

IEO

Independent Evaluation Office
of the International Monetary Fund

Reading the Fund: Analyzing IMF Fiscal Advice Using LLMs

Jantsankhorloo Amgalan, Jérémie Cohen-Setton, and Anton Korinek

Disclaimer: The views expressed in this presentation are those of the authors and do not necessarily represent those of Anthropic, the IEO, the IMF, or IMF policy.

NBER International Seminar on Macroeconomics | June 24, 2026

This Paper (1/2)

- ▶ **Background Paper for the Evaluation of IMF Advice on Fiscal Policy**
 - IEO [background paper](#) and [online appendix](#) available on ieo.imf.org
 - IMF Working Paper version and dataset expected in summer / fall 2026
- ▶ **Construct the first large-scale dataset of IMF fiscal advice**
 - Bilateral advice extracted from 5,600+ Article IV reports (1978-2025)
 - Fiscal stance recommended by IMF staff (for the near-term and medium-term)

This Paper (2/2)

- ▶ **Assemble a contemporaneous macro-fiscal dataset**
 - Skip information reported in Article IV appendix tables
 - Instead, collect indicators available at the time of each consultation cycle using internal and public sources
- ▶ **Estimate fiscal advice reaction functions**
 - Quantify how macroeconomic conditions and fiscal fundamentals shape recommended fiscal stance across income groups
 - Document the evolution of IMF advice over time

Broader Relevance (Anton's Personal Thoughts)

1. Template for AI-assisted policy evaluation

2. Automation of the policy workflow:


- ▶ AI can automatically evaluate policy
- ▶ ...
- ▶ AI may soon be able to automatically design policy
– subject to human judgement
- ▶ ...
- ▶ AI may soon be able to automatically judge policy

Outline of this paper

- 1. Dataset construction**
- 2. Empirical approach**
- 3. Results**

Fiscal Stance: Concept and Challenges

- **Objective:** Extract the direction of change in discretionary fiscal policy
- **Interpretation:** Terms like *loosening / stimulus* or *tightening / consolidation* are taken as signals of changes in discretionary fiscal policy
- **Benchmark:** Ideally corresponds to changes in the cyclically adjusted (structural) primary balance
- **Challenge:** Article IV Staff Reports lack a consistent definition of “fiscal stance,” so usage can be ambiguous


INTERNATIONAL MONETARY FUND

IMF Country Report No. 24/232

UNITED STATES

July 2024

2024 ARTICLE IV CONSULTATION—PRESS RELEASE; STAFF REPORT; AND STATEMENT BY THE EXECUTIVE DIRECTOR FOR THE UNITED STATES

Under Article IV of the IMF’s Articles of Agreement, the IMF holds bilateral discussions with members, usually every year. In the context of the 2024 Article IV consultation with the United States, the following documents have been released and are included in this package:

- A **Press Release** summarizing the views of the Executive Board as expressed during its July 17, 2024, consideration of the staff report that concluded the Article IV consultation with the United States.
- The **Staff Report** prepared by a staff team of the IMF for the Executive Board’s consideration on July 17, 2024, following discussions that ended on June 13, 2024, with the officials of the United States on economic developments and policies. Based on information available at the time of these discussions, the staff report was completed on July 1, 2024.
- An **Informational Annex** prepared by the IMF staff.
- A **Statement by the Executive Director** for the United States.

The IMF’s transparency policy allows for the deletion of market-sensitive information and premature disclosure of the authorities’ policy intentions in published staff reports and other documents.

Copies of this report are available to the public from

International Monetary Fund • Publication Services
PO Box 92780 • Washington, D.C. 20090
Telephone: (202) 623-7430 • Fax: (202) 623-7201
E-mail: publications@imf.org Web: <http://www.imf.org>

International Monetary Fund
Washington, D.C.

Preparing the corpus of Article IV reports

Pre-processing strategy

- ▶ Objective: select relevant parts of the document, removing or separating extra material (figures, ...), fixing formatting issues
- ▶ All AIV reports are converted from PDF to JSON/Markdown
 - Each format has (dis)advantages
 - Images/charts are excluded
 - Tables and annexes retained
- ▶ Extract ToC with page numbers
 - Makes it easy to access annexes and tables individually
 - Helpful in identifying where Staff Appraisal begins and ends

Preparing the corpus of Article IV reports

Build structured dataset of report text

- ▶ Split the document pre- and post-staff appraisal sections
 - Pre sections are policy discussions
 - Post sections are annexes with useful information
- ▶ Feed the sections before Staff Appraisal (inclusive) to LLMs
 - Bucket each paragraph and box into policy themes (e.g., fiscal, monetary, exchange rate, financial, structural, social, etc.)
 - Verbatim extraction with strict separation of voices
 - Bucket authorities' responses to policy themes as well

Extracting and Classifying Fund Advice from Article IVs

► Classification Strategy

- **Task:** Classify IMF advice on policy stance as *loosen*, *neutral*, *tighten*, or *unclear*.
- **Method:**
 - Select fiscal policy related paragraphs + Staff Appraisal section.
 - Models reason step-by-step: → justification → certainty score → classification.
 - Use multiple LLMs: OpenAI's GPT-5.4, Claude Opus 4.6, Gemini-3-Pro.
 - Majority voting across models for the final classification.
- **Manual validation** of 105 country-year pairs:
 - 86% match rate for near-term and medium-term fiscal stance classifications.

Extracting and Classifying Fund Advice from Article IVs

▶ Prompt

You will be given excerpts from an IMF Article IV report, which may include the policy discussion and/or the staff appraisal sections.

Based on the provided text, your task is to identify and analyze the IMF staff's advice on fiscal policies for both the near-term and medium-term.

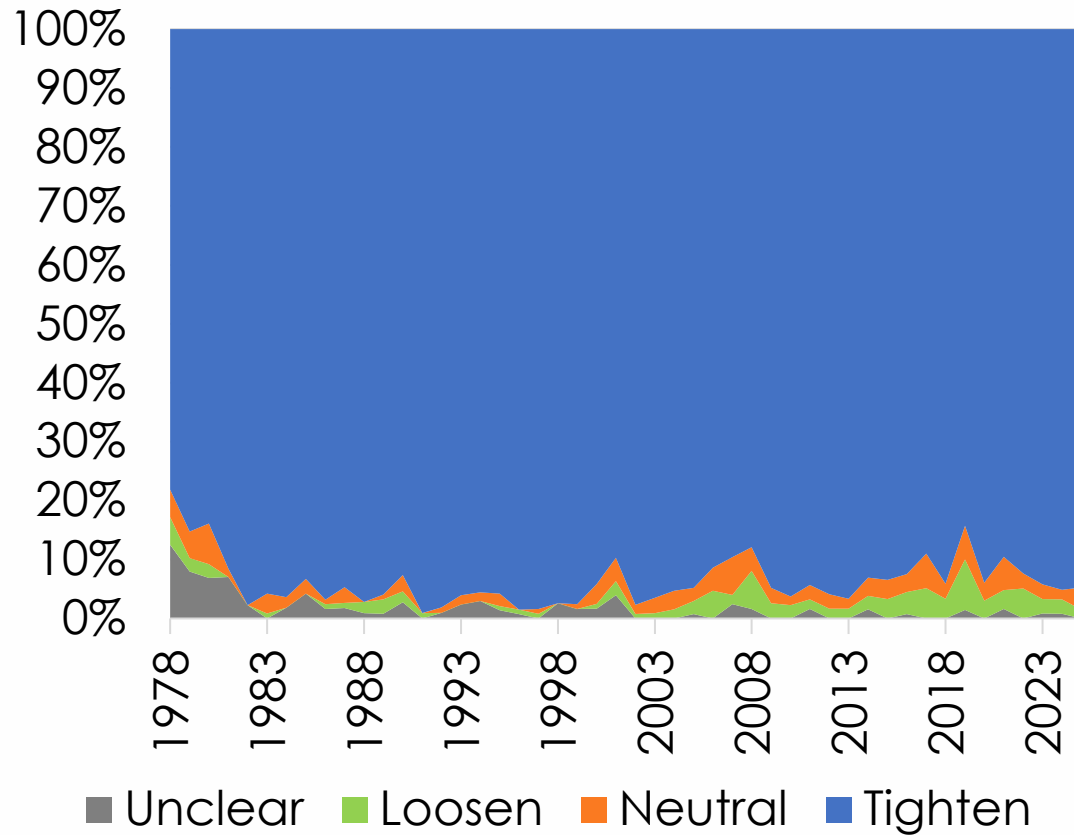
Near-term refers to...

Medium-term generally refers to...

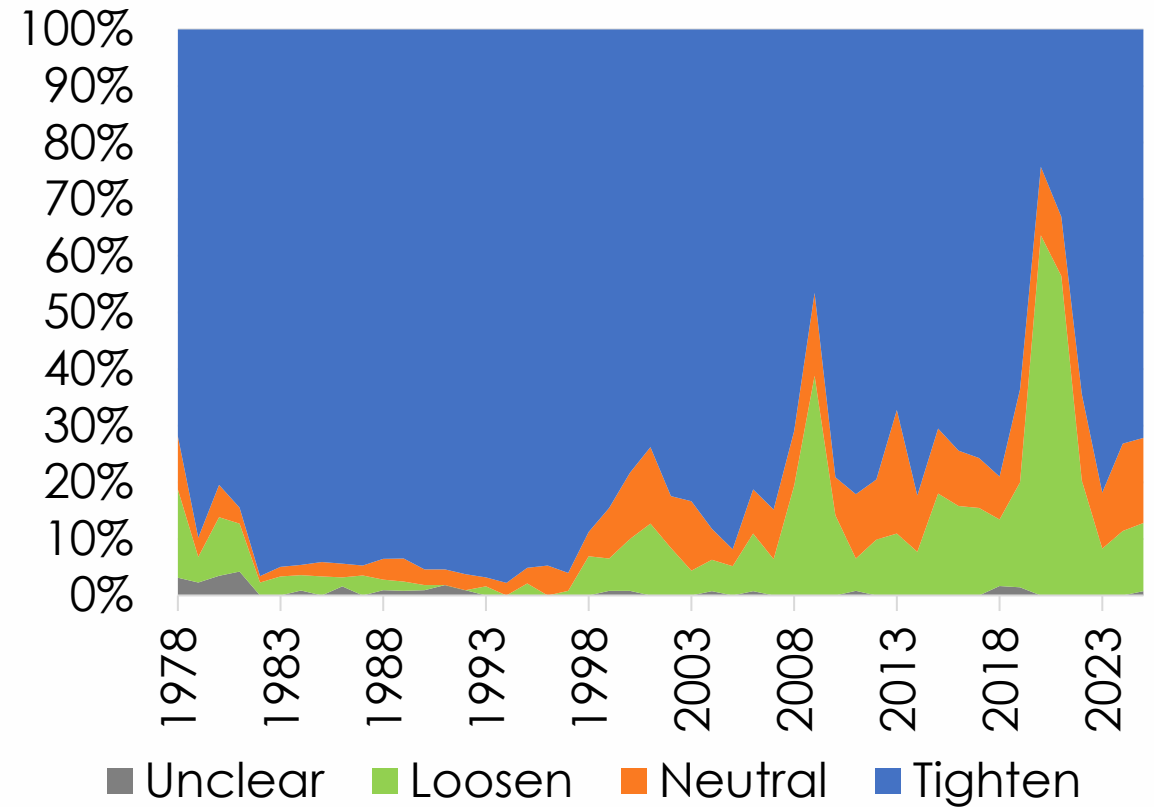
...

Preview of Results

Fiscal Stance Advice, Medium-Term
(In percent)



Fiscal Stance Advice, Near-Term
(In percent)



Macro and Fiscal Dataset

- ▶ **Objective:** approximate the macroeconomic information available to staff at the time of each Article IV
- ▶ **Ideally:** use staff macro-framework associated with each consultation cycle
 - In practice, difficult to collect systematically due to absence of a centralized repository until recently
 - Extracting data from appendix tables is feasible, but a separate project
- ▶ **Baseline approach:** use data vintages available at the time of each Article IV when possible. Supplement with fixed-vintage historical data when needed.

Macro and Fiscal Dataset

▶ **Additional Data Sources:**

- Financing and fiscal risk indicators: GFN (LIC-DSF, MAC-DSA), IMF Vulnerability Indicators; Fiscal space measures: Kose et al. (2022)
- Policy and institutional variables: exchange rate regimes (AREAER), fiscal Institutions (Alonso et al., 2025)
- Staff characteristics: mission chief background (Lang et al., 2024)

IMF Fiscal Policy Objectives

- ▶ **Maintain fiscal sustainability**

- Ensure solvency (sustainable debt dynamics) and liquidity (ability to meet financing needs)

- ▶ **Output Stabilization**

- Support countercyclical policy over the cycle
- Ensure consistency with monetary, financial, and external policies

- ▶ **Long-term growth and public goods**

- Promote potential output through investment in infrastructure, human capital, and institutions
- Support provision of public goods (e.g., social protection, climate, education)

Literature

▶ **Output stabilization vs. fiscal sustainability:**

- **Micro-founded models:** combine Keynesian stabilization motives with sovereign risk (Bianchi, Ottonello, Presno 2023)
- **Welfare-based policy frameworks:** forward-looking governments balance short-run output stabilization and long-term consequences of debt (Fournier 2019; Fournier and Lieberknecht 2020)
- **Operational rules:** Policy based on fixed weights for stabilization and sustainability objectives (Kanda 2011; Carnot 2014; European Commission 2015; ECB 2016; Bankowsky and Ferdinandusse 2017)

▶ **Long-term investment vs. fiscal sustainability:**

- **Debt-Investment-Growth (DIG) models:** Analyze how public investment supports growth while affecting debt dynamics (Buffie, Berg, Pattillo, Portillo, and Zanna 2012 and extensions)

Motivating the Empirical Specification

- ▶ **Ideal Approach: Model the full range of drivers of fiscal stance advice**
 - Whether financing constraints are binding
 - Extent of liquidity and solvency concerns
 - Amount of slack in the economy
 - Magnitude of medium-term spending pressures
- ▶ **Practical Limitation:**
 - Data gaps make it difficult to capture all dimensions
 - Trade-off between breadth and depth
- ▶ **Our Approach:**
 - Start with a parsimonious specification
 - Emphasize the core trade-off between fiscal sustainability and output stabilization
 - Explore additional factors or alternative dimensions through targeted, ad-hoc extensions

Empirical Model: Near-Term Advice

- ▶ We estimate the following near-term fiscal advice reaction function

$$\begin{aligned}
 & IMF\ Advice_{i,t}^* \\
 & = \beta Outputgap_{i,t} + \gamma Fiscalgap_{i,t-1} + \phi Debt_{i,t-1} + \boldsymbol{\theta}' \mathbf{X}_{i,t-1} + \varepsilon_{i,t}
 \end{aligned}$$

- ▶ **Outcome (latent):** $IMF\ Advice_{i,t}^*$ is a latent variable capturing the recommended fiscal stance for the near-term, which is mapped onto the observed advice categories: Loosen, Maintain, Tighten

Empirical Model: Near-Term Advice

▶ Key regressors

- *Output gap* captures the amount of slack in the economy
- *Fiscal gap* measures the adjustment in the fiscal balance required to stabilize the debt-to-GDP ratio at its current level
- *Debt* is the debt-to-GDP ratio
- X is a vector of other explanatory variables

▶ **Fixed effects:** ε , the error term, can be decomposed to include country and year fixed effects ($\varepsilon_{i,t} = \alpha_i + \lambda_t + u_{i,t}$)

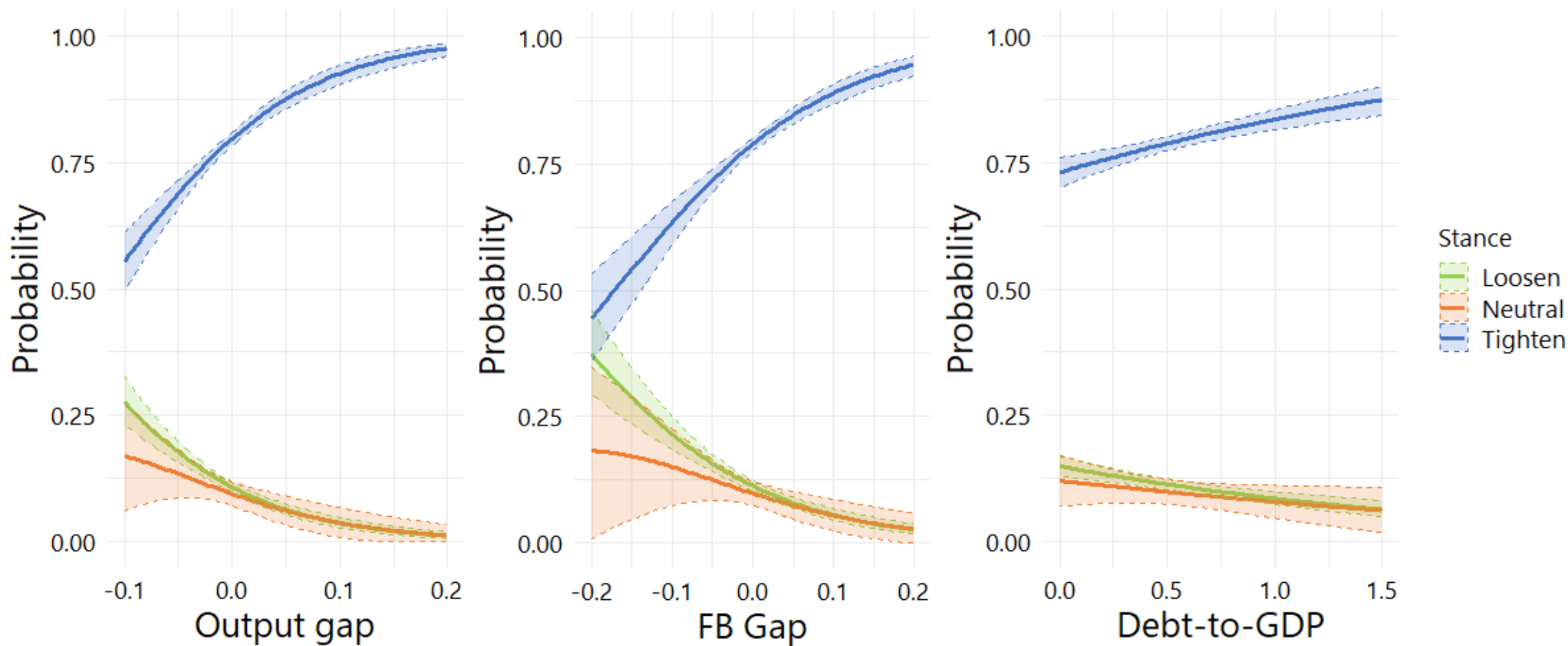
▶ **Heterogeneity:** interact key regressors with income group dummies (AE, EMMIE, LIC) to test differential advice

Baseline Results

Variables	(1)		(2)		(3)		(4)	
	AME	$\beta(se_{\beta})$	AME	$\beta(se_{\beta})$	AME	$\beta(se_{\beta})$	AME	$\beta(se_{\beta})$
Output gap	1.870	11.454*** (1.224)	2.291	15.594*** (1.596)	1.892	11.722*** (1.242)	1.287	10.707*** (1.871)
FB gap	1.254	7.680*** (0.896)	1.562	10.633*** (1.299)	1.254	7.774*** (0.906)	1.683	14.006*** (1.596)
Debt-to-GDP	0.102	0.625*** (0.124)	0.063	0.432*** (0.377)	0.097	0.603*** (0.125)	0.194	1.619*** (0.278)
N	3675		3434		3675		3434	
Country FE	No		Yes		No		Yes	
Year FE	No		No		Yes		Yes	

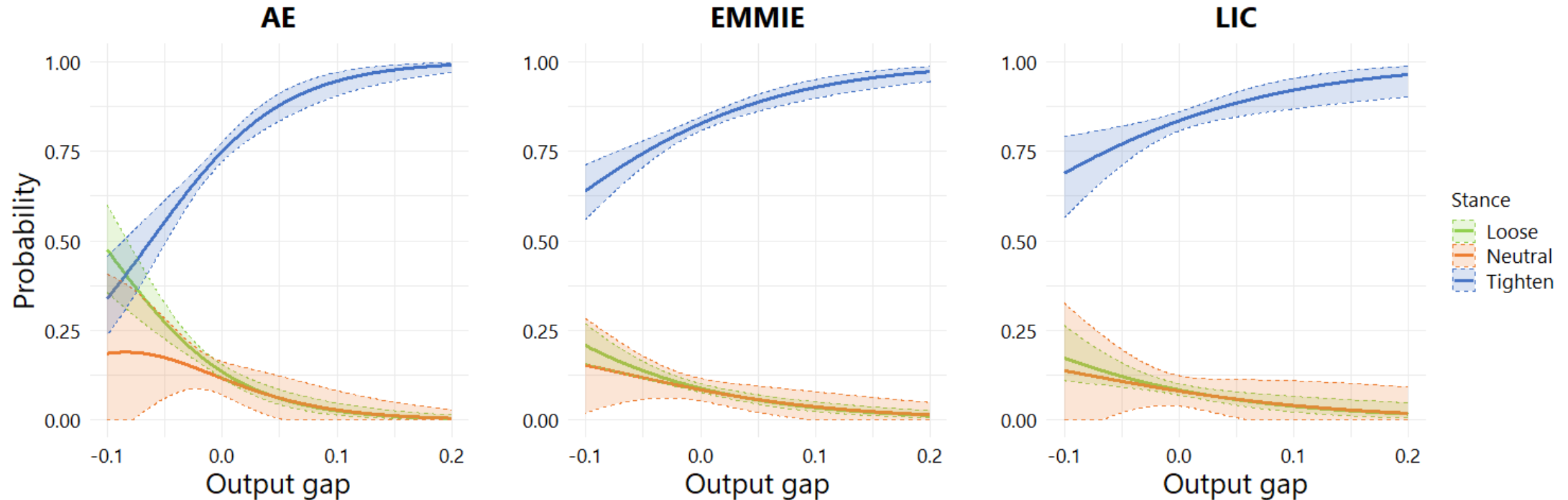
- The three variables are statistically significant
- We estimate the average marginal effect (AME) toward tightening advice

Baseline Results: Probability Plots



- ▶ **AME** (ppts): Output gap (1.87), FB gap (1.25), Debt-to-GDP (0.10).

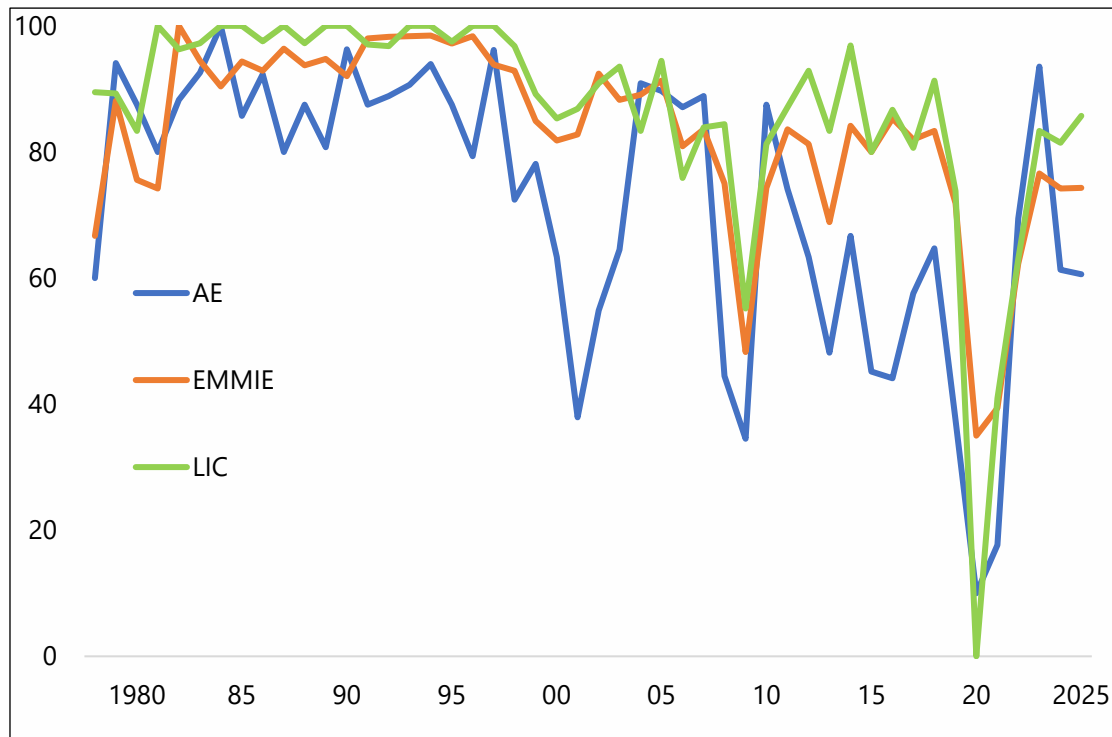
Results on Output Stabilization: by Income Group



- IMF fiscal advice is not procyclical
- Advice is more countercyclical for AEs than for EMs, but we obtain positive coefficients for both
- For LICs, it is mildly countercyclical
- ATE (pppts): AEs (2.8), EMs (1.6), LICs (1.3)

Result on Output Stabilization: Evolution

Fund Advice to Tighten Fiscal Stance in the Near Term (Percent of Total)



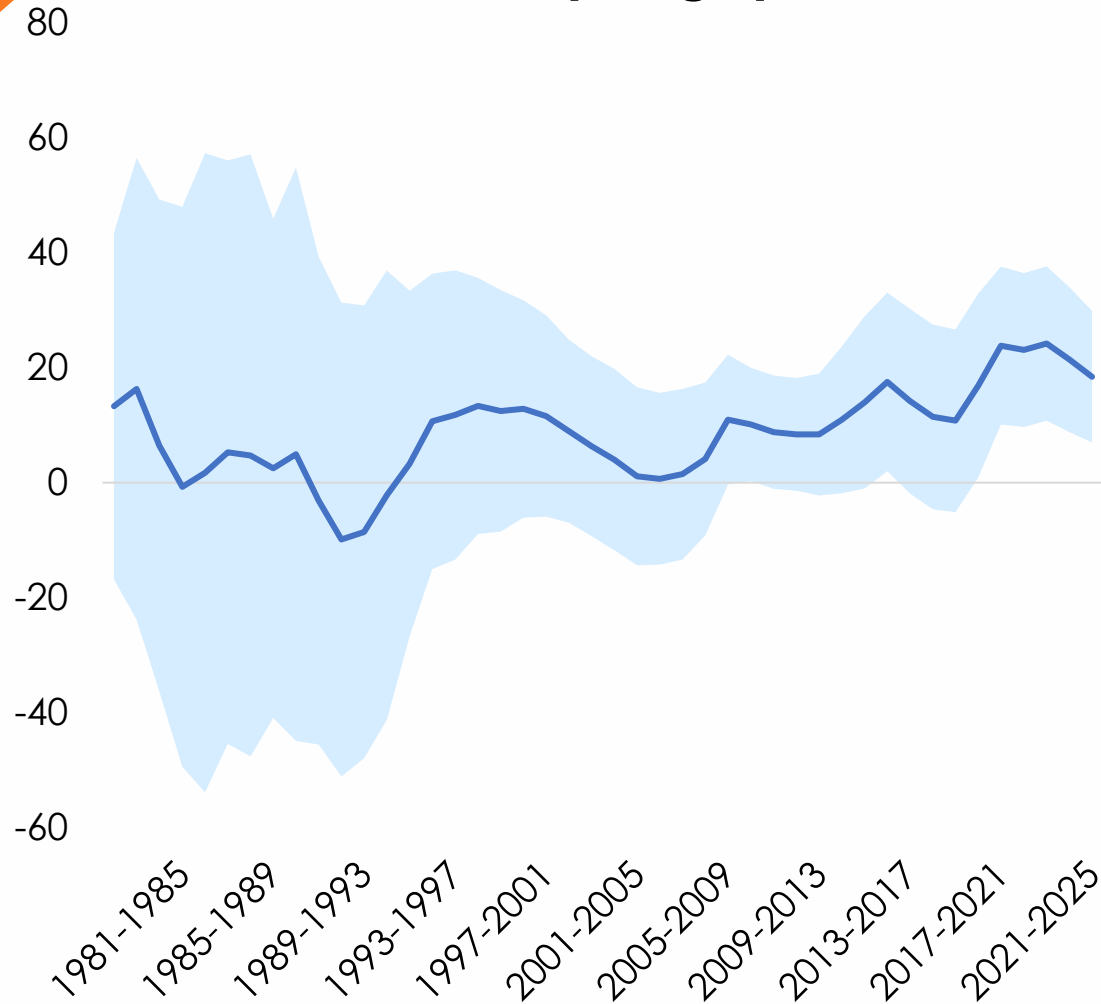
Source: [Amgalan, Cohen-Setton, and Korinek \(2025\)](#)

Note: The percentages are calculated as the share of reports recommending near-term tightening advice relative to the total number of Article IV reports published in that year for that income group

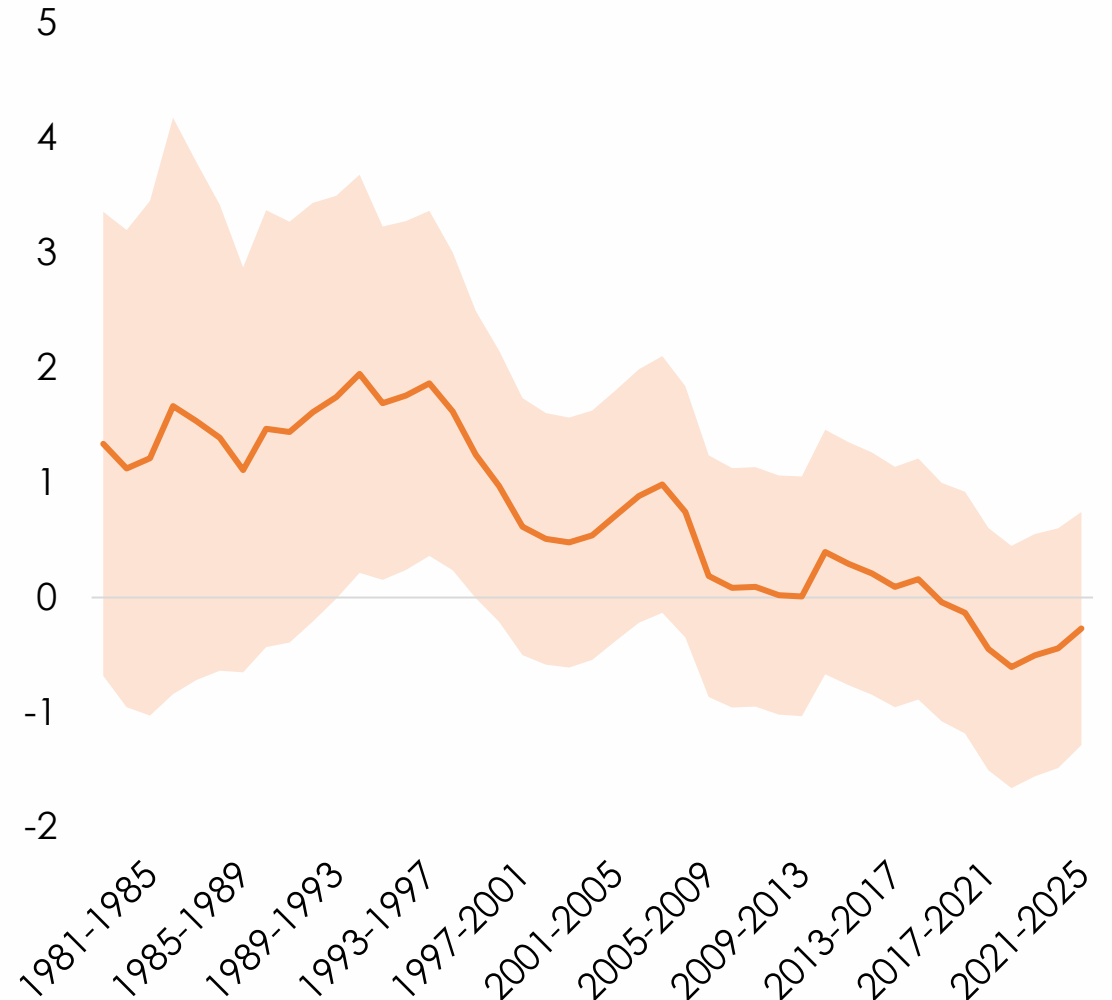
- ▶ Stronger countercyclicality in crises
 - Advice was consistently countercyclical across country groups in 2008-09 and 2020-21
 - Marks a clear shift from earlier episodes, when advice was often procyclical in less developed economies
- ▶ Fiscal suitability concerns dominate outside crisis periods, especially in less developed economies

Results on Output Stabilization: by Time Periods

Output gap



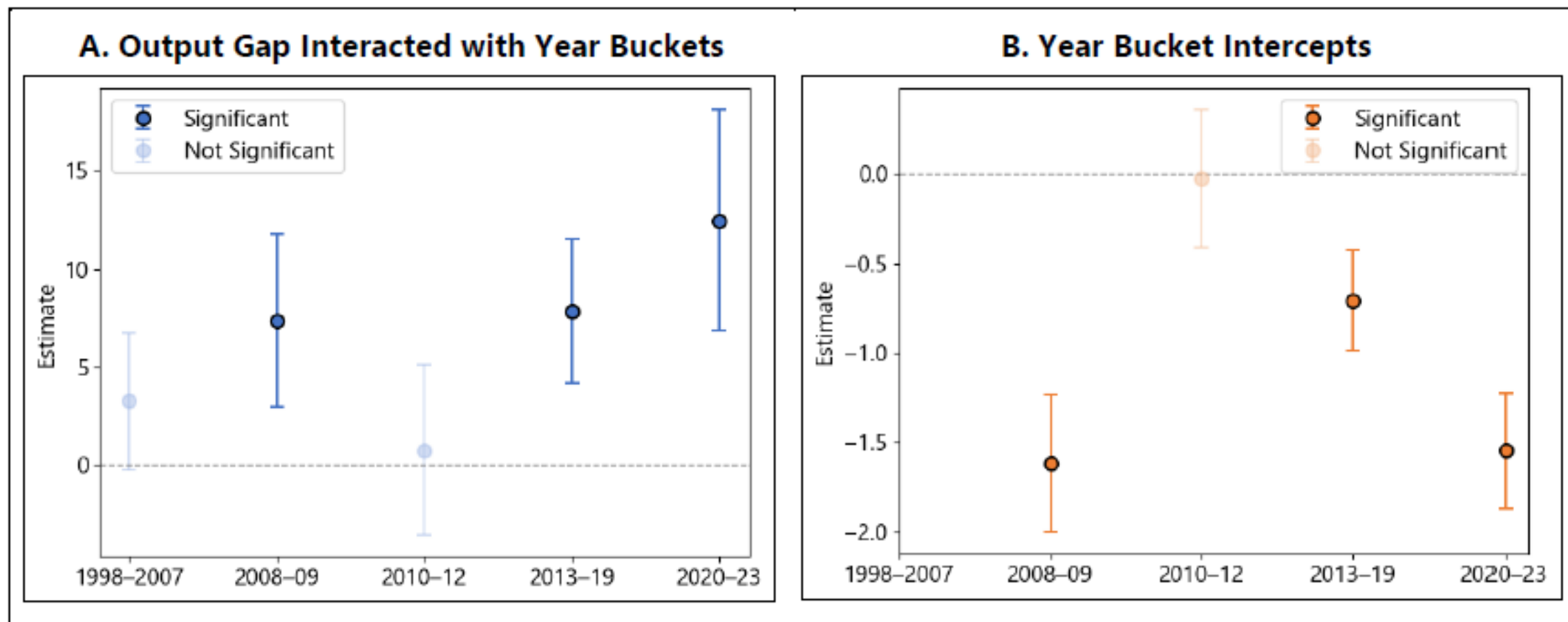
Year bucket constants



- Over time, advice has become more countercyclical

Results on Output Stabilization: by Time Periods

Figure 6. Evolution of the Countercyclicality of IMF Fiscal Advice



Output Stabilization: Robustness and Other Dimensions

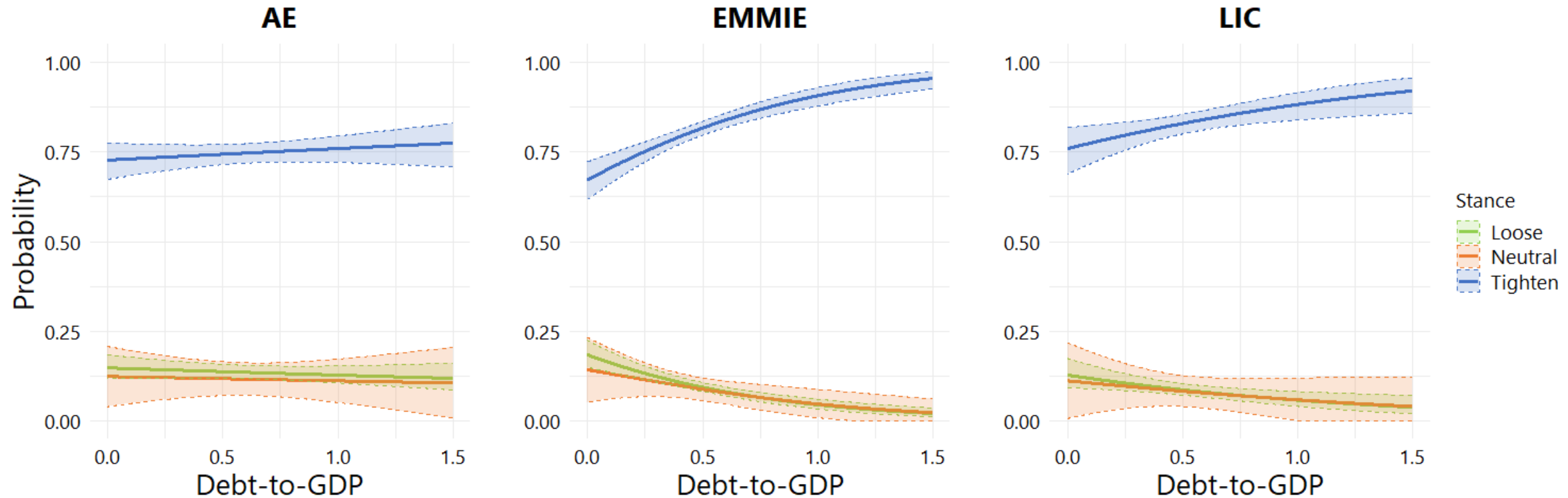
▶ Robustness Checks

- Alternative output gap measures: WEO estimates only; Hamilton-filter estimates only
- Growth measures: Growth relative to 5-year historical average; growth relative to 5-year-ahead WEO projections

▶ Additional Results

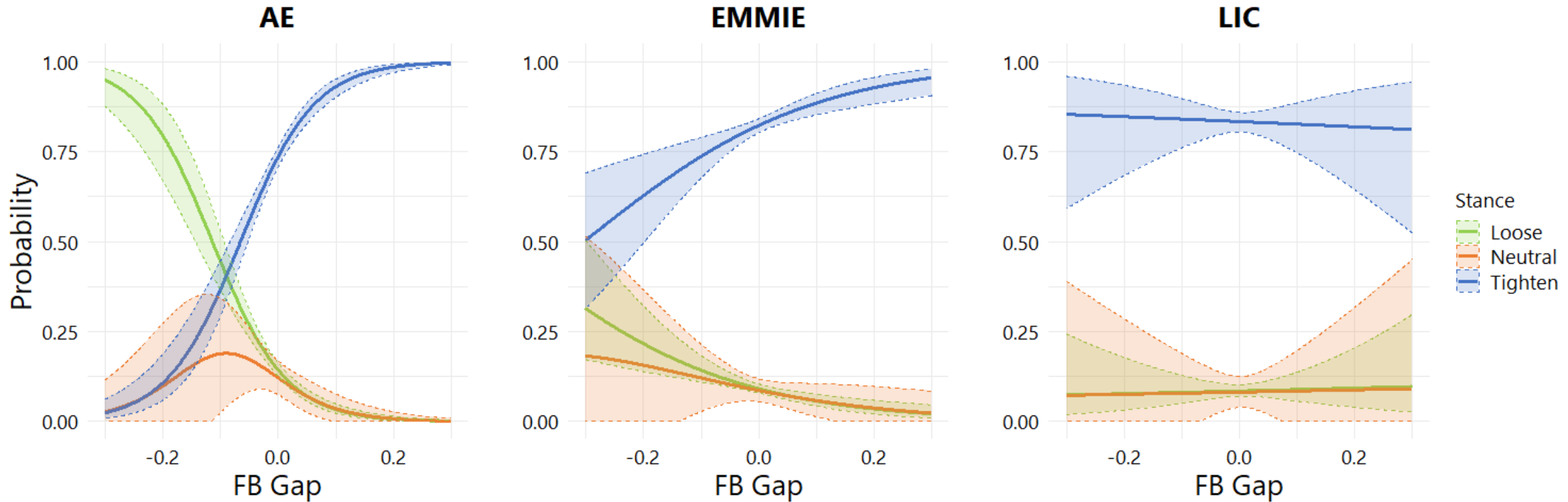
- **Fiscal Risks:** Advice is more countercyclical for countries with lower fiscal risks (as proxied by the Fund's Vulnerability Exercise)
- **Policy Mix:** Fiscal and monetary policies act as complements (not substitutes); no systematic link with macroprudential advice
- **External Imbalances:** CA deficits → tightening advice; CA surpluses → loosening advice; Asymmetry: stronger response to deficits than surpluses
- **Exchange Rate Regimes:** Advice more countercyclical under flexible ER regimes
- **Staff Background:** Mission Chiefs from “saltwater” schools give more countercyclical advice (Lang, Wellner & Kentikelenis, 2024)

Results on Fiscal Sustainability: Probability Plots of Debt-to-GDP by Income Group



- ▶ Nearly flat for AEs but steeper for EMs and LICs
- ▶ **AME** (ppts): AEs (0.03), EMs (0.24), LICs (0.13)

Results on Fiscal Sustainability: Probability Plots of Fiscal Gap by Income Group



- ▶ The fiscal gap is a more important explanatory variable for AEs
- ▶ **AME** (ppts): AEs (2.47), EMs (0.79), LICs (-0.08)

Results on Sustainability: Robustness and Other Dimensions

▶ **Robustness Checks**

- Alternative debt measures, alternative fiscal gap measures

▶ **Additional Results**

■ **Debt Risks:**

- Higher external debt-to-reserves ratio → tighter advice
- Weaker sovereign credit ratings → tighter advice

■ **Liquidity Measures:**

- Greater gross financing needs → tighter advice
- Higher effective interest rates → tighter advice
- Higher CDS spreads → tighter advice

■ **Balance Sheet Vulnerabilities:**

- Larger share of FX-denominated external debt
- Greater share of debt securities held by nonresidents

Near-Term Fiscal Advice

- ▶ Recommendations explicitly balance output stabilization and fiscal sustainability, with the relative weight evolving over the cycle
- ▶ Advice has become increasingly countercyclical over time, correcting earlier procyclical leanings
- ▶ Countercyclicity is strongest when monetary policy is constrained (zero lower bound), fiscal institutions are stronger, and staff graduated from salt-water universities

Medium-Term Fiscal Advice

- ▶ Persistent emphasis on consolidation to anchor debt sustainability
- ▶ Advice became more accommodative in advanced economies during secular stagnation

Outlook (Anton's Personal Thoughts)

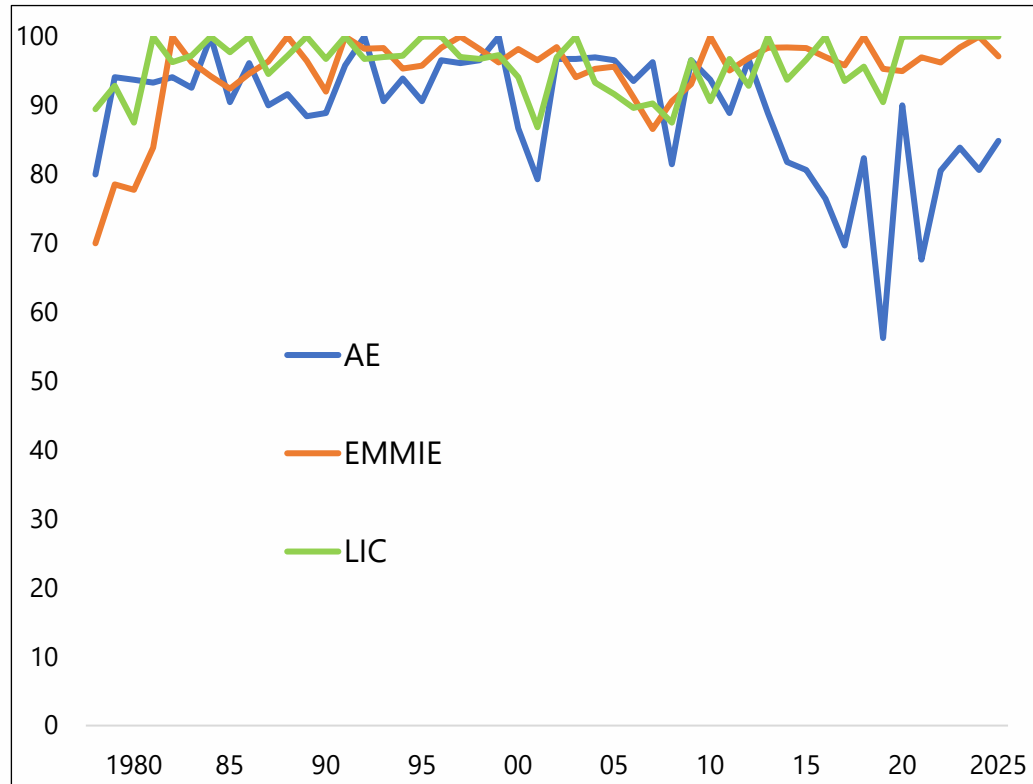
Near-Term Fiscal Advice

- ▶ AI Agents will soon be able to perform analyses like ours autonomously on their own
- ▶ This will be of tremendous value for decisionmakers
 - But also raises questions of AI oversight
 - Once AI systems good enough, will decisionmakers turn into AI rubber stampers?
- ▶ How can we design effective AI oversight systems?

THANK YOU!
Visit us at ieo.imf.org

Advice on Medium-Term Fiscal Stance

Advice to Tighten Fiscal Stance (Share of tightening, percent)



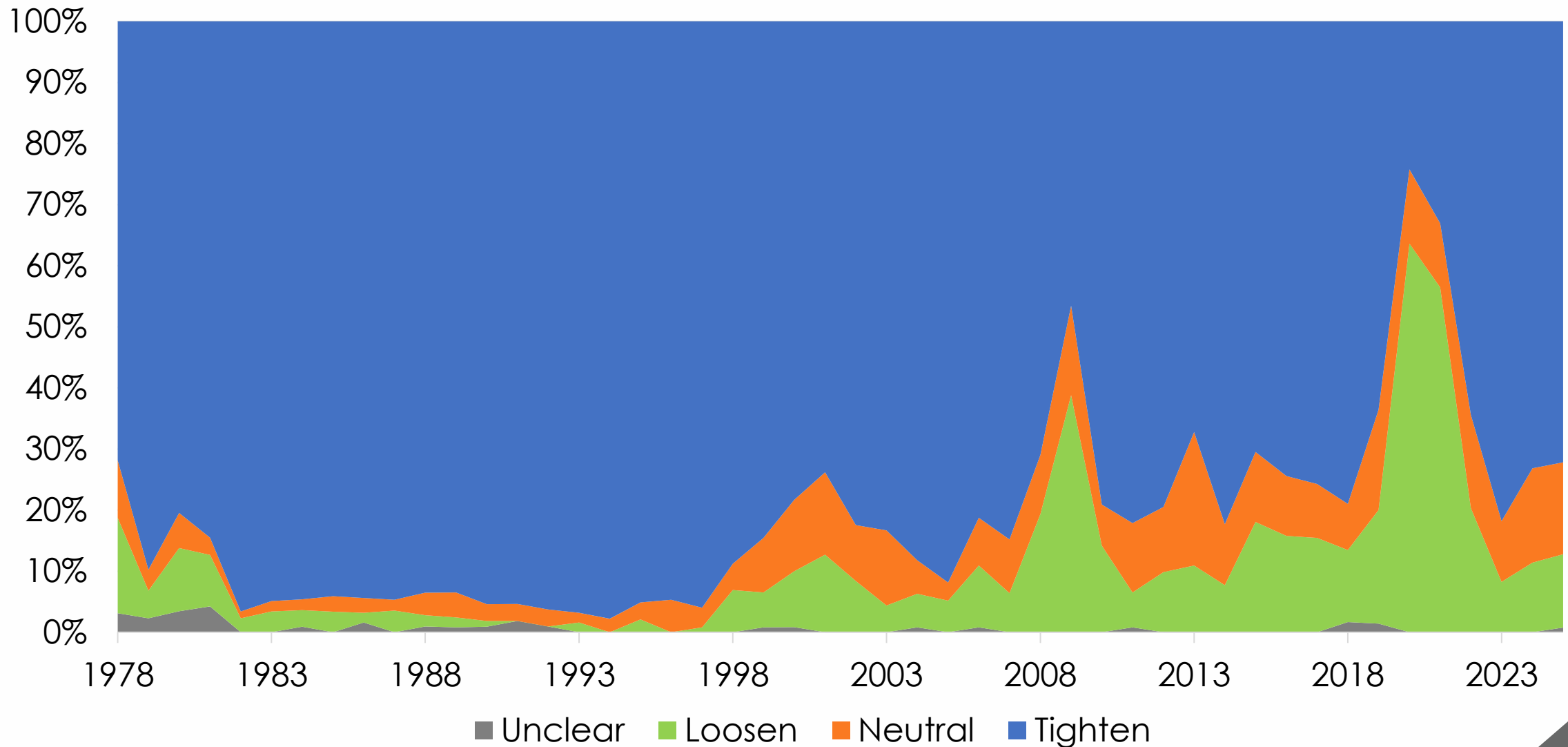
Source: [Amgalan, Cohen-Setton, and Korinek \(2025\)](#)

Note: The percentages are calculated as the share of reports recommending near-term (resp. .medium-term) tightening advice relative to the total number of Article IV reports published in that year for that income group

- ▶ Persistent focus on fiscal sustainability
 - Continued emphasis on consolidation to anchor fiscal sustainability
- ▶ Greater accommodation in AEs
 - For most countries, advice remained focused on gradual consolidation – through a mix of expenditure rationalization and revenue mobilization
 - Encouraged slower consolidation – and at times a neutral fiscal stance – to facilitate the financing of priority spending in countries with external surpluses and ample fiscal space
- ▶ Financing Long-Term Priorities in EMDEs
 - Emphasis on supporting investment through domestic adjustment
 - Mobilizing revenues and reprioritizing expenditures, rather than increasing debt

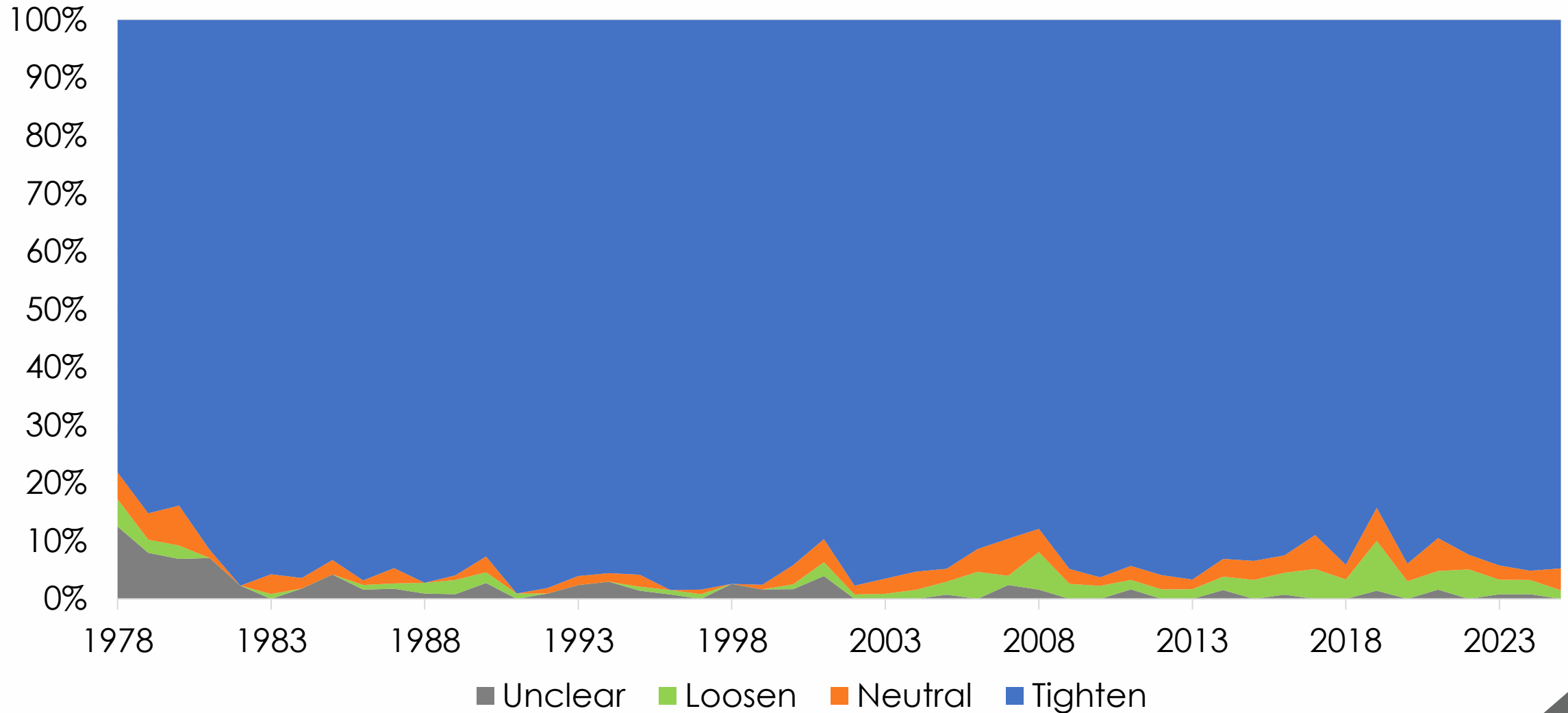
Preview of Results

Near-Term Advice on Fiscal Stance, 1978–2025



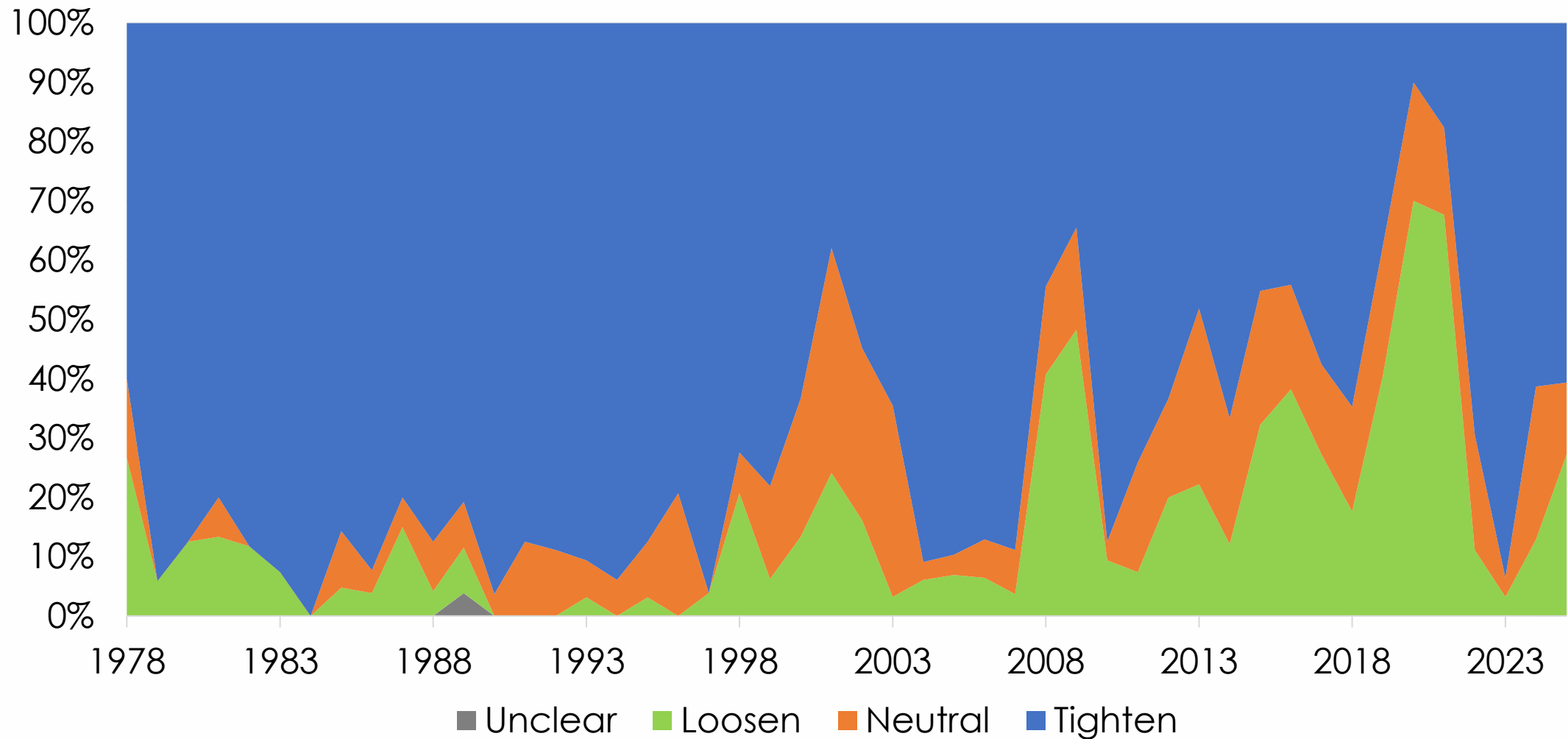
Preview of Results

Medium-Term Advice on Fiscal Stance, 1978–2025



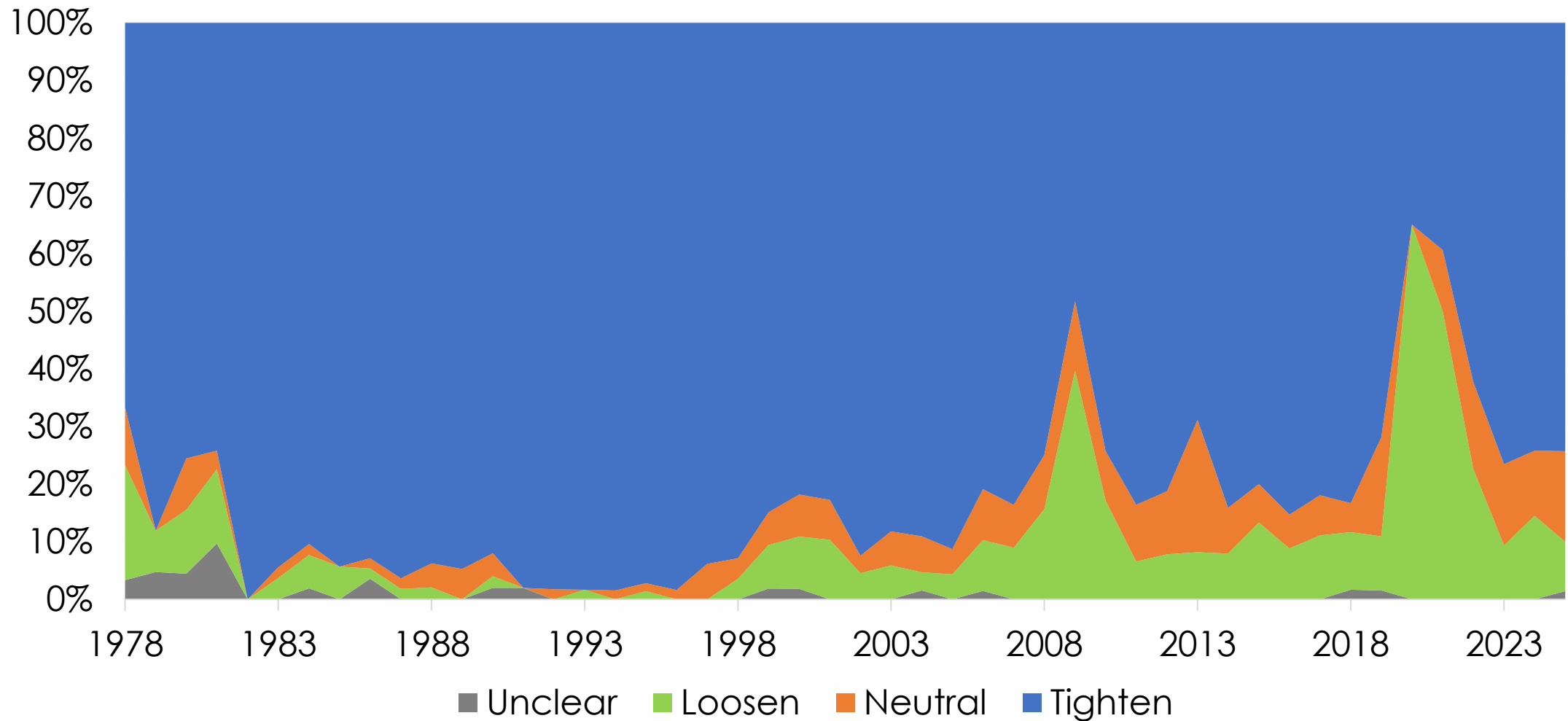
Preview of Results

Near-Term Advice on Fiscal Stance, AEs 1978–2025



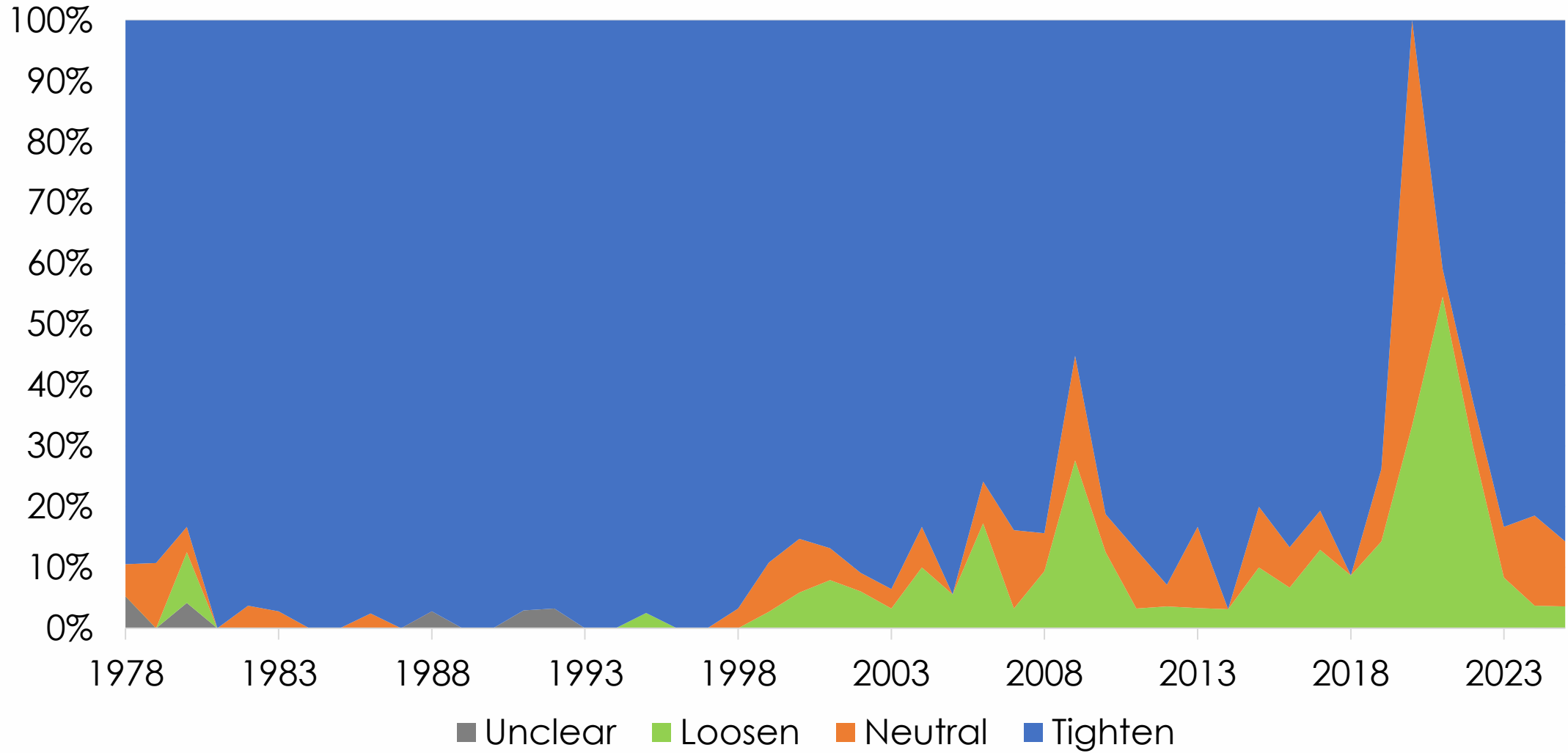
Preview of Results

Near-Term Advice on Fiscal Stance, EMMIEs 1978–2025



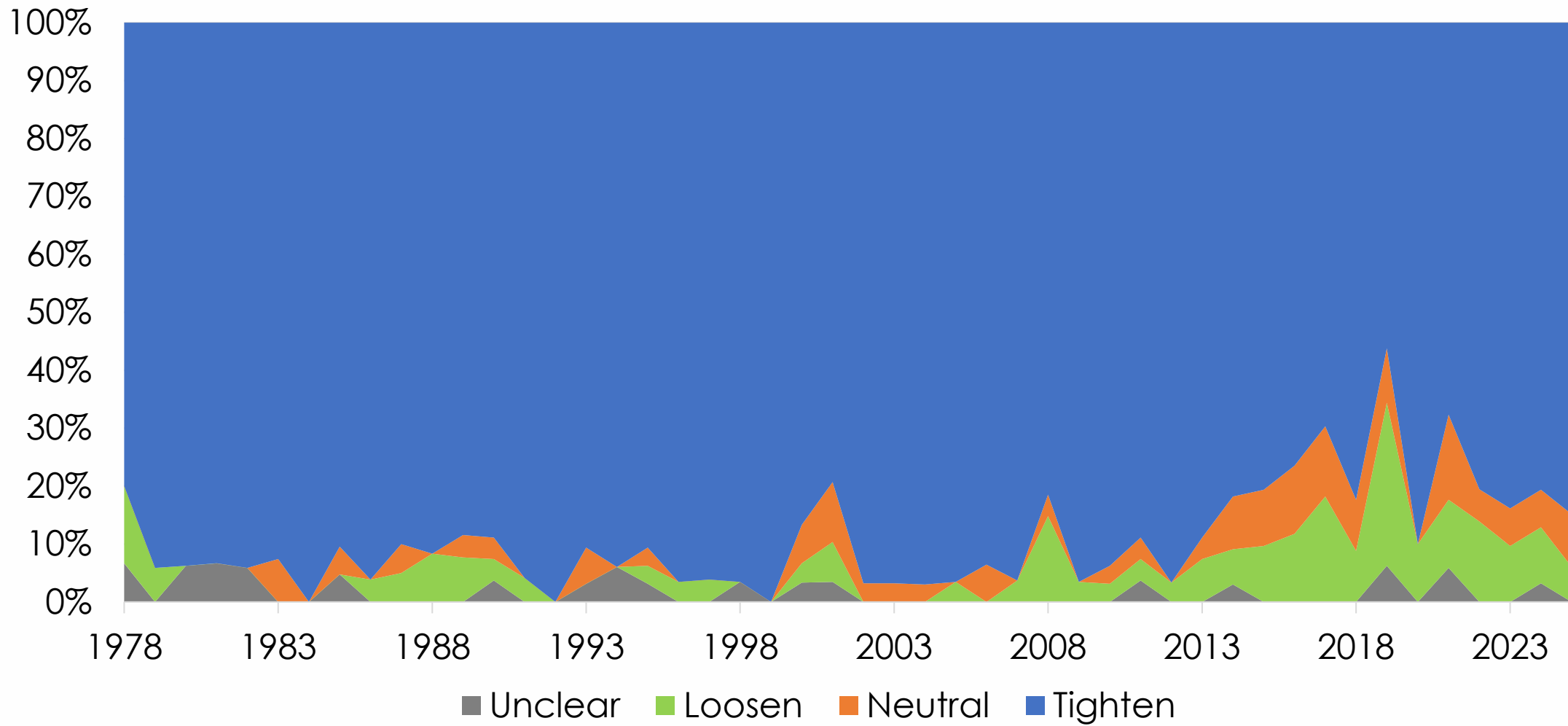
Preview of Results

Near-Term Advice on Fiscal Stance, LICs 1978–2025



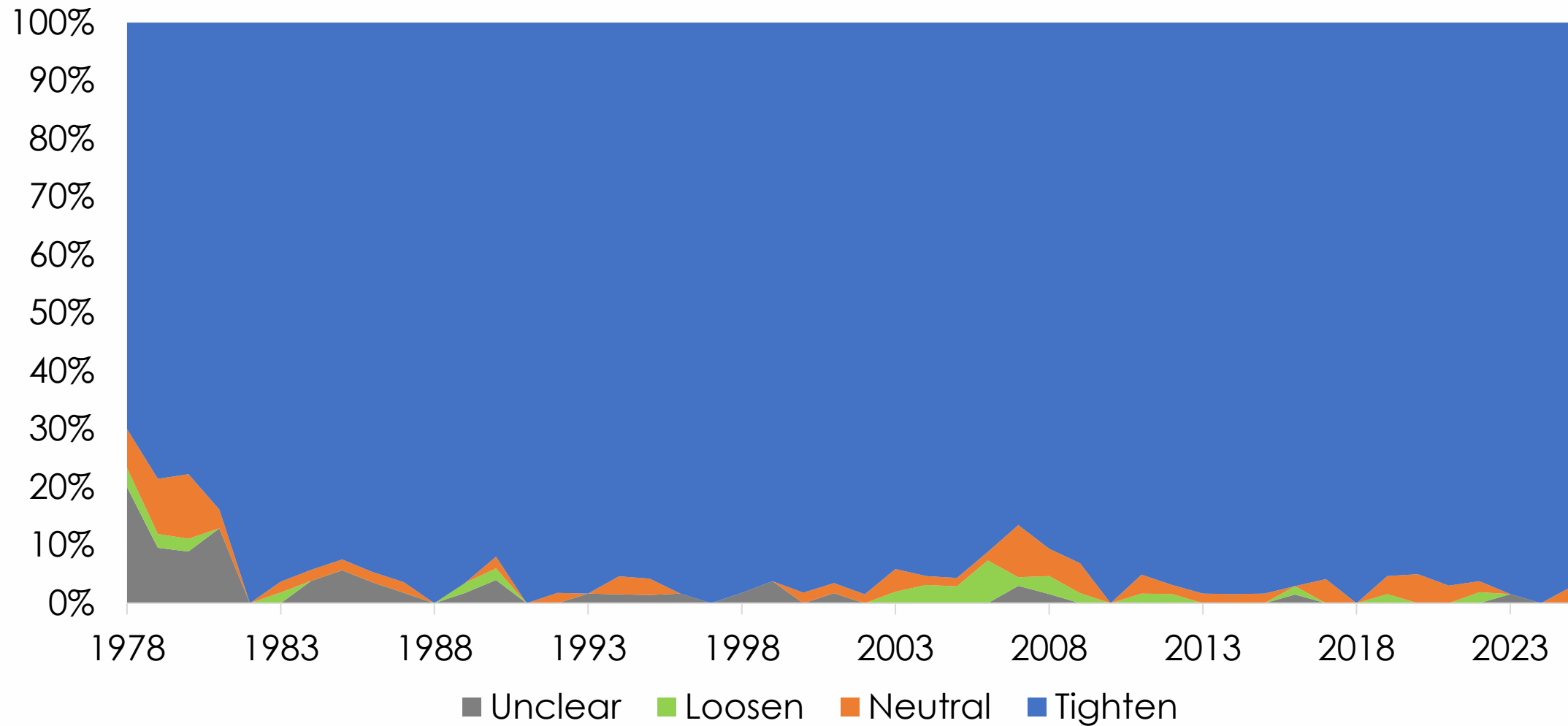
Preview of Results

Medium-Term Advice on Fiscal Stance, AEs 1978–2025



Preview of Results

Medium-Term Advice on Fiscal Stance, EMMIEs 1978–2025



Preview of Results

Medium-Term Advice on Fiscal Stance, LICs 1978–2025

