observation, a dummy variable for fiscal rules and another for fiscal councils are coded as 1 if present and 0 otherwise. These are then combined into a four-category indicator: None (no rule or council), Rule only (rule without council), Council only (council without rule), and Both (presence of both). "None" serves as the reference category in the regressions, meaning that the coefficients for the other categories represent their effects relative to countries with neither a rule nor a council.

Table AV.13. Fiscal Institutions Interaction				
Variables	(1)	(2)	(3)	(4)
Output gap × No Fiscal Institution	2.570**	8.737***	2.480**	5.817**
	(1.255)	(2.281)	(1.257)	(2.487)
Output gap × Fiscal Institution	7.646***	12.481***	7.820***	8.214***
	(1.597)	(2.505)	(1.618)	(2.652)
FB gap × No Fiscal Institution	2.548**	4.756***	2.691***	8.150***
	(1.000)	(1.437)	(1.004)	(1.624)
FB gap × Fiscal Institution	1.619	2.616	1.701	8.038***
	(1.098)	(1.603)	(1.106)	(1.847)
Debt-to-GDP × No Fiscal Institution	1.987***	3.066***	2.010***	3.086***
	(0.333)	(0.554)	(0.334)	(0.610)
Debt-to-GDP × Fiscal Institution	1.170***	1.824***	1.196***	3.354***
	(0.226)	(0.502)	(0.228)	(0.591)
Fiscal Institution	-0.227	-0.477	-0.239	-0.327
	(0.214)	(0.372)	(0.215)	(0.402)
τ_1	0.468***	2.912	0.776**	2.796
	(0.166)	(1.422)	(0.318)	(1.484)
τ_2	1.365***	4.016***	1.680***	4.077***
	(0.171)	(1.423)	(0.321)	(1.486)
N	2720	2556	2720	2556
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.153	0.168	0.159	0.249
Pseudo R^2 (Ugba & Gertheiss)	0.335	0.364	0.346	0.504
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); Alonso and others (2025a and 2025b); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data. The fiscal institutions variables are constructed based on the presence of fiscal rules and fiscal councils, based on the IMF Fiscal Rules and Fiscal Councils Datasets (Alonso and others, 2025a; 2025b). For each country-year observation, a dummy variable takes the value 1 if there is a presence of any fiscal rule or council, and 0 if neither exists.

Table AV.14. Non-Economic Factors: Saltwater Versus Freshwater Schools				
Variables	(1)	(2)	(3)	(4)
Output gap	3.438***	6.875***	3.382**	4.624*
	(1.225)	(2.418)	(1.238)	(2.586)
FB gap	4.552***	8.859***	4.788***	9.644***
	(0.891)	(1.534)	(0.908)	(1.763)
Debt-to-GDP	1.349***	3.603***	1.375***	2.859***
	(0.222)	(0.578)	(0.224)	(0.614)
Output gap × Saltwater universities	10.296***	12.861***	10.455***	11.132***
	(3.276)	(4.148)	(3.489)	(4.156)
Output gap × Freshwater universities	7.072**	7.546*	7.342**	6.395
	(3.276)	(4.379)	(3.259)	(4.689)
Saltwater universities	-0.131	0.290	-0.123	0.198
	(0.162)	(0.200)	(0.164)	(0.206)
Freshwater universities	-0.356**	-0.237	-0.360**	-0.355
	(0.179)	(0.224)	(0.182)	(0.234)
τ_1	0.702***	1.737	1.187***	1.156
	(0.131)	(1.436)	(0.380)	(1.459)
τ_2	1.654***	2.963**	2.151***	2.515
	(0.141)	(1.438)	(0.385)	(1.461)
N	1995	1873	1995	1873
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.167	0.214	0.176	0.267
Pseudo R^2 (Ugba & Gertheiss)	0.361	0.446	0.378	0.533
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); Lang, Wellner, and Kentikelenis (2024); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. Total effect of the interaction term is reported. All regressions use contemporaneous data. The result is based on data up to 2016, as information on mission chiefs is only available through that year.

Table AV.15. Robustness: Income Group Heterogeneity with Alternative Measures of Cyclical Conditions

Variables	(1)	(2)	(3)	(4)
5-year average growth deviation × AEs	17.462***	28.066***	17.286***	23.644***
	(4.809)	(5.969)	(4.870)	(7.024)
5-year average growth deviation × EMMIEs	9.811***	18.125***	9.807	11.427**
	(3.511)	(4.424)	(3.536)	(5.034)
5-year average growth deviation × LICs	4.217	3.645	4.864	-2.052
	(5.422)	(6.727)	(5.443)	(7.173)
FB gap × AEs	1.554	-3.322	1.819	3.237
	(1.500)	(2.257)	(1.511)	(2.513)
FB gap × EMMIEs	2.361**	5.323	2.460**	11.520***
	(1.040)	(1.527)	(1.044)	(1.783)
FB gap × LICs	-1.151	-0.796	-1.057	3.579*
	(1.610)	(2.003)	(1.613)	(2.163)
Debt-to-GDP × AEs	0.496*	-0.040	0.527*	2.266***
	(0.267)	(0.689)	(0.269)	(0.773)
Debt-to-GDP × EMMIEs	2.101***	2.143***	2.119***	3.713***
	(0.320)	(0.631)	(0.321)	(0.776)
Debt-to-GDP × LICs	1.648***	2.296***	1.636***	2.427***
	(0.490)	(0.707)	(0.491)	(0.792)
AEs	-0.064	0.556	-0.089	-0.897
	(0.294)	(1.021)	(0.296)	(1.091)
EMMIEs	-0.002	1.578	-0.014	1.076
	(0.278)	(1.576)	(0.279)	(1.619)
τ_1	0.467**	0.871	0.718**	1.246
	(0.232)	(0.719)	(0.360)	(0.834)
τ_2	1.382***	1.989***	1.642***	2.544***
	(0.236)	(0.721)	(0.363)	(0.836)
N	2732	2568	2732	2568
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.163	0.170	0.168	0.253
Pseudo R^2 (Ugba & Gertheiss)	0.354	0.367	0.364	0.511
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.16. Robustness: Income Group Heterogeneity with Alternative Measures of Cyclical Conditions

Variables	(1)	(2)	(3)	(4)
Growth deviation from 5-year ahead WEO projection × AEs	15.531***	19.565***	15.958***	15.642***
	(3.697)	(4.269)	(3.729)	(5.141)
Growth deviation from 5-year ahead WEO projection × EMMIEs	12.136***	19.457***	12.107***	12.744***
	(2.835)	(3.656)	(2.878)	(4.258)
Growth deviation from 5-year ahead WEO projection × LICs	1.245	0.235	1.320	-4.572
	(4.165)	(4.786)	(4.211)	(5.378)
FB gap × AEs	3.248**	-0.492	3.536**	5.261**
	(1.487)	(2.152)	(1.499)	(2.445)
FB gap × EMMIEs	2.214**	6.577***	2.348**	12.883***
	(1.069)	(1.611)	(1.078)	(1.880)
FB gap × LICs	-1.109	-0.754	-0.976	3.258
	(1.606)	(2.023)	(1.612)	(2.201)
Debt-to-GDP × AEs	0.621**	0.361	0.659**	2.579***
	(0.267)	(0.662)	(0.270)	(0.756)
Debt-to-GDP × EMMIEs	2.491***	2.049***	2.519***	3.581***
	(0.331)	(0.639)	(0.333)	(0.773)
Debt-to-GDP × LICs	1.632***	2.277***	1.633***	2.361***
	(0.483)	(0.727)	(0.483)	(0.816)
AEs	-0.129	0.569	-0.154	-0.976
	(0.2933)	(1.022)	(0.296)	(1.089)
EMMIEs	-0.222	1.577	-0.234	1.044
	(0.279)	(1.583)	(0.281)	(1.619)
τ_1	0.478**	0.857	0.770**	1.396*
	(0.231)	(0.720)	(0.360)	(0.835)
τ_2	1.407***	1.985***	1.709***	2.721***
	(0.235)	(0.722)	(0.363)	(0.838)
N	2606	2454	2606	2454
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.195	0.170	0.202	0.257
Pseudo R^2 (Ugba & Gertheiss)	0.413	0.366	0.425	0.517
Proportional odds assumption holds	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.17. Robustness: Alternative Measures of Solvency

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Output gap	4.558***	5.516***	5.851***	7.150***	7.060***	8.683***	7.295***	7.051***
	(0.950)	(0.955)	(0.951)	(1.242)	(1.242)	(1.579)	(1.307)	(1.243)
FB gap	1.930***	2.654***						
	(0.740)	(0.745)						
Debt-to-GDP			1.088***	1.421***	1.416***	1.469***	1.357***	1.407***
			(0.181)	(0.222)	(0.223)	(0.241)	(0.227)	(0.230)
General government gross debt as % of average tax revenues	0.074***							
	(0.019)							
General government gross debt as % of 10-year moving average GDP		0.952***						
		(0.130)						
Sustainability gap, fiscal balance as % of GDP			1.796**					
			(0.862)					
Sustainability gap, primary balance, country-specific conditions, % of GDP				4.515***				
				(1.699)				
Sustainability gap, primary balance, historical conditions, % of GDP					4.538***			
					(1.674)			
Sustainability gap, primary balance, current conditions, % of GDP						3.762***		
						(1.352)		
Sustainability gap, primary balance, stressed conditions, % of GDP							8.238***	
							(1.533)	
Sustainability gap, primary balance, benign conditions, % of GDP								0.420
								(1.430)
τ_1	0.778***	0.359***	0.455***	0.144	0.158	0.050	-0.156	0.169
	(0.080)	(0.101)	(0.108)	(0.125)	(0.126)	(0.138)	(0.139)	(0.149)
τ_2	1.666***	1.260***	1.344***	1.061***	1.082***	0.970***	0.769***	1.081***
	(0.090)	(0.107)	(0.115)	(0.131)	(0.132)	(0.143)	(0.143)	(0.154)
N	2691	2713	2765	1861	1860	1580	1838	1837
Country FE	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No
Pseudo R^2 (McFadden)	0.137	0.143	0.132	0.379	0.381	0.455	0.393	0.383
Pseudo R^2 (Ugba & Gertheiss)	0.303	0.315	0.294	0.688	0.691	0.774	0.705	0.694
Proportional odds assumption holds	No	No	No	No	No	No	Yes	Yes

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.18. Robustness: Indicators of Liquidity and Market Access Measures						
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Output gap	5.935***	7.054***	2.527**	10.506***	5.624***	13.028***
	(1.053)	(1.047)	(1.068)	(2.128)	(1.370)	(3.015)
FB gap	-0.173	1.317*	1.791**	1.177	2.494***	7.863***
	(0.972)	(0.793)	(0.879)	(1.299)	(0.921)	(1.918)
Debt-to-GDP	0.843***	1.554***	1.928***	1.791***	1.279***	1.868***
	(0.241)	(0.203)	(0.254)	(0.345)	(0.210)	(0.322)
Gross financing needs	3.354***					
	(0.864)					
Effective nominal interest rate		9.884***				
		(2.310)				
Debt service as % of general government revenue			-0.334**			
			(0.138)			
Central government debt maturing in 12 months or less as % of GDP				-2.643*		
				(1.535)		
Foreign currency long-term sovereign debt ratings (index from 1 to 21 [best])					-0.089***	
					(0.012)	
5-year sovereign CDS spreads (basis points)						41.265***
						(7.353)
τ_1	0.240**	-0.289*	0.490***	-0.134	1.547***	-0.725***
	(0.118)	(0.158)	(0.142)	(0.177)	(0.209)	(0.208)
τ_2	1.109***	0.582***	1.494***	0.744***	2.420***	0.212***
	(0.124)	(0.160)	(0.152)	(0.182)	(0.215)	(0.209)
N	2211	2284	2009	931	2011	867
Country FE	No	No	No	No	No	No
Year FE	No	No	No	No	No	No
Pseudo R^2 (McFadden)	0.301	0.274	0.424	0.668	0.343	0.694
Pseudo R^2 (Ugba & Gertheiss)	0.585	0.544	0.741	0.933	0.647	0.945
Proportional odds assumption holds	No	No	No	No	Yes	No

Sources: WEO; Kose and others (2022); DSA/DSF database; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.19. Robustness: Indicators of Balance Sheet Vulnerabilities					
	(1)	(2)	(3)	(4)	(5)
Output gap	9.342***	15.506***	9.442***	4.798***	11.777***
	(3.429)	(4.463)	(2.710)	(1.268)	(3.086)
FB gap	2.385	3.849**	1.522	1.250	0.641
	(2.558)	(1.675)	(1.596)	(1.188)	(1.436)
Debt-to-GDP	2.388***	0.835***	1.821***	2.464***	1.177***
	(0.452)	(0.284)	(0.306)	(0.326)	(0.324)
General government debt in foreign currency as % of total	0.583				
	(0.411)				
Debt securities held by nonresidents as % of total		0.795			
		(2.698)			
General government debt held by nonresidents as % of total			1.390***		
			(0.399)		
Concessional external debt stocks as % of external public debt				-0.246	
				(0.276)	
Sovereign debt average maturity in years at T					-0.044
					(0.032)
τ_1	-0.609**	0.039	-0.843***	-0.173	0.280
	(0.289)	(0.197)	(0.242)	(0.210)	(0.305)
τ_2	0.386	0.907***	0.053	0.842***	1.209***
	(0.293)	(0.203)	(0.242)	(0.216)	(0.310)
N	543	668	941	1453	696
Country FE	No	No	No	No	No
Year FE	No	No	No	No	No
Pseudo R^2 (McFadden)	0.811	0.742	0.648	0.572	0.737
Pseudo R^2 (Ugba & Gertheiss)	0.983	0.963	0.923	0.874	0.962
Proportional odds assumption holds	Yes	No	No	Yes	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.20. Robustness: Indicators of External and Private Sector Debt Vulnerabilities

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Output gap	5.876***	15.394***	5.795***	2.486**	4.173***	6.822***	7.176***	7.085***
	(1.068)	(4.542)	(1.179)	(1.061)	(1.061)	(1.120)	(1.147)	(1.146)
FB gap	1.402	5.612**	2.020**	2.330***	1.068	2.135**	2.040**	2.034**
	(0.898)	(2.708)	(0.971)	(0.786)	(0.913)	(0.915)	(0.927)	(0.922)
Debt-to-GDP	1.643***	2.218***	1.471***	1.599***	1.527***	1.369***	1.362***	1.383***
	(0.220)	(0.530)	(0.227)	(0.195)	(0.211)	(0.216)	(0.218)	(0.219)
Total external debt stocks as % of GDP	-0.187***							
	(0.046)							
External debt in foreign currency as % of total		1.673***						
		(0.427)						
Private external debt stocks as % of GDP			-0.160***					
			(0.043)					
Domestic credit to private sector as % of GDP				-0.755***				
				(0.083)				
Short-term external debt stocks as % of total					-1.610***			
					(0.344)			
Short-term external debt stocks as % of reserves						-0.001		
						(0.004)		
Total external debt stocks as % of reserves							0.0006	
							(0.001)	
Total external debt stocks as % of reserves excl. gold								-0.0001
								(0.0007)
τ_1	0.258**	-1.777***	0.251**	0.831***	0.548***	0.223*	0.203	0.208*
	(0.121)	(0.487)	(0.127)	(0.123)	(0.142)	(0.124)	(0.124)	(0.124)
τ_2	1.153***	-0.897*	1.127***	1.760***	1.449***	1.106***	1.083***	1.087***
	(0.127)	(0.481)	(0.133)	(0.130)	(0.148)	(0.129)	(0.130)	(0.130)
N	2216	465	1852	2567	2233	2020	2014	2015
Country FE	No	No	No	No	No	No	No	No
Year FE	No	No	No	No	No	No	No	No
Pseudo R^2 (McFadden)	0.290	0.818	0.380	0.203	0.380	0.342	0.344	0.343
Pseudo R^2 (Ugba & Gertheiss)	0.568	0.984	0.690	0.427	0.691	0.641	0.644	0.643
Proportional odds assumption holds	No	Yes	No	No	No	No	No	No

Sources: WEO; Kose and others (2022); IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

Table AV.21. Interaction with Vulnerability Exercise Assessment for Fiscal at T				
Variables	(1)	(2)	(3)	(4)
Output gap × Low	6.337***	9.361***	6.355***	6.483***
	(1.433)	(2.212)	(1.440)	(2.391)
Output gap × Medium	2.204	12.899***	2.052	9.129***
	(1.666)	(2.631)	(1.684)	(2.901)
Output gap × High	-0.418	9.513**	0.045	6.594
	(2.938)	(3.853)	(2.938)	(4.220)
FB gap × Low	3.339***	4.084***	3.423***	7.926***
	(0.973)	(1.438)	(0.983)	(1.688)
FB gap × Medium	1.439	4.167**	1.642	9.031***
	(1.508)	(1.907)	(1.512)	(2.080)
FB gap × High	-1.097	0.324	-0.871	4.927
	(2.409)	(3.012)	(2.381)	(3.096)
Debt-to-GDP × Low	0.790***	2.139***	0.807***	3.667***
	(0.222)	(0.514)	(0.224)	(0.608)
Debt-to-GDP × Medium	1.998***	1.808***	2.018***	2.477***
	(0.436)	(0.605)	(0.437)	(0.702)
Debt-to-GDP × High	0.777	1.764**	0.826	2.081**
	(0.608)	(0.772)	(0.614)	(0.838)
Medium	0.064	0.588	0.072	1.315***
	(0.276)	(0.403)	(0.278)	(0.444)
High	1.087**	1.080**	1.045**	1.480**
	(0.432)	(0.545)	(0.436)	(0.596)
τ_1	0.388***	2.305	0.616**	2.215
	(0.124)	(1.411)	(0.297)	(1.472)
τ_2	1.297***	3.410**	1.532***	3.510**
	(0.129)	(1.412)	(0.299)	(1.473)
N	2707	2543	2707	2543
Country FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Pseudo R^2 (McFadden)	0.164	0.168	0.169	0.256
Pseudo R^2 (Ugba & Gertheiss)	0.355	0.463	0.364	0.515
Proportional odds assumption holds	No	No	Yes	No

Sources: WEO; Kose and others (2022); Vulnerability Exercise (VE) assessment; IEO staff calculations.
Note: Standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01. All regressions use contemporaneous data.

REFERENCES

Alonso, Virginia and others, 2025a, "Fiscal Rules Dataset: 1985-2024," (Washington: International Monetary Fund).

Alonso, Virginia and others, 2025b, "Fiscal Councils Dataset: The 2024 Update," (Washington: International Monetary Fund).

International Monetary Fund, 2018, "Assessing Fiscal Space—An Update and Stocktaking," April (Washington).

Kilic Celik, Sinem, Ayhan M. Kose, Franziska Ohnsorge, and F. Ulrich Ruch, 2023, "Potential Growth: A Global Database," Policy Research Working Paper 10354, (Washington: World Bank).

Kose, Ayhan M., Segio Kurlat, Franziska Ohnsorge, and Naotaka Sugawara, 2022, "A cross-country database of fiscal space," Journal of International Money and Finance, 128(November): 102682.

Lang, Valentin, Lukas Wellner, and Alexandros Kentikelenis, 2024, "Biased bureaucrats and the policies of international organizations." American Journal of Political Science 1–19.

Liu, Dungang, and Heping Zhang, 2017, "Residuals and Diagnostics for Ordinal Regression Models: A Surrogate Approach," Journal of the American Statistical Association, 113 (522), 845-854.