



SEPTEMBER 2026

REVIEW OF PROGRAM DESIGN AND CONDITIONALITY: INITIAL CONSIDERATIONS – BACKGROUND PAPER

IMF staff regularly produces papers proposing new IMF policies, exploring options for reform, or reviewing existing IMF policies and operations. The following document has been released and is included in this package:

- A **Staff Supplement**: Background paper on the Review of Program Design and Conditionality: Initial Considerations, prepared by IMF staff and completed on September 15, 2025.

The report prepared by IMF staff has benefited from comments and suggestions by Executive Directors following the informal session on October 6, 2025. Such informal sessions are used to brief Executive Directors on policy issues and to receive feedback from them in preparation for a formal consideration at a future date. No decisions are taken at these informal sessions. The views expressed in this paper are those of the IMF staff and do not necessarily represent the views of the IMF's Executive Board.

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.

Electronic copies of IMF Policy Papers
are available to the public from
<http://www.imf.org/external/pp/ppindex.aspx>

International Monetary Fund
Washington, D.C.



September 15, 2025

REVIEW OF PROGRAM DESIGN AND CONDITIONALITY: INITIAL CONSIDERATIONS – BACKGROUND PAPER

EXECUTIVE SUMMARY

This paper provides additional analytical background for Executive Directors' first informal discussion on the 2026 Review of Program Design and Conditionality (ROC). It covers the following topics:

- Program Success During 2002–24
- Realism of Growth Forecasts
- Assessing Fiscal Policy Conditionality
- Assessing Social Conditionality
- Public Debt and Forecast Accuracy
- Evolving Monetary Policy Conditionality
- Exchange Rate and External Adjustment
- Assessing Structural Conditionality
- Fragile and Conflict-Affected States (FCS) and Small Developing States (SDS)
- Follow-on Program Engagement
- Enterprise Risk Assessments in Use of Fund Resources (UFR)
- Stakeholder Survey: Preliminary Views

Approved By
Ceyla Pazarbasioglu
(SPR)

The Background Paper was prepared by Yazan Al-Karablieh, Juliana Araujo, Anil Ari, Rimtautas Bartkus, Anja Baum, Jocelyn Boussard, Thomas J. Carter, David Florian, Karina Garcia, Jaime Guajardo, Mahmoud Harb, Iacovos Ioannou, Mahamoud Islam, Fah Jirasavetakul, Robin Koepke, Lisa Kolovich, Tobias Krahne, Maxym Kryshko, Gabriel Leost, Giulio Lisi, Vivian Malta, Alexei Miksjuk, Sanghamitra Warrier Mukherjee, Monique Newiak, Liam O’Sullivan, Hyunmin Park, Gaelle Pierre, Faezeh Raei, Marina Rousset, Yulia Ustyugova, Agustin Velasquez, Peter Wankuru, and Irene Yackovlev with research assistance from Paula Arias, Luisa Ballen, Chen Chen, Yiran Chen, Juliana Dueñas Villabona, Maxwell Kushnir, Salvador Traettino, Joyce Saito, Holt Williamson, Lavinia Zhao, Usama Zafar, and Helen Zheng. Contributions were received from Kevin Bailey, Alex Grohovsky, Yuko Hashimoto, Alvar Kangur, Eric Pondi, Marta Spinella, Maxwell Tuuli, and Alexander Zaborovskiy. Administrative assistance was provided by Linda Bisman and Merceditas San Pedro-Pribram.

CONTENTS

Abbreviations and Acronyms	5
SECTION I. PROGRAM SUCCESS DURING 2002–24	8
A. Methodology	8
B. Program Success and Empirical Results	11
C. Further Considerations for PRGT-supported Programs	19
SECTION II. REALISM OF GROWTH FORECASTS	22
A. Improving Forecast Realism: Gains and Remaining Gaps	22
B. Enhancing Resilience: Tailored Assumptions and Contingency Planning	24
SECTION III. ASSESSING FISCAL POLICY CONDITIONALITY	25
A. Calibrating Fiscal Adjustment to Shocks: Balancing Flexibility and Realism	25
B. Fiscal Performance under Strain: Initial Shortfalls and Recent Gains	27
SECTION IV. ASSESSING SOCIAL CONDITIONALITY	29
A. Program Design and Implementation	30
B. Preliminary Evidence on Social Outcomes	32
C. Gender in IMF-Supported Programs	33
SECTION V. PUBLIC DEBT AND FORECAST ACCURACY	34
A. Debt Trends and Risks	34

B. Debt Forecast Errors and Drivers	38
SECTION VI. EVOLVING MONETARY POLICY CONDITIONALITY	42
A. Inflation Outcomes and Drivers	42
B. Experience with Review-Based Monetary Policy Conditionality	47
SECTION VII. EXCHANGE RATE AND EXTERNAL ADJUSTMENT	53
A. External Vulnerabilities and Exchange Rate Adjustment	53
B. Large Exchange Rate Depreciation and the External Balance Assessment	55
SECTION VIII. ASSESSING STRUCTURAL CONDITIONALITY	56
A. Overview	57
B. Impact of Structural Conditionality on Growth	61
C. Role of Capacity Development	64
SECTION IX. FRAGILE, CONFLICT-AFFECTED STATES AND SMALL DEVELOPING STATES	67
SECTION X. FOLLOW-ON PROGRAM ENGAGEMENT	71
SECTION XI. ENTERPRISE RISK ASSESSMENTS IN UFR	76
SECTION XII. STAKEHOLDER SURVEY: PRELIMINARY VIEWS	79
A. Effective and Timely Support	79
B. Focused, Evenhanded, and Tailored Conditionality	80
C. Trusted Advisor	80
BOXES	
1. Program Ownership and Macroeconomic Outcomes	12
2. Alternative PRGT Success Methodology	21
3. Trends and Integration of Capacity Development in IMF-Supported Programs	65
FIGURES	
1. GRA Framework in the 2018 ROC	9
2. PRGT Framework in the 2018 ROC	9
3. Stylized Facts of Fund-Supported Programs	14
4. Stylized Facts of Successful Programs: Pre-Program Conditions and Adjustment	15
5. Stylized Facts of Successful Programs: Program Design and Implementation	17
6. Correlates of Program Success: Existing Conditions, Shocks, and Program Design	18
7. Success of PRGT-Supported Program: Alternative Methodology	20
8. Growth Forecast Errors	22
9. Fiscal Adjustments – Program Design	26

10. Fiscal Adjustments – Outcomes	28
11. Social Spending Conditionality – Main Indicators	31
12. Dynamic Effects of IMF-Supported Programs on Social Spending, Poverty, and Inequality	32
13. Public Debt to GDP	34
14. Financing Conditions and Debt Composition	35
15. Assessment of Risk of Debt Distress	37
16. The Use of Debt Conditionality in IMF-Supported Programs	38
17. Empirical Decomposition of Debt Forecast Errors	41
18. Expected Change in Inflation Under IMF-Supported Programs	43
19. Inflation Outcomes Under IMF-Supported Programs	44
20. Inflation Outcomes Under IMF-Supported Programs	45
21. A Decomposition of Inflation Forecast Errors (FEs)	46
22. Monetary Policy Conditionality under IMF-Supported Programs	47
23. Distribution of ICC and MPCC Programs by Income Level and Policy Regime	48
24. Central Bank Transparency Index	48
25. RBC Target Levels and Bandwidths Interval by Time Period	49
26. Inflation Outcomes Relative to ‘No-Program’ Counterfactuals	50
27. RBC Compliance Rates	51
28. Key Drivers of Breaches of Review Based Conditionality	52
29. Exchange Rate and External Adjustment	54
30. EBA CA Model: Realignment Gap	56
31. Institutional Quality at Program Approval	57
32. Structural Focus of IMF-Supported Programs and Related Conditionality	58
33. SC – Time to Implementation from Set to Test Date	59
34. Average High Structural Debt by Sector	59
35. Design and Implementation of Structural Conditionality	60
36. Strength of Structural Conditionality, Institutional Quality, and Growth	62
37. Counterfactual Analysis of the Effect of IMF Arrangements on Growth	64
38. CD Integration into Fund-supported Programs	66
39. Program Outcome in FCS and SDS	68
40. Program Design in FCS and SDS	70
41. Trends in Follow-on Program Engagement	71
42. Distinguishing Follow-on Program Engagement	72
43. GRA Successor vs Repeated Programs	74
44. PRGT Successor vs Repeated Programs	75
45. Enterprise Risk Assessments for Having a New Arrangement or Proceeding with a Review	77
46. IMF Survey of Stakeholders: Views on Outcomes, Programs, and Trust	81

TABLES

1. Regression Results	40
2. Decomposition for Different Sub-Samples	42

References	82
------------	----

Abbreviations and Acronyms

AREAER	Annual Report on Exchange Arrangements and Exchange Restrictions
AEs	Advanced Economies
BOP	Balance of Payments
BRM	Base or Reserve Money
CA	Current Account
CB	Central Bank
CD	Capacity Development
CDMAP	Capacity Development Management and Administration Program
CEMAC	Central African Economic and Monetary Community
CFM	Capital Flow Management
COVID-19	Coronavirus Disease of 2019
CPI	Consumer Price Index
DD	Debt Distress
DRM	Domestic Revenue Mobilization
DSA	Debt Sustainability Assessment
DSF	Debt Sustainability Framework
EA	Exceptional Access
EBA	External Balance Assessment
ECF	Extended Credit Facility
EFF	Extended Fund Facility
EM	Emerging Market
EMDEs	Emerging Market and Developing Economies
ERA	Enterprise Risk Assessment
ERM	Enterprise Risk Management
ESA	External Sector Assessment
FCS	Fragile and Conflict-Affected States
FDI	Foreign Direct Investment
FCL	Flexible Credit Line
FTE	Fiscal Transparency Evaluation
FE	Forecast Error
FX	Foreign Exchange
FXI	Foreign Exchange Intervention
GAS	Global Assumptions
GDP	Gross Domestic Product
GEE	Global Economic Environment
GFC	Global Financial Crisis
GFCF	Gross Fixed Capital Formation
GRA	General Resources Account
ICCs	Inflation Consultation Clauses
IDS	International Debt Statistics

IEO	Independent Evaluation Office
IFIs	International Financial Institutions
IMF	International Monetary Fund
ITs	Indicative Targets
LFP	Labor Force Participation
LICs	Low-Income Countries
LP	Local Projections
LTPE	Longer-Term Program Engagement
MAC	Market Access Countries
MCPs	Multiple Currency Practices
MEFP	Memorandum of Economic and Financial Policies
MONA	Monitoring of Fund Arrangements
MP	Monetary Policy
MPCCs	Monetary Policy Consultation Clauses
NDA	Net Domestic Assets
NEER	Nominal Effective Exchange Rate
NIIP	Net International Investment Position
NIR	Net International Reserves
OLS	Ordinary Least Squares
PAs	Prior Actions
PC	Performance Criteria
PCI	Policy Coordination Instrument
PCL	Precautionary Credit Line
PFM	Public Financial Management
PLL	Precautionary and Liquidity Line
PPG	Public and Publicly Guaranteed
PSI	Policy Support Instrument
PRGT	Poverty Reduction and Growth Trust
QPCs	Quantitative Performance Criteria
QTs	Quantitative Targets
RBC	Review-Based Conditionality
REER	Real Effective Exchange Rate
ROC	Review of Program Design and Conditionality
RTs	Reform Targets
SBs	Structural Benchmarks
SBA	Stand-by Arrangement
SC	Structural Conditionality
SCF	Stand-by Credit Facility
SCID	Structural Conditionality Implementation and Depth
SCIDG	Structural Conditionality Implementation and Depth Growth
SCM	Synthetic Control Method
SDI	Structural Development Indicator
SDS	Small Developing States

SLL	Short-term Liquidity Line
SOEs	State-Owned Enterprises
SSC	Social Spending Conditionality
SRDSF	Sovereign Risk and Debt Sustainability Framework
TIMS	Travel Information Management System
UFR	Use of Fund Resources
VE	Vulnerability Exercise
WAEMU	West African Economic and Monetary Union
WB	World Bank
WDI	World Development Indicators
WEO	World Economic Outlook
XR	Exchange Rate

SECTION I. PROGRAM SUCCESS DURING 2002–24¹

A longer-term empirical analysis of the key factors driving program performance is conducted. It builds on the 2018 ROC methodology of program success and notably expands the set of programs assessed using machine learning techniques. The analysis underscores the importance of timely, realistic, and sustained adjustment, alongside strong ownership, as robust contributors to program success. Conversely, adverse shocks—particularly in the presence of significant underlying vulnerabilities—substantially reduce the likelihood of achieving program objectives.

A. Methodology

1. The methodology to measure program success has evolved to address empirical challenges inherent in evaluating Fund-supported programs. These challenges include selection bias, confounding factors, and the difficulty in identifying counterfactuals. While case studies and instrumental variable approaches offer insights, their validity remains limited. Over time, various approaches have been used to assess program performance. [Ivanova and others \(2003\)](#) emphasized implementation metrics such as conditionality compliance, program interruptions, and disbursement rates. The [IEO \(2008\)](#) focused on the role of structural conditionality (SC) in advancing reforms, while the 2011 ROC examined macroeconomic outcomes relative to program design ([IMF, 2012a, 2012b](#)). The 2018 ROC introduced a framework aligned with primary macroeconomic goals to resolve BOP problems without measures destructive of national and international prosperity and achieve medium-term external viability while fostering sustainable economic growth, as per the Conditionality Guidelines ([IMF, 2019a](#)).

2. This analysis applies the 2018 ROC framework, which introduced a differentiated approach to assessing program performance under GRA and PRGT facilities. Reflecting their distinct ways to operationalize their objectives, the 2018 ROC developed tailored methodologies to classify programs as “successful,” “partially successful,” or “unsuccessful.”

- *The GRA methodology focuses on Vulnerability Exercise (VE) assessments and post-program engagements. An improvement in the VE final overall assessment during the program contributes to a better success rating,² while a drawing follow-on arrangement contributes to a worse rating and precludes the possibility of classifying a program as successful (Figure 1).^{3 4}*

¹ Prepared by Anil Ari, Anja Baum, Jocelyn Boussard, Giulio Lisi, Gaëlle Pierre and Faezeh Raei, with research assistance from Juliana Dueñas Villabona, Salvador Traettino, Helen Zheng, and Max Kushnir.

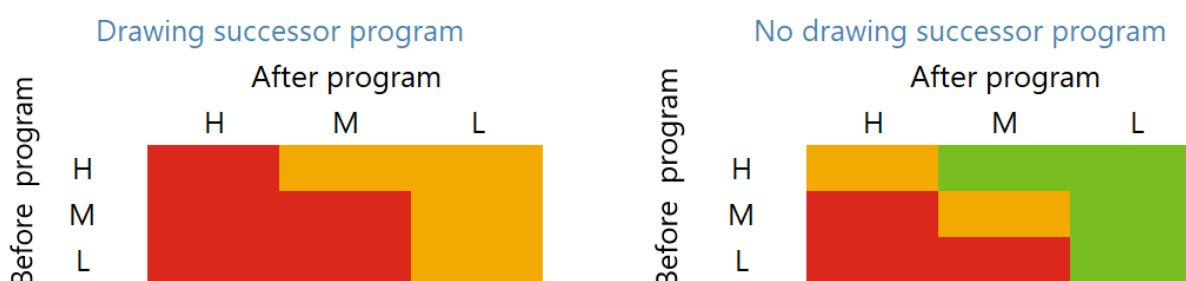
² The VE is a multi-sectoral assessment of risks that could make a country vulnerable to a BOP crisis or other sectoral crises. The overall VE assessment is computed as an average of external, fiscal, financial and non-financial corporate and household sector VE model assessments, normalized by percentile ([IMF, 2021](#)).

³ Post-program engagement refers to the two-year period following the end of a Fund arrangement. A drawing follow-on arrangement, in this section, refers to a financing arrangement involving a drawing, excluding those with access below a quarter of the annual exceptional access threshold. Lower access programs include annualized access below 50 percent of quota for programs approved before January 2016 and 33.75 percent thereafter.

⁴ The empirical results presented here are consistent with those based on a variant of success ratings that excludes follow-on arrangements from consideration.

- The PRGT methodology focuses on external debt vulnerabilities and indicators of poverty reduction policies and macroeconomic stability—core objectives of PRGT facilities. It employs a two-step approach (Figure 2): Step 1 uses changes in LIC-DSA ratings (a proxy for external debt vulnerabilities) to identify unsuccessful programs. Step 2 assesses remaining programs against five indicators: (i) social spending on health and education, (ii) government capital expenditure, (iii) non-grant fiscal revenues, (iv) real GDP growth, and (v) inflation. Programs are assessed based on how many indicators over-perform forecasts at program approval.⁵

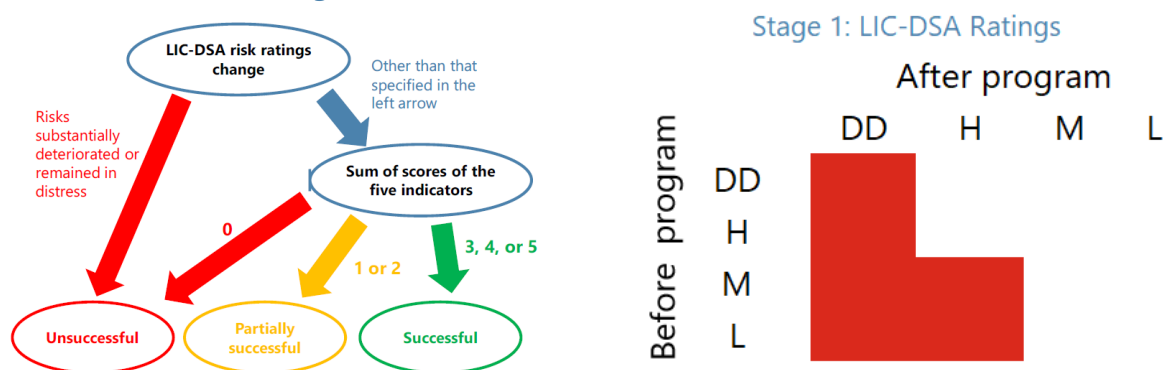
Figure 1. GRA Framework in the 2018 ROC



Sources: IMF staff calculations.

Notes: The matrices show the assignment of program performance ratings according to the mapping of VE final overall risk ratings before and after the program. Green, yellow and red color coding respectively refer to successful, partially successful and unsuccessful performance ratings. H, M, L respectively refer to high, medium and low VE risk ratings. The matrix on the left (right) is applied to programs with (without) a drawing follow-on arrangement. See [IMF \(2019a\)](#) for further elaboration on the methodology and its motivation.

Figure 2. PRGT Framework in the 2018 ROC



Sources: IMF staff calculations.

Notes: The left panel shows the two stages. The right panel shows the matrix used to select unsuccessful programs in Stage 1, which are denoted in red color coding. DD, H, M and L respectively refer to debt distress and high, medium, and low LIC-DSA risk ratings.

⁵ Compared with the GRA methodology, performance is measured relative to program design rather than past outcomes. This avoids penalizing PRGT programs given the historic volatility of growth, inflation, and budget-related indicators. When forecast data are unavailable, performance is assessed relative to pre-program outcomes.

3. Relative to the 2018 ROC, the analysis significantly extends the coverage of GRA program performance by applying machine learning methods. To enrich the assessment, it includes programs previously excluded owing to missing VE ratings in the pre- or post-program period. A random forest algorithm ([Hastie and others, 2009](#)) is used to predict these missing assessments, allowing for a broader set of GRA arrangements in the analysis.⁶

4. As a robustness check, this analysis presents an alternative PRGT methodology to further align the measure of success with the PRGT objectives (Section C). The alternative PRGT methodology offers refinements including accounting for external imbalances, beyond external public debt, and avoiding penalizing marginal target misses when overall progress is meaningful.

5. Taking advantage of the expanded constructed sample, this analysis conducts an empirical assessment of the determinants of program performance. Using the 2018 ROC methodology and the expanded GRA sample, empirical analysis is conducted on the relationship between the macroeconomic conditions prior to and during the programs, program design and implementation, and program performance.⁷

First, fixed effects panel regressions are used to explore stylized facts about IMF programs and their associations with program performance. The regression specifications are as follows:

$$Y_{i,t} = \alpha_i + \gamma_t + \beta P_{i,t} + \theta X_{i,t} + \epsilon_{i,t} \quad (1)$$

$$Y_{i,t} = \alpha_i + \gamma_t + \beta_S S_{i,t} + \beta_U O_{i,t} + \theta X_{i,t} + \epsilon_{i,t} \quad (2)$$

where i, t denote country and year, and α_i and γ_t control for country and time fixed effects. The regressions are conducted for a range of macroeconomic variables ($Y_{i,t}$). $X_{i,t}$ is a vector of control variables. In specification (1), $P_{i,t}$ is a binary variable which takes the value of one when there is a Fund-supported program such that the coefficient β estimates the mean difference in Y between programs and non-programs, controlling for fixed effects and $X_{i,t}$. Similarly, in specification (2), $S_{i,t}$ and $O_{i,t}$ are binary variables for fully successful and other (i.e., partially successful and unsuccessful) programs respectively, with the difference between their coefficients β_S and β_U estimating the mean difference in Y between these two groups, controlling for fixed effects and $X_{i,t}$. $\epsilon_{i,t}$ is the residual.⁸

⁶ The method proceeds by, first, gathering key indicators of emerging market (EM) crisis vulnerability indicators, based on those identified by ([IMF, 2021](#)). These variables are used to train the random forest model on observed VE ratings. The underlying assumption—validated by pairwise regressions—is that EM crisis vulnerability indicators (e.g., current account deficits, reserves, and public debt) correlate with and can therefore help to predict VE scores. As a last step, results from the training group are used to predict missing VE ratings.

⁷ The analysis focuses on empirical associations and should not be interpreted causally given the challenges described in ¶11 and the endogenous nature of success ratings.

⁸ The regressions use WEO data. Partially successful and unsuccessful programs are grouped as the latter make up only 25 percent of the sample. When focusing on initial (pre-program) conditions, program data are excluded and the binary variables equal one in the year before program approval. When focusing on macroeconomic outcomes during programs, data are transformed in difference-in-difference to capture changes or average changes on a 3-year horizon, depending on the variable. Fund program data are included only once to avoid repeat sampling. The vector $X_{i,t}$ is empty in the baseline but includes a binary variable for follow-on programs, interactions of VIX with pre-

(continued)

When Y reflects program design or implementation (e.g., the number or implementation of SCs), there is no time dimension or non-program data, and the regression is specified as follows:

$$Y_p = a + \gamma S_p + \theta X_p + \epsilon_p \quad (3)$$

where p denotes a cross-section of Fund-supported programs, a is the intercept, and S_p is a binary variable for a (fully) successful program such that its coefficient γ estimates the mean difference between successful and other programs.

Second, logistic regression models are used to gauge predictors of program success. S_p becomes the dependent variable and a vector ϕ_p encompasses covariates of interest, including proxies for external shocks during the program (such as the GFC, the COVID-19, and changes in terms of trade), pre-program macroeconomic characteristics, and elements of program design and implementation. Formally, the regression specification is the following:

$$\frac{\Pr(S = 1)}{1 - \Pr(S = 1)} = e^{(\alpha + \phi \beta' + \epsilon_p)} \quad (4)$$

Estimated probabilities from specification (4) are used to identify potential predictors of program success, and how expected performance varies in response to Fund engagement.

B. Program Success and Empirical Results

6. The empirical analysis highlights the role of initial conditions, adjustment paths, and external shocks underlying program success.⁹ The findings include:

- **Fund-supported programs** are often preceded by adverse shocks, and approved after macroeconomic conditions have already deteriorated, giving rise to BOP problems and fiscal vulnerabilities (Figure 3). While programs typically achieve some external adjustment and fiscal consolidation without further deteriorating macroeconomic conditions, their success in achieving catalytic effects and improving structural outcomes is mixed.
- **Successful programs** tend to start with lower public debt-to-GDP ratios and less erosion of market confidence, while unsuccessful ones are often preceded by protracted macroeconomic imbalances—including domestic credit booms and currency devaluations—and adverse external shocks during the program (Figures 4 and 6). Once underway, successful programs show stronger macroeconomic performance and sizable catalytic effects, leading to exchange rate appreciation

program debt-to-GDP, current account balances and household credit-to-GDP and interactions of oil prices with pre-program oil trade balances in robustness checks.

⁹ These stylized facts are largely robust to different time horizons, controls for follow-on programs, and the interaction of exogenous trade, financing and oil shocks with pre-program vulnerabilities (time fixed effects capture shocks that do not interact with pre-existing conditions). They are also broadly robust to excluding the nature of post-program Fund engagement from the definition of success for GRA arrangements.

and less (need for) external adjustment (Figure 4). They also tend to achieve stronger structural reform outcomes, alongside increases in private credit and investment.

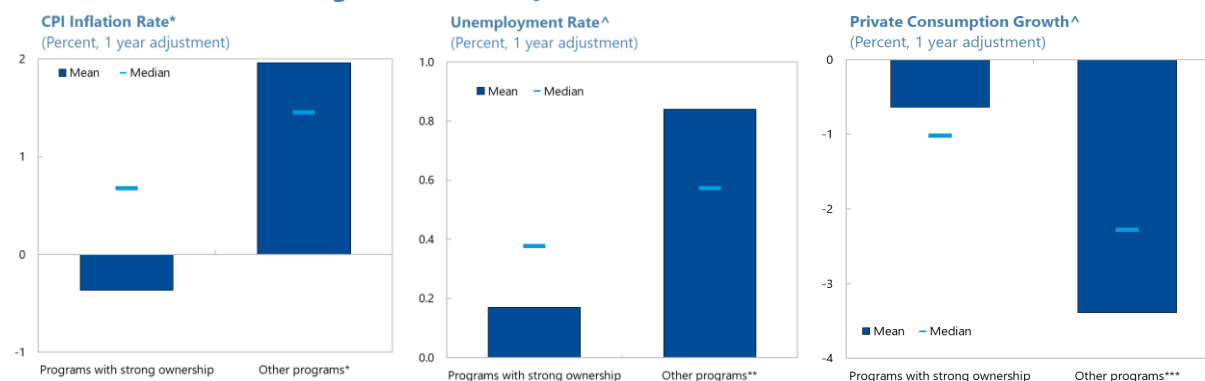
- **Program design and implementation—and their interaction with program ownership and exogenous shocks—play a key role in determining program outcomes.** Successful programs typically feature more parsimonious conditionality (fewer structural benchmarks and quantitative targets) and higher implementation rates (Figure 5). Program success is associated with a balanced mix of adjustment and financing, as well as frontloaded fiscal measures and realistic fiscal forecasts, even after controlling for shocks and pre-program conditions (Figure 6), while larger or more frontloaded access is generally not a strong predictor of success. Successful programs also show stronger ownership (proxied by review completion rates, Box 1) and longer duration, which—along with higher implementation rates—may reflect sustained adjustment, especially amid elevated pre-program debt vulnerabilities and external shocks.

Box 1. Program Ownership and Macroeconomic Outcomes

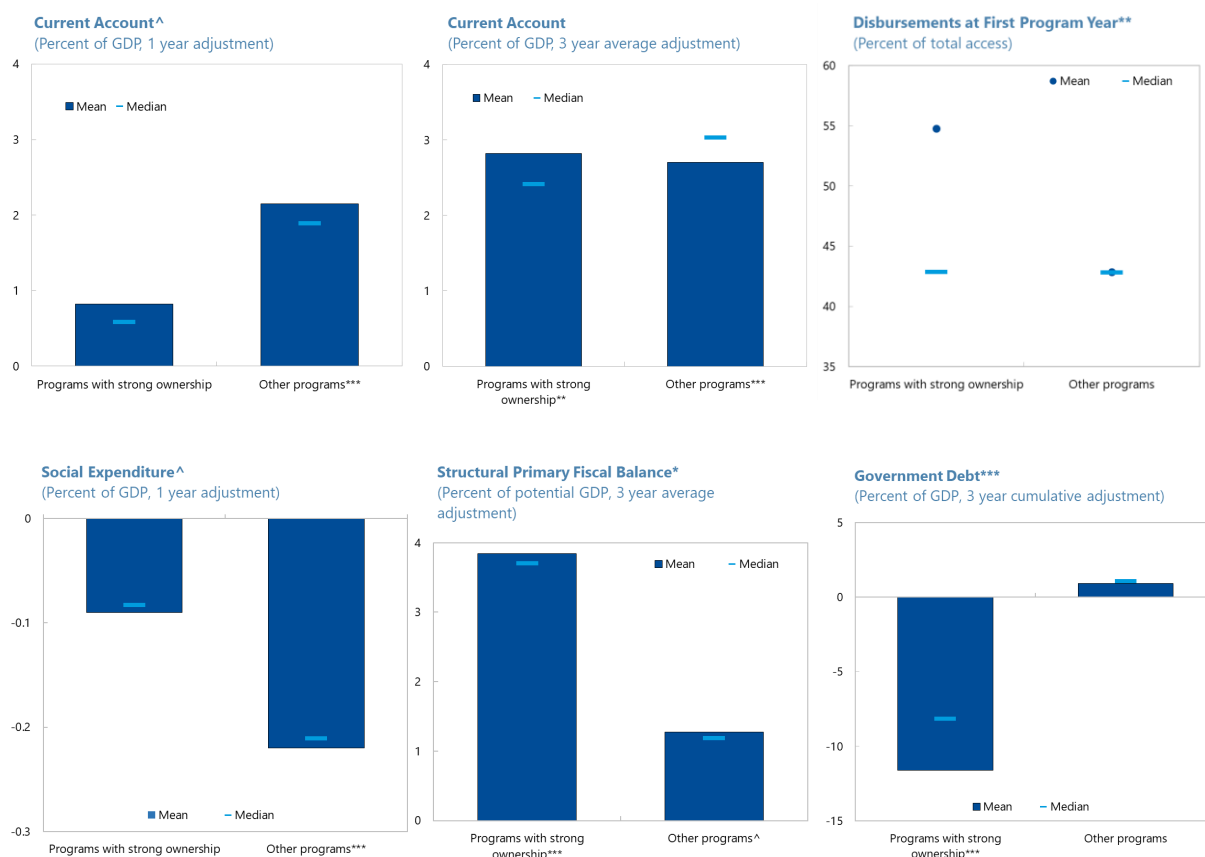
To better understand the association between program ownership and macroeconomic outcomes, this analysis draws on the framework used to examine the determinants of program success (15). Ownership is proxied by a binary variable indicating above/below median review completion rates, constructed for 322 Fund-supported programs over the period 2002–2024.^{1/ 2/} This variable substitutes for program success indicators in the empirical specifications.^{3/}

Strong ownership is associated with gradual—but not less—external adjustment and protection of social spending during fiscal consolidation. Programs with strong ownership typically have better unemployment, inflation, and consumption outcomes in the first year of the program. Such programs also tend to have more frontloaded financing coupled with gradualism in external adjustment, while protecting social expenditures despite undertaking more substantial fiscal consolidation.^{4/}

Program Ownership and Macroeconomic Outcomes



Box 1. Program Ownership and Macroeconomic Outcomes (concluded)



Sources: MONA database; IMF staff calculations.

Notes: Blue bars denote mean differences from non-program cases controlling for country and fixed effects per specification (2) with $S_{i,t}$ and $O_{i,t}$ re-defined as above/below median review completion rate binary variables. Light blue bars denote median differences residualized from the same fixed effects. Statistical significance of the mean differences from non-program cases are shown next to the X axis labels. Adjustment refers to differences from pre-program values. In the center right chart, statistical significance of differences in means is estimated per specification (3) with S_p re-defined as an above median review completion rate binary variable and shown next to the chart title. Robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^ $p < 0.25$.

^{1/} Program ownership is a multi-faceted concept that cannot be fully captured with a single metric. While the 2018 ROC considers a number of alternative metrics, this section focuses on review completion rates due to data availability constraints. Sensitivity analysis is conducted using a dummy variable for programs which quickly go off-track, and yields results which are consistent with the baseline.

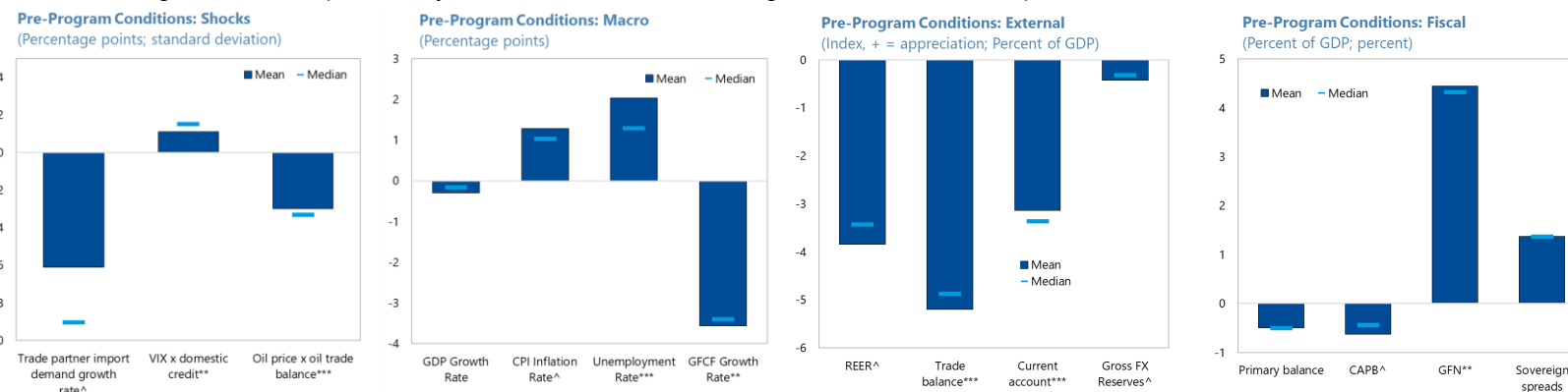
^{2/} This includes all program with available data and includes a larger set of programs than the program success analysis due to the reduced data requirements.

^{3/} As with the program success analysis, the empirical results should not be interpreted causally given the endogeneity of review completion rates.

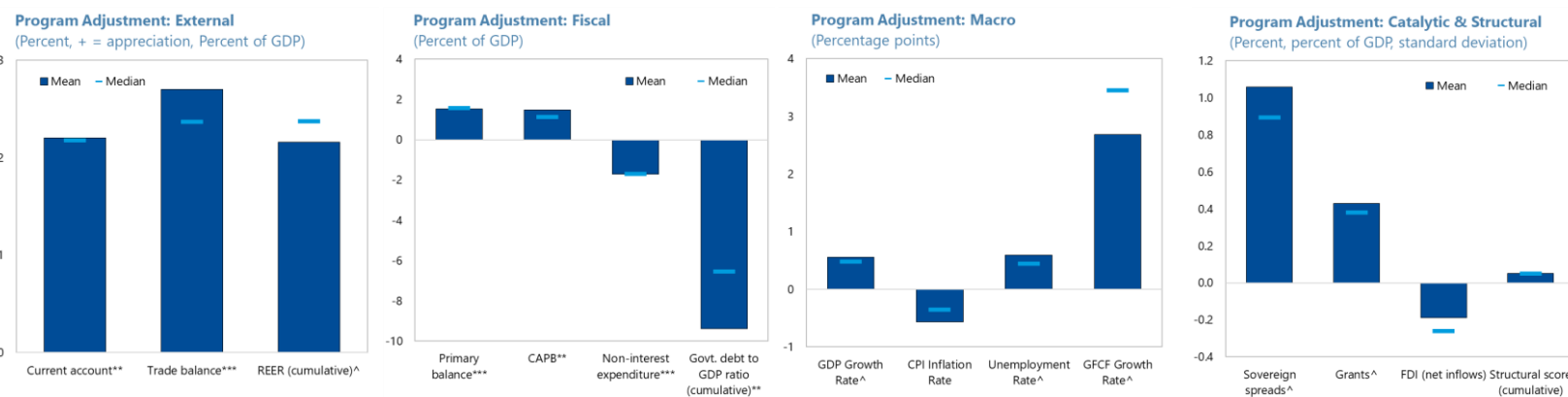
^{4/} Higher review completion rates are empirically associated with less stringent structural conditionality, lower reliance on prior actions, and increased flexibility in adjusting and postponing structural benchmarks and quantitative targets. These associations are likely driven by reverse causality as these attributes could also lower the bar for completing reviews.

Figure 3. Stylized Facts of Fund-Supported Programs

Programs are often preceded by shocks, which lead to worsening macro conditions, BOP problems, and elevated fiscal vulnerabilities.



Programs typically involve external and fiscal adjustment, without undermining macroeconomic outcomes though presenting a mixed picture for catalytic effects and structural outcomes.

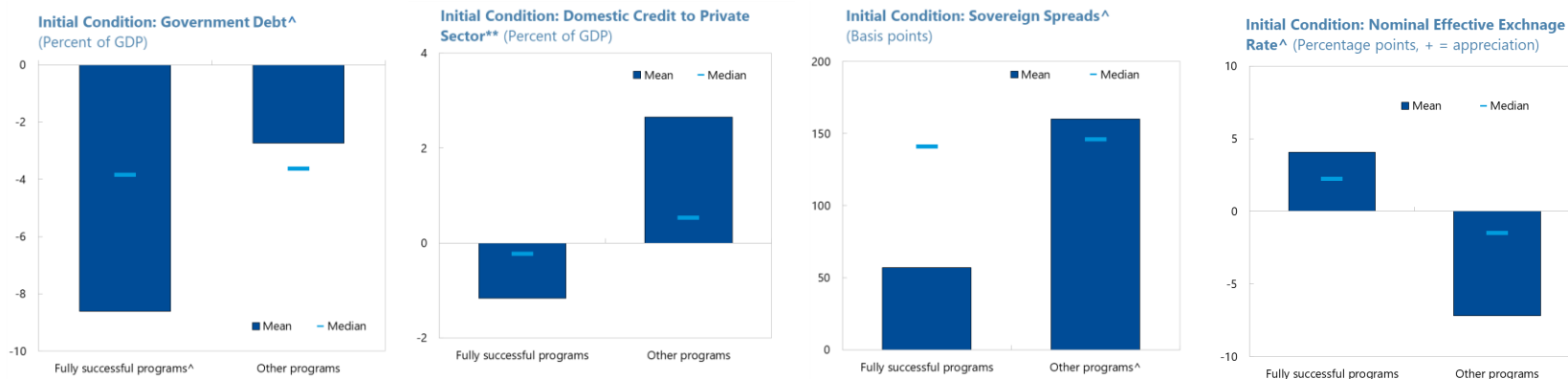


Sources: IMF staff calculations.

Notes: Pre-program refers to the year before program approval. Program adjustment refers to differences between 3-year averages after program approval and pre-program values. For variables labelled "cumulative," program adjustment refers to differences between 3 years after program approval and pre-program values. Blue bars denote mean differences between program and non-program cases controlling for country and fixed effects per specification (1). Light blue bars denote median differences residualized from the same fixed effects. Statistical significance of the mean differences is shown next to the X axis labels. Robust standard errors. *** p<0.01, ** p<0.05, * p<0.1, ^ p<0.25.

Figure 4. Stylized Facts of Successful Programs: Pre-Program Conditions and Adjustment

Successful programs typically start with lower government debt ratios, slower domestic credit growth, and less substantial erosion of confidence, as indicated by lower spreads and stronger nominal exchange rates.



Successful programs are characterized by stronger macroeconomic performance and more substantive catalytic effects, through lower spreads...

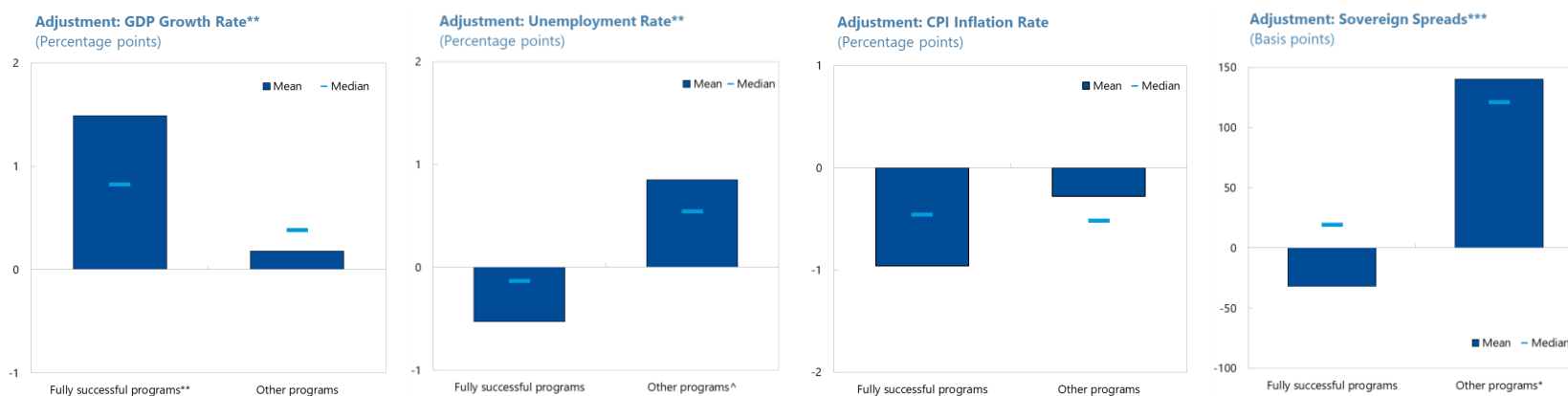
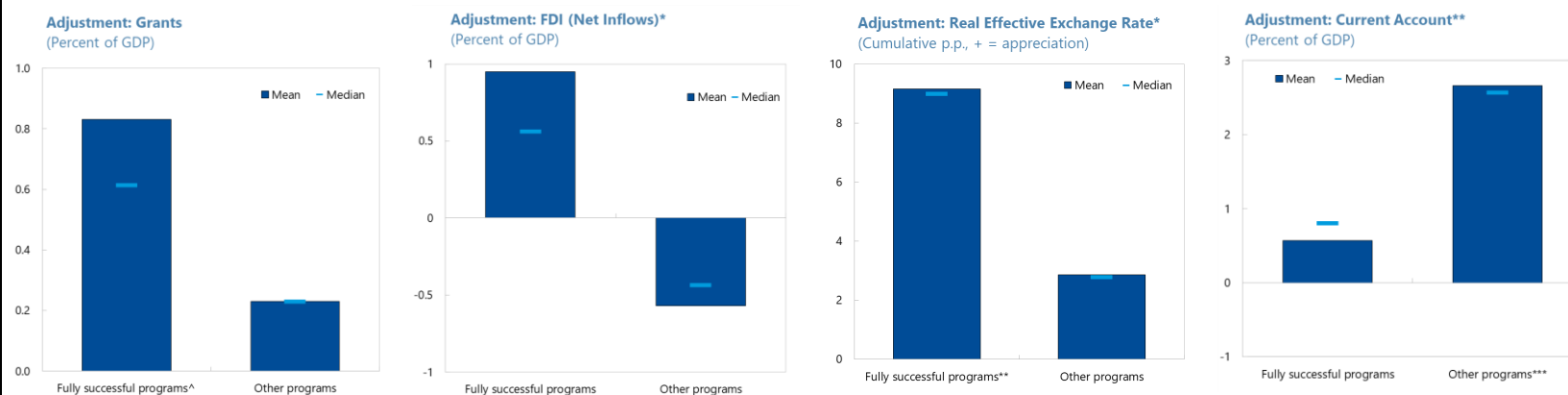
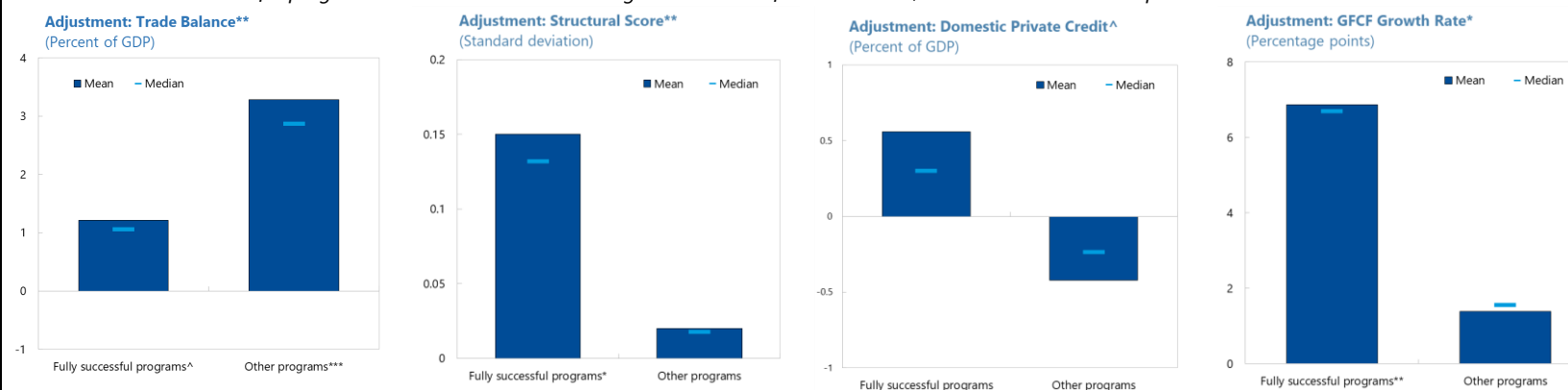


Figure 4. Stylized Facts of Successful Programs: Pre-Program Conditions and Adjustment (concluded)

... and higher grants and private inflows, leading to exchange rate appreciation and less (need for) external adjustment.

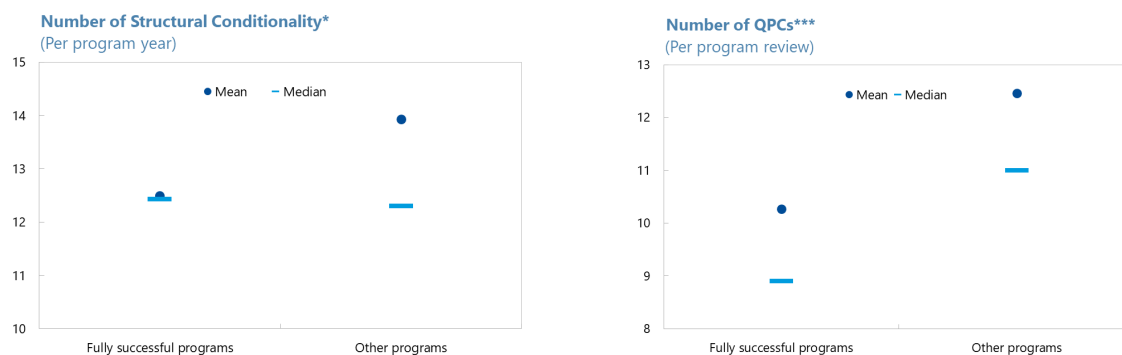
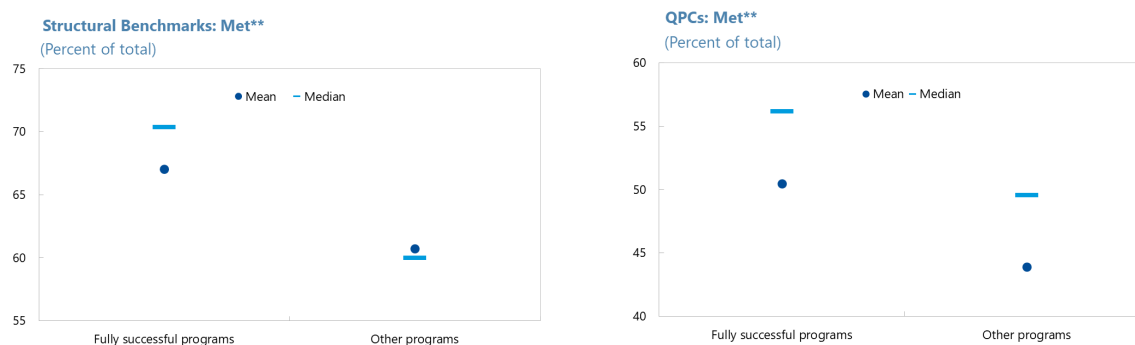


Successful programs also tend to achieve stronger structural reform outcomes, as well as a rebound in private credit and investment.



Sources: IMF staff calculations.

Notes: For initial condition charts, pre-program refers to the year before program approval. Blue bars denote mean differences from non-program cases controlling for country and fixed effects per specification (2). Light blue bars denote median differences residualized from the same fixed effects. Statistical significance of the mean differences from non-program cases are shown next to the X axis labels. For adjustment charts, program adjustment refers to differences between 3-year averages after program approval and pre-program values. Blue bars denote mean differences from non-program cases controlling for country and fixed effects. Light blue bars denote median differences residualized from the same fixed effects. Statistical significance of the mean differences from non-program cases are shown next to the X axis labels. Statistical significance of the mean differences between fully successful and other programs are shown next to the chart titles. Robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^ $p < 0.25$.

Figure 5. Stylized Facts of Successful Programs: Program Design and Implementation*Successful programs tend to have more parsimonious conditionality...**...and a stronger track record at implementing structural benchmarks and quantitative targets.*

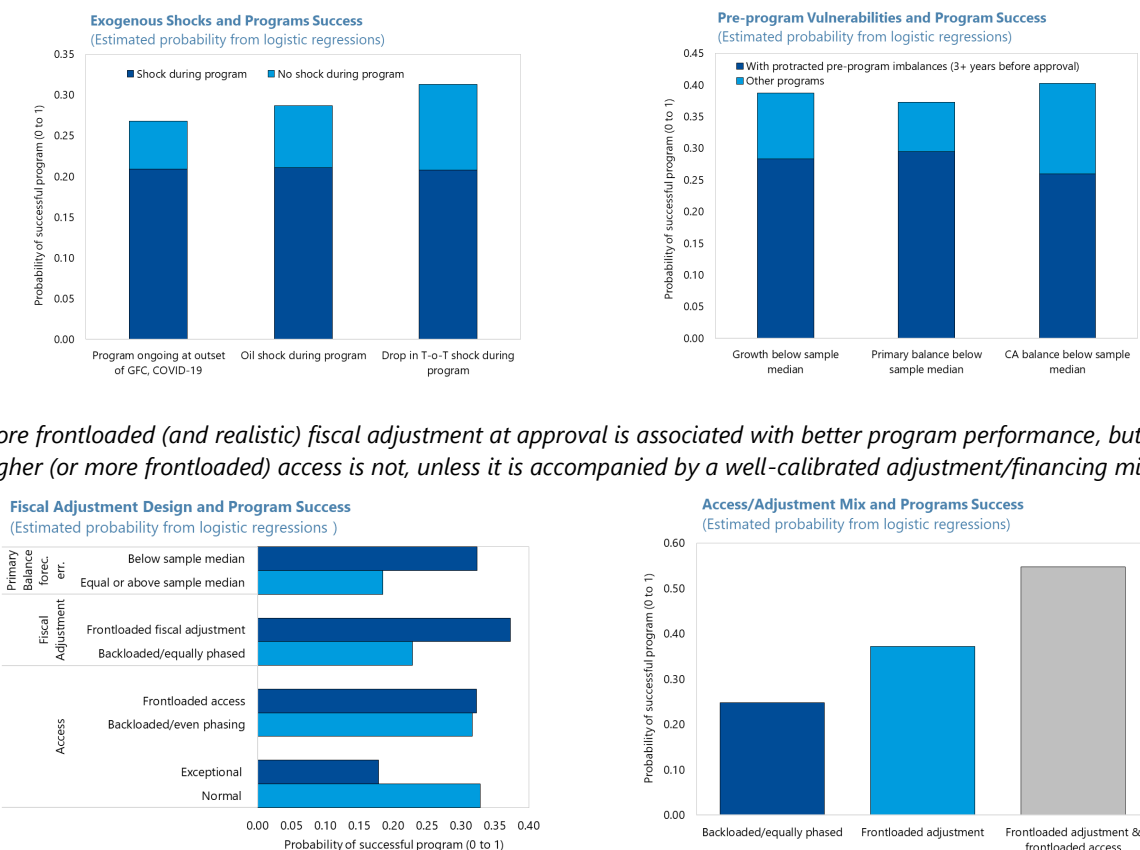
Sources: IMF staff calculations.

Notes: Structural conditionality includes structural benchmarks and prior actions at program approval and all program reviews. Blue bars denote means and light blue bars denote medians. Statistical significance of differences in means is estimated per specification (3) and shown next to the chart titles. Robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, ^ $p < 0.25$.

Figure 6. Correlates of Program Success: Existing Conditions, Shocks, and Program Design

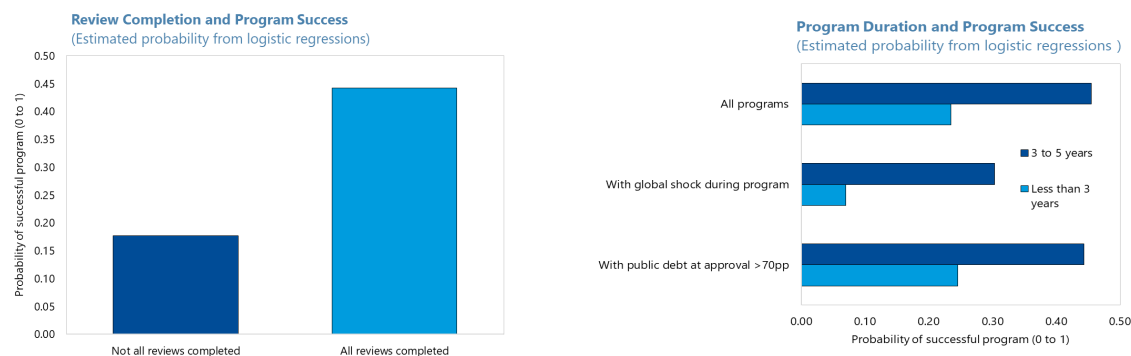
Protracted macroeconomic imbalances in the years leading to program approval

...and external shocks during programs are associated with lower probability of success.



More frontloaded (and realistic) fiscal adjustment at approval is associated with better program performance, but higher (or more frontloaded) access is not, unless it is accompanied by a well-calibrated adjustment/financing mix.

Strong implementation and ownership proxies are strong predictors of success and longer engagements tend to be associated with better performance, particularly in cases of adverse shocks or pre-program vulnerabilities.



Sources: IMF staff calculations.

Notes: Estimated probabilities from specification (4). Controls include proxies for shocks during programs, lagged macroeconomic variables, and gross debt and IMF credit outstanding. *Exceptional access*: binary variable equal to 1 if the program had exceptional access at approval. *Frontloaded access*: binary variable equal to 1 if available Fund resources at approval exceed 20 percent of total available under program at approval. *Frontloaded adjustment*: Binary variable equal to 1 when projected average primary balance in the first two years of program (approval and subsequent year) exceeds the projected average of the first 4 years from approval. *Primary balance forecast errors*: difference between observed and projected primary balance (mean of three years from approval, 1 if the difference is above sample median).

C. Further Considerations for PRGT-supported Programs

7. An alternative success methodology for PRGT-supported programs could increase the accuracy of program performance evaluation by optimizing the weights of program goals. It places greater emphasis on macro stabilization, while maintaining consideration of poverty reduction and growth (PRG) objectives. This alternative methodology aims at:

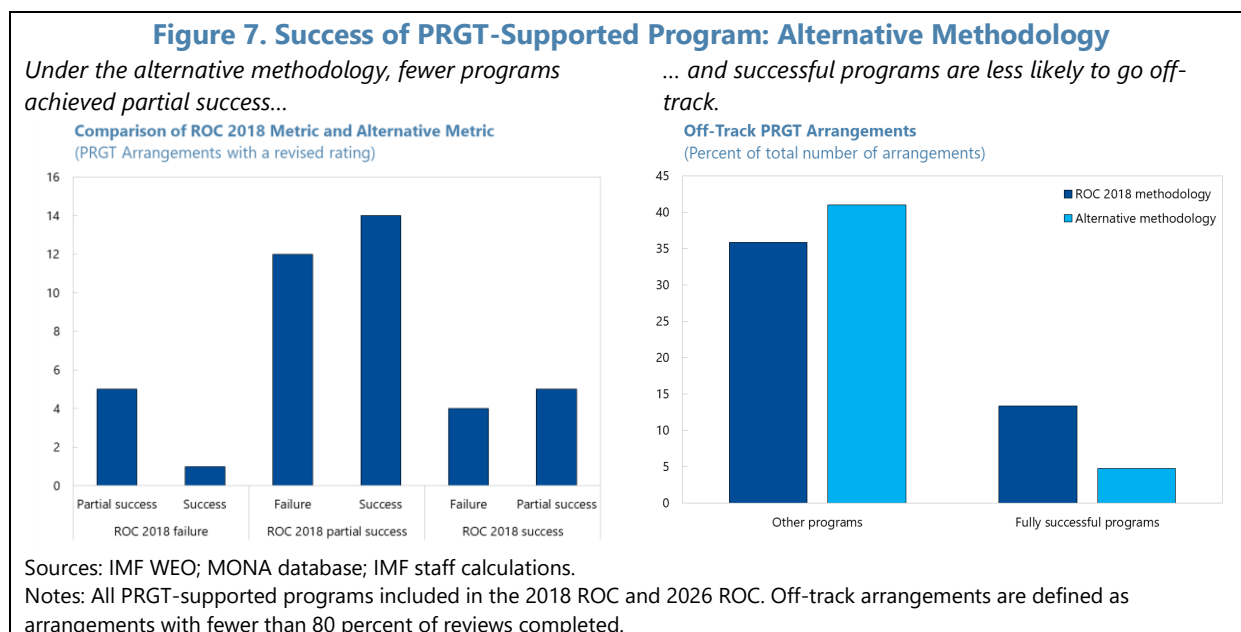
- **Accounting for external imbalances beyond external public debt.** The first step of the 2018 ROC methodology categorizes countries as unsuccessful if they remain in debt distress or move into high risk of debt distress. This ignores cases where most or all other economic indicators improve noticeably during the program (e.g., Grenada ECF 2014). Such cases might be better rated as partial success. Given the relevance of broader external sector rebalancing, greater weight could be assigned to external sector indicators when evaluating program success.¹⁰
- **Developing a more holistic approach across sectors.** In the 2018 ROC methodology, achieving the program target in just one indicator can lead to partial success, even if all other indicators worsen noticeably. This shows the importance of better balancing developments across sectors.
- **Adding more poverty indicators.** Including social sector outcomes instead of only spending (sometimes poorly targeted) could be beneficial.
- **Reducing the impact of overoptimism.** The 2018 ROC methodology only assigns positive ratings to indicators with outcomes better than targets set at program request. As shown in this and other ROCs, original targets can be overly optimistic, particularly on growth and revenue, and are often revised down during the program, contributing to excessively negative ratings.

8. The alternative methodology remains aligned with PRGT objectives (Box 2). It assigns greater weights to external public debt risk ratings, supplemented by current account and reserves indicators to capture macroeconomic stability. The social block includes a poverty indicator alongside spending inputs. Ratings are based on improvements relative to pre-program conditions, with three possible outcomes per each indicator: “unsuccessful” if an indicator is worse than at program request, “neutral” if better but below program target, and “successful” if at or better than its program target.

9. Under the alternative methodology, fewer programs achieved partial success, with a shift towards failure and success ratings. Overall, 70 percent of PRGT-supported programs show some degree of success (Figure 7). This is mostly driven by partial success cases in the 2018 ROC methodology for which the only improved indicator was social expenditure, while the overall macro

¹⁰ Frequently, external public debt and reserve accumulation are set as two distinct program goals/targets. One also does not necessarily follow the other. It is not uncommon to have low foreign reserves without a public external debt crisis. Similarly, a public external debt crisis is possible in a country with adequate reserve levels (e.g., Zambia) or the program tackles a high CA deficit while reserves are adequate owing to underlying FXI/surrender requirements.

environment deteriorated noticeably. Under the alternative methodology, many of these programs are classified as unsuccessful, as a worsening macroeconomic environment outweighs select social sector improvements. Moreover, the success ratings in the alternative methodology show stronger alignment with the completion of program reviews.



10. Most of Section B's empirical results are robust, with the following key exceptions:

- **Pre-program conditions are less relevant.** Successful programs still show lower credit growth prior to program approval, although the difference with other programs is smaller and less statistically significant than with the ROC 2018 methodology. The finding that successful programs start with lower public debt-to-GDP ratios becomes statistically insignificant.
- **Successful programs undergo external adjustment on par with other programs.** The difference in trade and current account adjustment between successful and other programs becomes statistically insignificant, although successful programs achieve external adjustment without compressing investment, unlike other programs. External shocks (such as adverse terms of trade dynamics) also lose statistical significance as (negative) predictors of program success.
- **Further refinements could be considered.** Some caveats are critical. The alternative methodology does not address all weaknesses of the 2018 ROC, presenting both challenges and opportunities for further policy and academic research. Further refinements remain possible, given flexibility in weighting indicators against each other, scoring, and rating thresholds. Including poverty rate could also raise some concerns related to comparability and timeliness. The existence of correlations across variables also begs the question of the weight of different economic concepts in the overall assessment. Currently, a mitigating factor is that despite correlations—such as between social spending and poverty or between external debt and reserves—they appear to be generally minor and likely do not distort the overall

assessment. Finally, neither the current nor alternative methodology fully address base effects, leaving some concerns that post-crisis rebounds may reflect recovery, not reform efforts—highlighting the need for sound program design and forecasting to account for these effects.

Box 2. Alternative PRGT Success Methodology

The list of indicators is extended to include additional dimensions of program success. The alternative methodology additionally includes external indicators (current account and reserves) and the poverty rate.

More granularity is introduced in the scoring of performance along each dimension. Two key changes in the proposed methodology are to account for improvements even if targets are missed, and to penalize very poor performance relative to targets and past outcomes. When forecast data are available, this methodology gives scores of:

- 1 point if the indicator is better than expected at program approval, or if the gap is smaller than the 10th percentile of the distribution of gaps and the indicator has improved relative to past outcomes.
- 0 points if the indicator has improved relative to past outcomes but the target is missed by a wider margin than the 10th percentile of the distribution of gaps.
- -1 point if the indicator is both worse than target and past outcomes.

A similar approach is taken for indicators where forecast data are not available: the target is calibrated as the median performance across PRGT countries. Finally, performance on the external debt risk rating is scored according to a more granular transition matrix, with more weight given to extreme shifts or failure to exit from debt distress (DD).

A program is deemed successful if sufficient progress is made toward macroeconomic stabilization, consistent with PRG objectives. Indicators are grouped by their progress on stabilization or PRG, and scores summed within each group. Given the asymmetry between objectives, a program failing to improve most macroeconomic stabilization indicators is, at best, a partial success, while one with stagnation in most PRG indicators can be rated successful.

Box 2. Figure 1. Proposed PRGT Success Rating

Panel A. Scoring the External LIC-DSF Risk

		After Program		
Before Program		DD ^{1/}	H	M,L
	DD	-2	1	2
	H	-2	0	2
	M,L	-2	-1	0

Panel B. Final Success Rating

		Sum of stabilization scores			
Sum of PRG scores		>1	1	0	<0
	>1	S	S	PS	PS
	1	S	S	PS	F
	0	S	PS	PS	F
	<0	PS	PS	F	F

Source: IMF staff calculations.

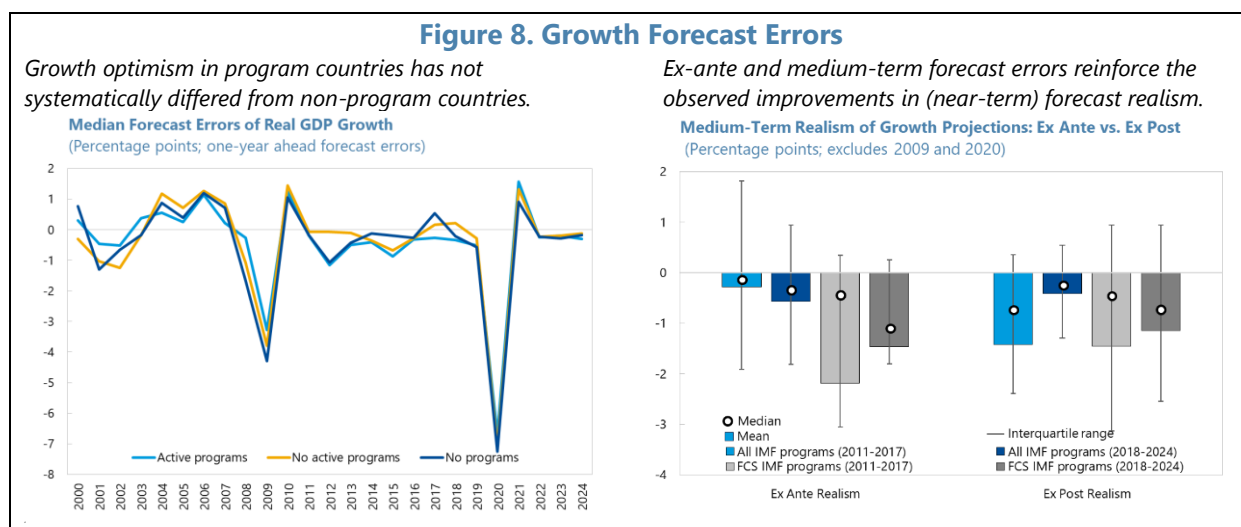
Note: A program's success rating is at most partial if the external debt risk is rated in distress after the program. S stands for success, PS for partial success, F for failure.

SECTION II. REALISM OF GROWTH FORECASTS¹¹

Growth forecasting accuracy has improved significantly since the 2018 ROC, with major errors primarily occurring during the pandemic years. However, persistent optimism remains a challenge in fragile and conflict-affected states (FCS), where forecast accuracy is hindered by institutional and security issues. Severe shocks, such as global events, have been the primary drivers of large growth forecast errors, with the pandemic alone explaining at least half of the medium-term forecast deviations in the 2026 ROC period.

A. Improving Forecast Realism: Gains and Remaining Gaps

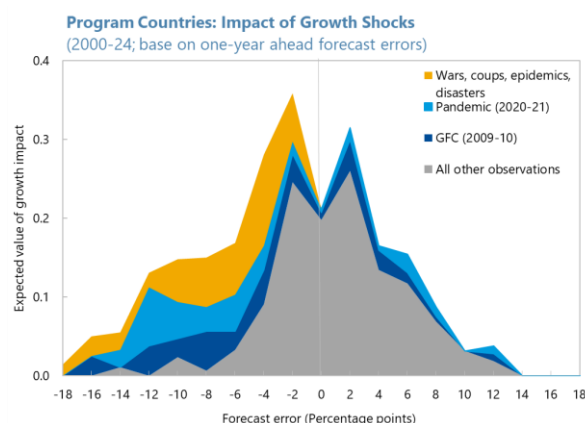
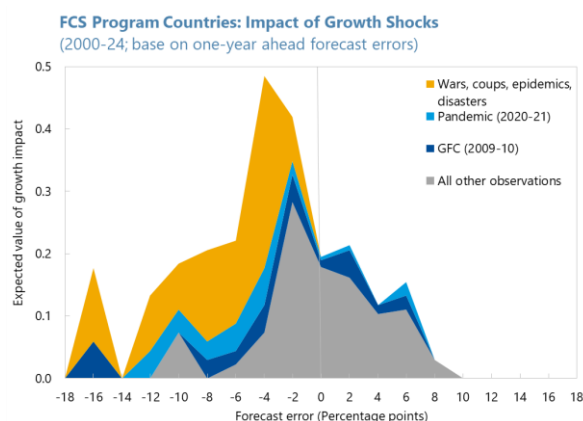
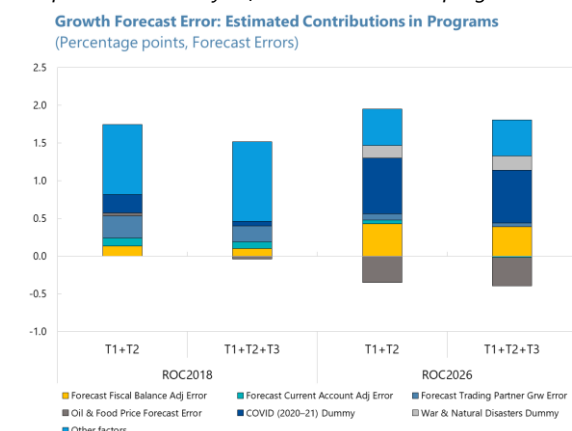
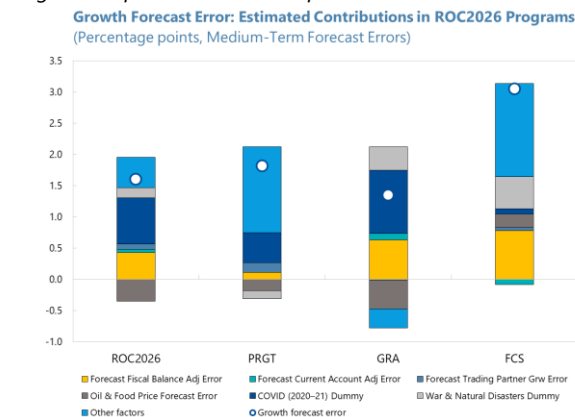
11. Growth forecasting accuracy has improved significantly since the last ROC, enhancing the credibility of macro frameworks in program design. Overly optimistic projections can understate required adjustment and financing needs, while pessimistic ones risk undue tightening. In the 2026 ROC period, major forecast errors were limited to pandemic years, with projections otherwise broadly aligning with actual outcomes. Forecast accuracy has improved across both short and medium-term horizons relative to the 2018 ROC sample. (Figure 8).¹² Nonetheless, forecast still exceed countries' historical growth record—by 0.6 percentage points in the 2026 ROC, compared to 0.3 points in the 2018 ROC.¹³ Importantly, forecast errors in program cases closely tracks those in non-program peers, suggesting that global shocks and cyclical factors—rather than program design—have been the dominant influence in forecast performance in recent years.



¹¹ Prepared by Robin Koepke, Irene Yackovlev, with research assistance from Helen Zheng.

¹² While medium-term growth forecast errors (defined as the deviation between actual growth outcomes and projections made at the time of program approval) shed light on overall optimism in program design, one-year-ahead forecast errors are better suited to isolate the impact of unforeseen shocks.

¹³ Staff analysis uses average growth during the prior 10 years to capture countries' historical growth record, excluding the years 2009 (GFC recession) and 2020 (pandemic recession). While growth forecasts in the 2018 ROC sample were only modestly above growth outturns in the preceding 10 years, these historical outturns were marked by strong growth, notably in the years leading up to the GFC.

Figure 8. Growth Forecast Errors (concluded)*Severe growth shortfalls came mainly from global shocks...**... except in FCS where local factors dominated.**The pandemic heavily influenced 2026 ROC programs...**...growth optimism was most prevalent in PRGT and FCS.*

Sources: MONA database; IMF WEO; GEE; GAS; and IMF staff calculations.

Notes: Forecast errors are calculated as the difference between actual outcomes and projections. Ex ante errors compare the pre-program 10-year average to projected post-program growth (at approval). Ex post errors compare the outturn for the average growth rate during the three years after program approval minus the projection (at approval). Expected value of growth is estimated by multiplying the historical frequency of forecast errors by their average size. The last 2 charts show the estimated contributions of external and domestic factors to growth forecast errors using a regression-based approach. The baseline regression model is described by the equation (building on the methodology in the Technical Notes of the 2018 ROC): $\text{GrowthErr}_{jit} = \text{Constant} + \beta_e \text{ExternalErr}_{jit} + \beta_d \text{DomesticPolicy}_{jit} + \text{fe}_{jit} + \text{ej}_{jit} \text{GrowthErr}_{jit} = \text{Constant} + \beta_e \text{ExternalErr}_{jit} + \beta_d \text{DomesticPolicy}_{jit} + \text{fe}_{jit} + \text{ej}_{jit}$. The underlying forecast data are based on April and October WEO vintages from 2008 to 2024, with April 2025 as the source for “actual” values. Baseline regressions are then run on a panel where the time period for each WEO vintage spans T0 to T5, and where T0 is the year that corresponds to the publication of that WEO vintage.

12. Despite recent improvements, persistent optimism remains a challenge in fragile and conflict-affected states (FCS).

Forecast accuracy in countries facing persistent institutional and security issues remains an issue. While recent projections for non-FCS programs have generally aligned with outcomes, PRGT-eligible FCS countries have consistently underperformed for both near- and medium-term forecasts, even outside the pandemic period. Post-pandemic, one-year-ahead growth forecasts for PRGT-eligible FCS countries have exceeded actual outcomes by around 1 percentage point on average, with three-quarters of such cases showing optimism. This bias persists even in medium-term forecasts, underscoring the need for more tailored forecasting approaches in these settings.

13. Severe shocks have been the primary drivers of large growth forecast errors in IMF-supported programs, with global events affecting most countries and frequent local disruptions weighing on FCS.

To gauge the impact of large shocks on forecast accuracy, forecast errors were weighted by their historical frequency and average magnitude, and observations were grouped by the nature of the disruption—whether global, regional, or local. Historical patterns show that major global shocks—such as the GFC and COVID-19—coincided with the largest forecast misses, while in FCS settings, repeated shocks like natural disasters, health crises, and armed conflicts drove persistent underperformance. Regression analysis confirms that the pandemic alone explained at least half of the medium-term forecast deviations in the 2026 ROC period, with other shocks like wars and natural disasters adding another tenth (Figure 8, bottom charts).

B. Enhancing Resilience: Tailored Assumptions and Contingency Planning

14. To improve forecasting accuracy and program realism, the Fund is advancing efforts to refine baseline assumptions and better account recurring shocks. The realism tools under the debt sustainability framework have helped benchmark growth projections against historical outcomes and peer comparators and assess the contribution of projected growth to debt dynamics. Work is also underway to incorporate the average impact of recurring shocks, such as natural disasters, into baseline projections or adjusted downside scenarios. This shift entails adopting a probabilistic forecasting approach, where the baseline reflects a probability-weighted average of alternative macroeconomic paths. Such an approach is particularly relevant for countries with a high and well-characterized likelihood of divergent growth paths, often driven by recurrent shocks. Initial modelling efforts have focused on climate-related shocks in countries frequently affected by natural disasters (see, for example, [Seychelles 2023 EFF](#)).

15. Contingency planning has gained traction as a key risk management tool, helping countries to respond flexibly to unexpected shocks. Illustrative cases from the pandemic period show how pre-established frameworks facilitated timely adjustment. In Armenia (2019 SBA), a pre-existing contingency plan enabled swift program recalibration during the pandemic—access increased, targets relaxed, and policies eased—to support urgent health and social spending while preserving macroeconomic stability. Similarly, The Gambia’s 2020 ECF was adjusted within weeks of approval, with modified conditionality and emergency financing creating space for immediate needs. More broadly, a recent survey of IMF senior country reviewers in the Strategy, Policy, and Review Department (SPR) found that contingency plans are now standard practice, with 7 out of 10 programs including provisions to address policy slippages, growth-related shocks, or commodity price volatility. Revenue and financing shortfalls were the most frequent anticipated triggers, with several teams conducting quantitative sensitivity analyses of the most salient risks facing programs.

SECTION III. ASSESSING FISCAL POLICY CONDITIONALITY¹⁴

Fund-supported programs have adapted fiscal efforts in response to major shocks, with generally well calibrated and front-loaded adjustment paths, differentiated strategies across country groups, and more granular conditionality. While initial performance fell short amid the pandemic, recent outcomes show some improvement. Deviations from targets were largely driven by revenue shortfalls and spending pressures, with severe external shocks being the main source of fiscal slippages over time.

16. The assessment draws on program experience in the 2026 ROC, focusing on the evolution of fiscal policy frameworks in response to shocks, heightened uncertainty, and shifting macroeconomic priorities. It highlights trends in adjustment size and phasing, the composition of policy measures, the use of fiscal conditionality, and performance outcomes and realism. It distills lessons for program design, particularly in balancing stabilization needs with inclusive and growth-friendly fiscal strategies.

A. Calibrating Fiscal Adjustment to Shocks: Balancing Flexibility and Realism

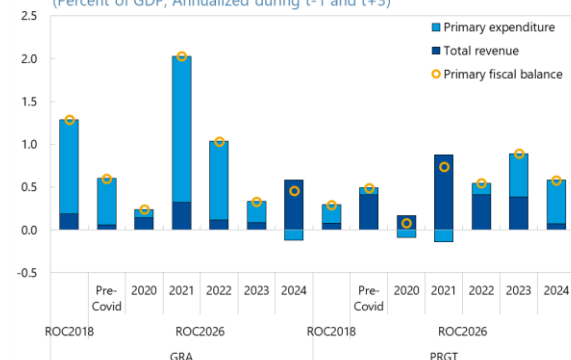
17. Fiscal targets were calibrated to account for large external shocks. On average, GRA-supported programs envisaged annual fiscal adjustment of about $\frac{1}{2}$ percent of GDP over the three years following program approval, compared to $1\frac{1}{4}$ percent in the 2018 ROC (Figure 9). PRGT programs aimed for a slightly larger annual fiscal adjustment of about $\frac{3}{4}$ percent of GDP, exceeding the $\frac{1}{4}$ percent targeted in the 2019 ROC, a pattern that holds across FCS and non-FCS. The evolution of the targets reflected the sizable impact of Covid-19. Programs approved in 2020 featured lower adjustment paths to accommodate the need for exceptional support, while those approved in 2021 incorporated higher targets to facilitate the withdrawal of pandemic-related measures. Subsequent programs generally returned to lower targets, particularly in GRA arrangements, where primary balance targets were close to debt-stabilizing levels. Fiscal consolidation targets were frontloaded in GRA-supported programs and became more frontloaded in PRGT-supported programs, with nearly half of the adjustment planned for the year of program approval and two-thirds within the first two years.

¹⁴ Prepared by Fah Jirasavetakul and Robin Koepeke, with research assistance from Juliana Dueñas Villabona, Salvador Traettino, and Helen Zheng.

Figure 9. Fiscal Adjustments – Program Design

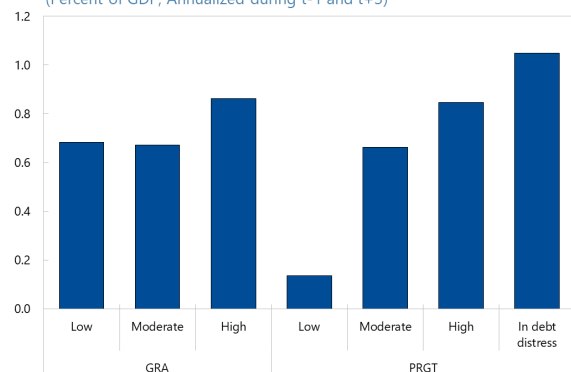
Planned adjustments incorporated the effect of Covid, with variation across GRA and PRGT programs, ...

Change in Primary Fiscal Balance, Targets
(Percent of GDP, Annualized during t-1 and t+3)



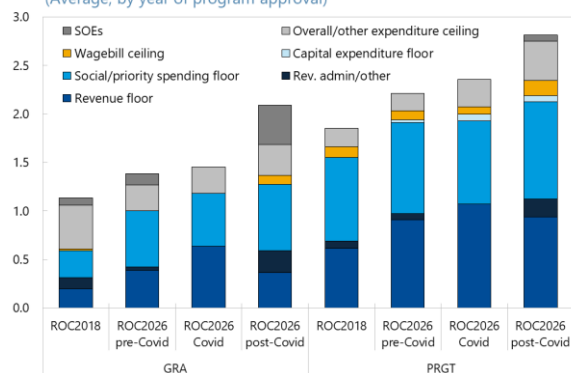
Fiscal adjustment targets have been more ambitious in programs with high debt vulnerabilities.

Change in Primary Fiscal Balance, Targets
(Percent of GDP, Annualized during t-1 and t+3)



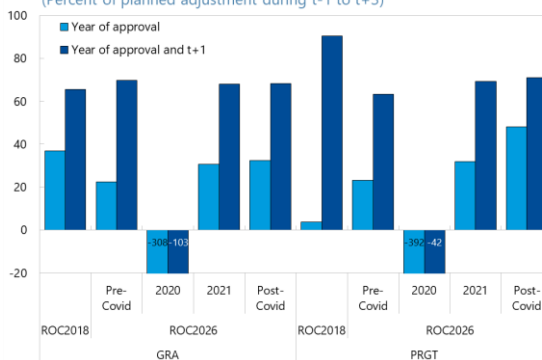
Programs increasingly adopted granular targets on priority spending in GRA and revenue in PRGT programs...

Average Number of Granular Fiscal QPC and ITS
(Average, by year of program approval)



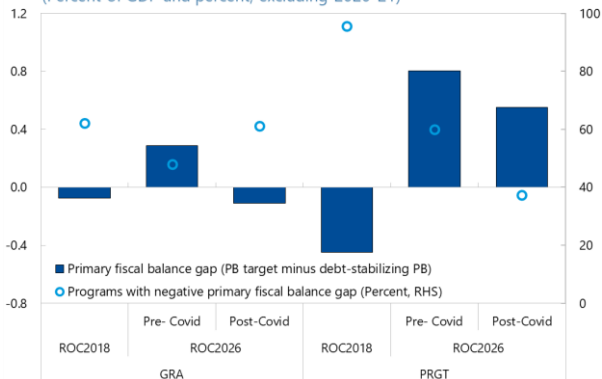
... though remain frontloaded.

Primary Fiscal Balance Adjustment Targets, Phasing
(Percent of planned adjustment during t-1 to t+3)



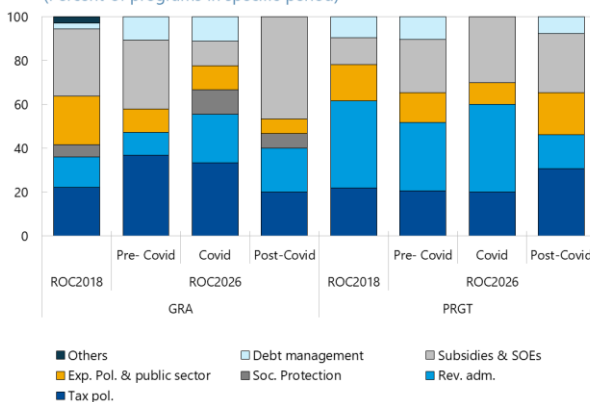
Primary balance targets broadly aimed to stabilize debt in GRA programs and reduce debt in PRGT programs.

Primary Fiscal Balance Gap: Targets vs Debt-stabilizing Levels
(Percent of GDP and percent; excluding 2020-21)



... while the focus of fiscal SCs has been more evenly distributed, with some attention to subsidies and SOEs.

Focus of Fiscal Structural Conditionality
(Percent of programs in specific period)



Sources: MONA database; IMF WEO; and IMF staff calculations.

18. Fiscal adjustment was intended to be attained through different sets of policy measures for emerging markets (EMs) and low-income countries (LICs). In GRA programs, fiscal consolidation targets have continued to rely primarily on expenditure measures, though their contribution has declined significantly in recent years. By contrast, most programs under the PRGT have shifted their adjustment strategy towards revenue mobilization. Among programs approved in the post-pandemic period, GRA arrangements aimed to facilitate higher capital spending, a trend that is less pronounced in PRGT programs. Adjustment efforts are also tailored to country-specific vulnerabilities, typically by being more ambitious in cases of high risks of debt distress.

19. Enhanced granularity has been evident through the use of detailed fiscal targets and structural conditionality. Detailed quantitative targets have generally aligned with the overarching fiscal strategy. In GRA programs, there has been an increase in the number of quantitative targets related to overall spending, social spending, and state-owned enterprises (SOEs). In PRGT programs, the focus has increasingly shifted toward targets on revenue and revenue administration. The focus of fiscal structural conditionality (SCs) has also evolved, though with a somewhat looser connection to the main adjustment priorities. Both GRA and PRGT programs have increased their fiscal SC attention to subsidy reforms while maintaining a strong emphasis on revenue measures.

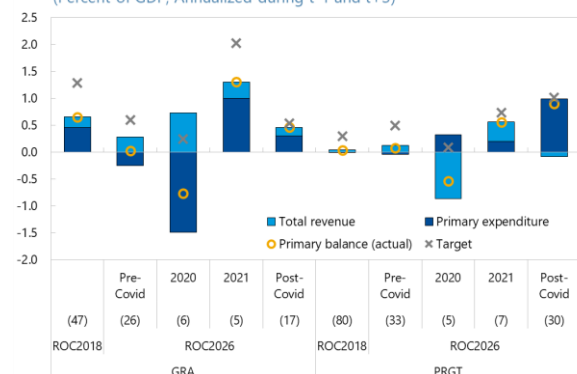
B. Fiscal Performance under Strain: Initial Shortfalls and Recent Gains

20. Actual fiscal adjustments fell short of expectations during the period of major external shocks, but performance has improved in more recent programs (Figure 10). Amid the pandemic, underperformance of fiscal adjustments among GRA programs largely reflected delayed spending rationalization as resources were directed towards Covid-related measures. Although revenue collection subsequently exceeded expectation as economic activity rebounded, the larger-than-expected spending persisted. In PRGT programs, fiscal adjustments also underperformed during the pandemic, largely driven by revenue shortfalls. While adjustment targets were generally lower for programs approved after 2021, performance has improved notably in both GRA and PRGT programs with actual annualized fiscal adjustment averaging about 0.1 percent of GDP below target. However, as many of these programs are still ongoing, a full assessment of their fiscal outcomes will require additional time and data.

Figure 10. Fiscal Adjustments – Outcomes

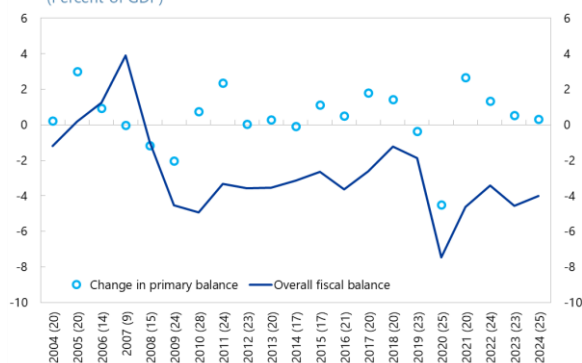
Actual adjustments fell short of expectations set at program approval, in part owing to the Covid-19 shock...

Change in Primary Fiscal Balance, Actual
(Percent of GDP, Annualized during t-1 and t+3)



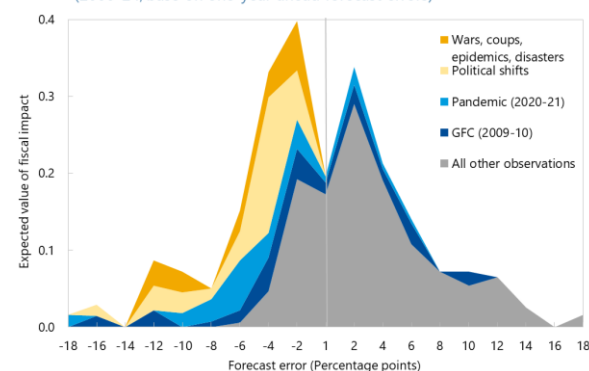
Notwithstanding significant overall fiscal deficits, post-pandemic adjustment efforts in GRA programs have generally lost momentum...

One-year Fiscal Adjustments - GRA Programs
(Percent of GDP)



A longer-term view shows that severe shocks played key roles in explaining large fiscal underperformance...

Program Countries: Impact of Shocks on Primary Balance
(2000-24; base on one-year ahead forecast errors)

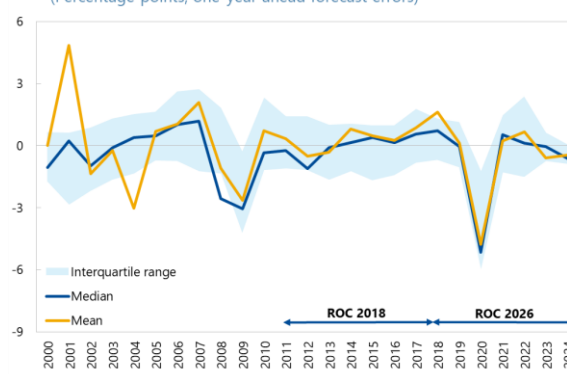


Sources: MONA database; IMF WEO; IMF staff calculations.

Notes: For the charts presenting annualized fiscal adjustment from t-1 to t+3, the time horizon is shorter for 2022–2024 due to the limited availability of actual outcomes. Specifically, these charts show annualized adjustments over t-1 to t+2 for 2022, over t+1 for 2023, and over t for 2024. Number of programs in parentheses. Shocks include political, war, coups, epidemics, and natural disaster shocks.

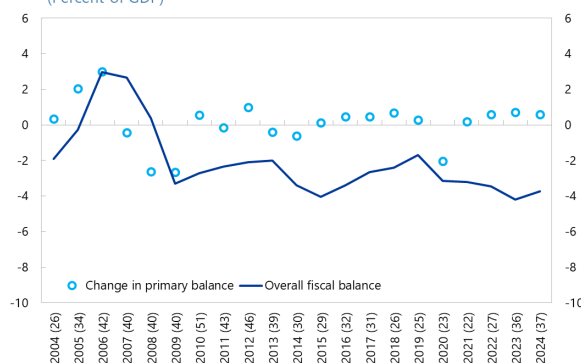
... and one-year fiscal forecast errors highlight the impact of large external shocks, which prompted a recalibration of fiscal policies.

Active Program Countries: Forecast Errors of Fiscal Balance
(Percentage points; one-year ahead forecast errors)



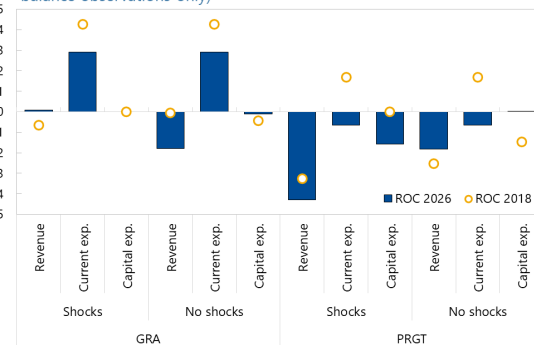
... while some adjustments have continued in PRGT programs.

One-year Fiscal Adjustments - PRGT Programs
(Percent of GDP)



... primarily through revenue shortfalls, with GRA programs able to maintain current spending while PRGT programs often had to cut both current and capital expenditures.

Median Forecast Errors of Primary Fiscal Balance
(Percent deviation from nominal target; bottom quartile of primary balance observations only)



21. Evidence from one-year-ahead forecast errors in the primary fiscal balance underscores how exogenous shocks contributed to fiscal underperformance and prompted program recalibration. Excluding 2020, fiscal outcomes were broadly in line with one-year-ahead projections, suggesting that short-term targets were generally realistic. This near-term fiscal realism, alongside some deviations from fiscal targets set at program approval, particularly in GRA programs during early-post shock recovery, points to important adaptation in program frameworks. This pattern, along with efforts to balance trade-offs between fiscal adjustment and adverse growth impacts, is further reflected in the evolution of structural primary balances, especially in GRA programs, which shows implementation shortfalls post-pandemic, though with gradual improvement over time. In PRGT programs, although both one-year-ahead and medium-term targets were largely met outside 2020, the composition of adjustment diverged from initial plans: countries frequently underperformed on revenue mobilization and instead achieved fiscal balance targets through reductions in current and capital spending. This trend was observed across both FCS and non-FCS, with performance across these groups converging somewhat in recent years.

22. A longer-term perspective over the past 25 years reveals that severe and unexpected shocks have been the primary drivers of large fiscal deviations. Most significant shortfalls in fiscal performance are linked to global events—such as the COVID-19 pandemic and the GFC—or acute country-specific shocks, including conflicts, epidemics, and natural disasters. This is consistent with the above-mentioned findings on macroeconomic realism, which highlight the role of large shocks in explaining residual growth optimism. An analysis of the bottom quartile of primary balance forecast errors sheds further light on the composition of fiscal underperformance. In years marked by shocks, shortfalls were predominantly driven by revenue underperformance. While GRA programs were generally able to maintain current spending to mitigate the impacts of shocks, PRGT programs often had to reduce both current and capital expenditures [potentially owing to tighter financing conditions]. In contrast, during non-shock years, deviations from fiscal targets were more evenly split between revenue shortfalls and slippages in current spending.

SECTION IV. ASSESSING SOCIAL CONDITIONALITY¹⁵

Social spending conditionality has increased in Fund-supported programs, with implementation rates remaining high. Social outcomes in program countries are broadly comparable to, or better than, those in non-program countries. Analysis also finds that IMF-supported programs have not systematically impacted differences in economic opportunities and outcomes between men and women.

23. The use of social spending conditionality (SSC) expanded following the adoption of the 2019 [Strategy for Engagement on Social Spending](#). The strategy's implementation is supported by various how-to notes¹⁶ and the staff Operational Guidance Note ([IMF, 2024a](#)). A key consideration for social spending conditionality (IT/PCs, SB/PAs) is its criticality in achieving the

¹⁵ Prepared by Iacovos Ioannou, Lisa Kolovich, Vivian Malta, Sanghamitra Warrier Mukherjee, Monique Newiak, and Peter Wankuru, with research assistance from Juliana Dueñas Villabona, Salvador Traettino, and Usama Zafar.

¹⁶ See How-to-Notes on health ([IMF, 2023a](#)), pensions ([IMF, 2022a](#)), education ([IMF, 2025a](#)), and social safety nets ([IMF, 2022b](#)).

program objectives or for monitoring implementation. For this analysis, social spending encompasses social protection, education, and health. Social protection distinguishes social assistance programs, such as universal and targeted transfers, child benefits, and active labor market programs that are aimed at mitigating poverty, from social insurance, notably pensions, unemployment insurance, and sickness and maternity leave to smooth income shocks. The analysis draws on data from the Monitoring of Fund Arrangements (MONA) database and World Bank social indicators while acknowledging data limitations due to inconsistent classifications and inadequate coverage in LICs.

A. Program Design and Implementation

24. The use of SSC rose slightly during the ROC 2026 period. About 88 percent of programs (116 out of 132) included some form of SSC, up slightly from 84 percent in the last ROC (Figure 11). Coverage remained higher in PRGT programs (97 percent) than in GRA program (75 percent), with no material change since the last ROC. This reflects the recognition of the impact of adjustment on inclusive growth and the need to protect vulnerable populations. SSC is most commonly delivered through indicative targets (ITs), present in 75 percent of programs. About 39 percent of programs combined ITs with structural benchmarks (SBs), while 31 percent relied solely on ITs and 9 percent used SBs alone. QPCs—either alone or combined with ITs, SBs, or prior actions (PAs)—were included in 5 percent of programs. PRGT programs relied more heavily on ITs (93 percent) than GRA programs (51 percent), while SBs were more common in GRA programs.

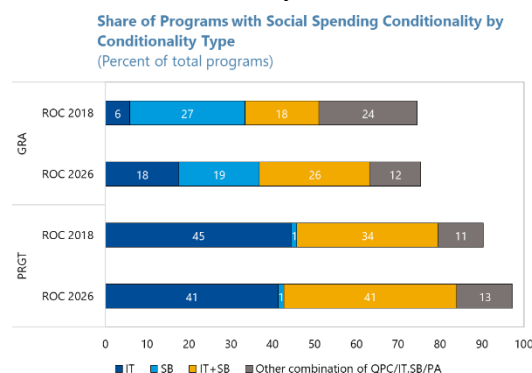
25. Social assistance has continued to dominate SSC. About 68 percent of all SSC is related to social assistance with a more pronounced trend in PRGT programs, where it accounts for 81 percent of the total (an increase from 65 percent in the previous ROC). By contrast, education, health, and social insurance-related conditionality collectively accounted for about one-third of SSC. This underscores both the near-term priority of protecting the most vulnerable and the medium-term focus on human capital investment and pension sustainability.

26. SSC implementation remained broadly consistent with past performance. During the 2026 ROC period, 70 percent of SSCs were met, mirroring the performance in the 2018 ROC. Implementation rates were highest for QPCs (100 percent) and lower for ITs (69 percent) and SBs (63 percent). Including delayed implementation, the implementation rate for SBs increased to 81 percent, reflecting capacity constraints and the challenges of implementing socially and politically sensitive reforms.

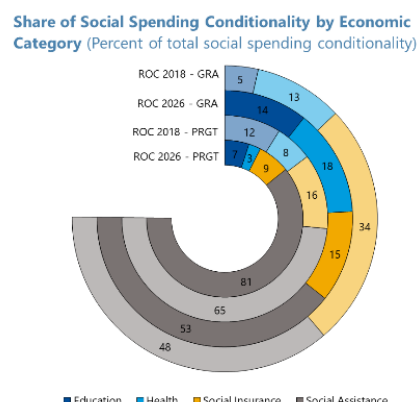
27. Across program types, performance was similar, though sectoral patterns varied. Health-related conditionality had the highest implementation rate (71 percent), followed by social assistance (70 percent), and education (69 percent). Social insurance reforms lagged significantly. Health and education conditionality was stronger in PRGT programs, whereas GRA programs saw better outcomes in social insurance-related measures.

Figure 11. Social Spending Conditionality – Main Indicators

Social spending conditionality increased marginally from the 2018 ROC to the 2026 ROC, mostly in IT+SBs...



...and conditionality on social assistance...

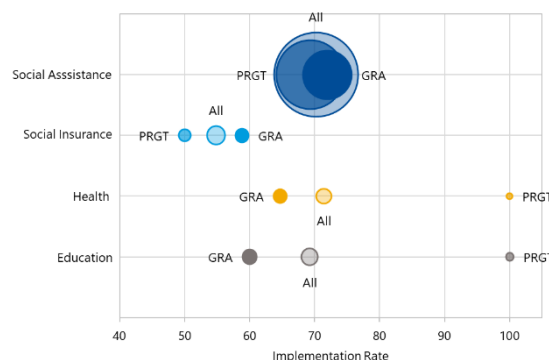


...while implementation rates increased for social assistance conditionality in GRA programs from the 2018 ROC to the 2026 ROC, they fell for pension conditionality in both GRA and PRGT programs.

ROC 2018: Implementation Rate for Social Spending Conditionality by Conditionality Type



ROC 2026: Implementation Rate for Social Spending Conditionality by Conditionality Type



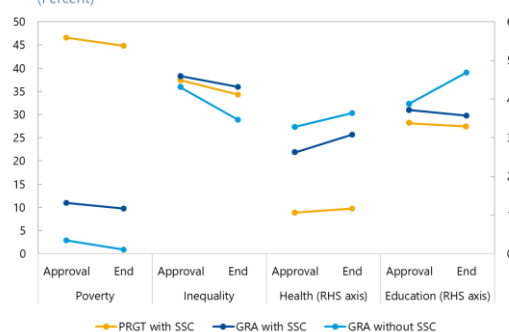
Implementation rates remained lower for SBs than ITs.

ROC 2026: Implementation Rate for Social Spending Conditionality by Conditionality Type 2/



On social outcomes, there were improvements for both GRA and PRGT with SSC, but less than for GRA without SSC (partly driven by level differences).

Changes in Socio-Economic Outcomes (Percent)



Sources: MONA database; IMF staff calculations.

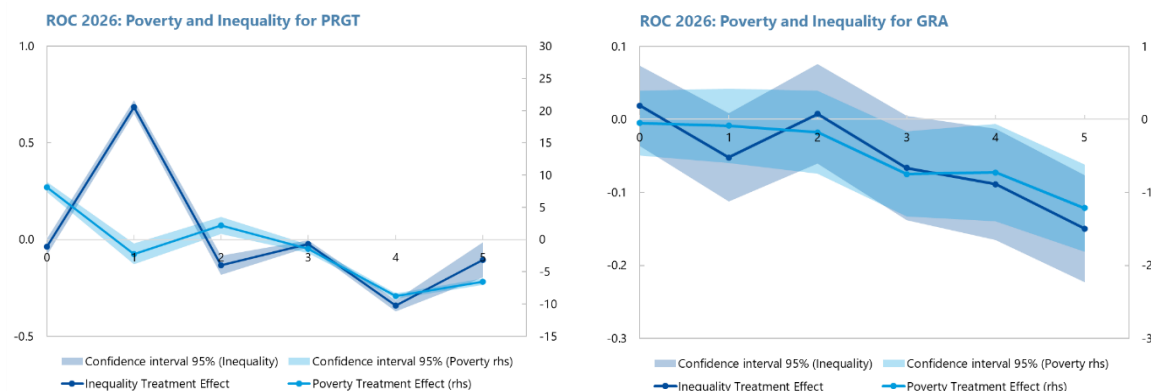
Notes: Programs may have multiple SSC classified in different categories. Social assistance includes conditionality on universal and targeted transfers, child benefits, and active labor market programs that are aimed at mitigating poverty. Social insurance includes conditionality most notably on pensions, but also other employment reforms to smooth income shocks. In the bubble charts, colors represent the conditionality type and sizes show the number of conditionalities of a given category in each of the samples. Transparency levels identify the sample (GRA 0 percent transparency, PRGT 33 percent transparency, all 67 percent transparency). Poverty and inequality are measured as percent of the population, and health and education expenditures as percent of GDP. Social assistance includes conditionality on universal and targeted transfers, child benefits, and active labor market programs that are aimed at mitigating poverty. Social insurance includes conditionality most notably on pensions, but also other employment reforms to smooth income shocks.

B. Preliminary Evidence on Social Outcomes

28. IMF-supported programs are generally associated with considerable reductions in poverty and inequality, particularly in PRGT countries, and limited effects on social spending levels¹⁷. Wankuru and others (forthcoming) analyze the dynamic effects of IMF-supported programs on poverty, inequality, and social spending from 2000 to 2022, using local projections (LP) with inverse probability weighting and regression adjusted estimators. This methodology addresses selection bias and treatment effect heterogeneity by reweighting observations based on propensity scores to create comparable groups of program and non-program countries.¹⁸ Based on World Bank data, the results show that, over the medium term, PRGT programs are associated with a 0.34 point decline in inequality and a 6.5 p.p. drop in poverty, while GRA programs see a 0.15 point decline in inequality and 1.2 p.p. reduction in poverty (Figure 12).¹⁹ By contrast, program effects on social spending are less pronounced: education spending increases slightly in PRGT countries (0.25 percent of GDP by year five), while health spending is marginally lower (0.17 percent of GDP) compared to control groups of countries. For GRA programs, differences in education and health spending are statistically insignificant. Program conditionality influences social outcomes primarily by expanding fiscal space for increased spending, promoting macroeconomic stability, and establishing substantial social spending floors, though these effects may take time to materialize.

Figure 12. Dynamic Effects of IMF-Supported Programs on Social Spending, Poverty, and Inequality

IMF-supported programs are associated with lower poverty and inequality in PRGT and GRA programs...



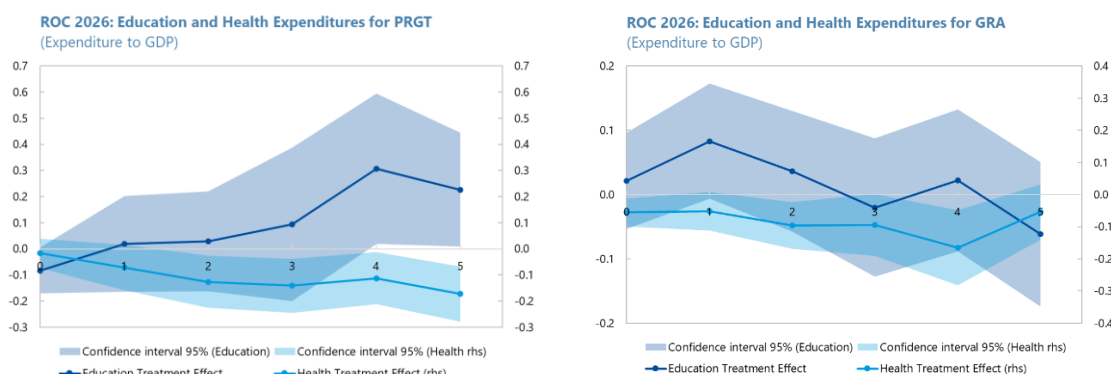
¹⁷ Social spending includes health and education spending. Due to lack of data, it does not include social assistance spending which is a significant component.

¹⁸ The LP method estimates *h-step* ahead cumulative forecasts of an outcome variable in response to treatment (approval of program) at time $t = 0$, up to 5 years out and accounts for staggered adoption using a clean control approach (Cengiz and others, 2019). The method is doubly robust to misspecification (Jordà and Taylor, 2013), requiring only either the propensity or the outcome model to be correctly specified, but not both.

¹⁹ The results hold against several robustness tests and are aligned with those found by Bird and others (2021) and Clements and others (2013).

Figure 12. Dynamic Effects of IMF-Supported Programs on Social Spending, Poverty, and Inequality (concluded)

... while program effects on social spending are less pronounced.



Sources: MONA database; IMF staff calculations and local projection regressions.

Notes: Local projections regressions and their 95 percent confidence intervals. The horizontal axis shows the years since treatment (or program approval), while the vertical axis shows the coefficients of the regressions. The controls include primary balance, current account balance, population, GDP per capita, gross external debt to GDP, urban share of population, trade openness, IMF liquidity ratio, global oil price index, electoral democracy index, country, and year fixed effects. Data from 2000 to 2022 for all countries are used in the analysis.

C. Gender in IMF-Supported Programs

29. IMF-supported programs have not systematically impacted differences in economic opportunities and outcomes between men and women. Most past studies associate IMF-supported programs with worse economic outcomes for women. Yet, these studies generally do not differentiate the impact of crises from the impact of programs. A recent paper by [Eicher and others \(2024\)](#) separates the program impact by constructing custom-tailored control groups for each of the more than 120 examined programs during 1994–2022. They find that consistently across arrangement types and time, about 9 in 10 programs do not exhibit a statistically significant impact on male versus female labor force participation (LFP) or education, health, or a composite index that captures several dimensions of inequality between men and women. For the remainder, the impacts are not consistently positive or negative.

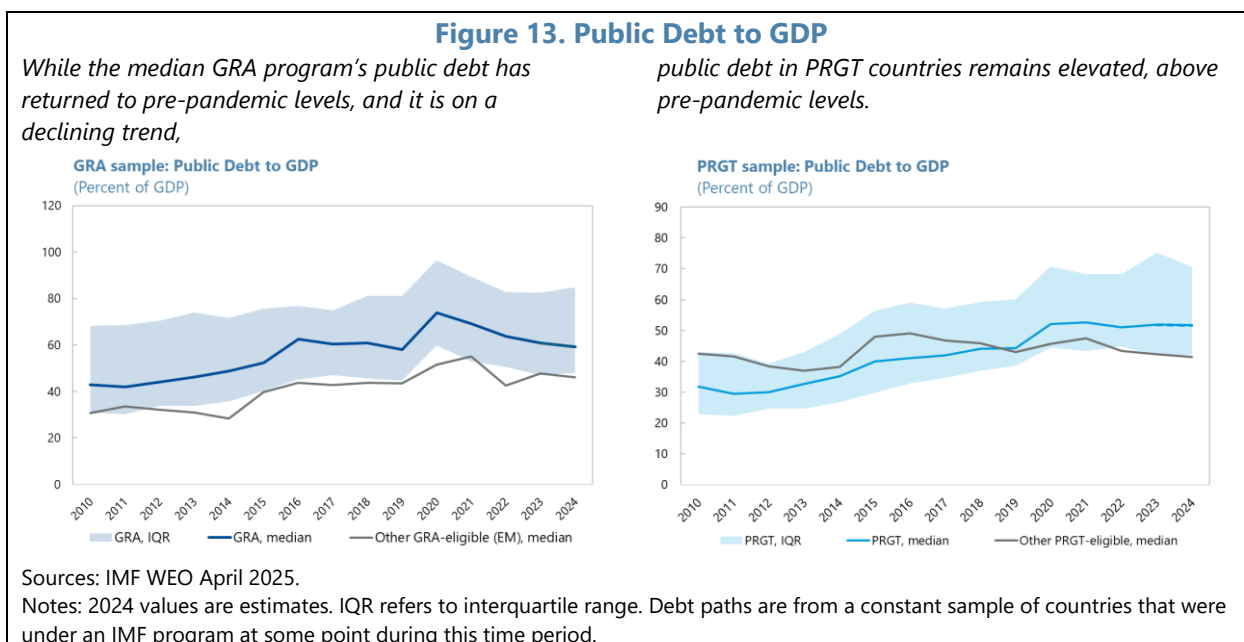
30. Since 2016, a few programs have included very limited conditionality related to women’s economic empowerment to foster stability and growth. Gender coverage in programs is in its very early stages and is considered where it is essential to achieving the core objectives of the program. Interviews with IMF staff working on nine cases with direct conditionality highlight that countries’ commitment to support stability and growth through women’s economic empowerment has generally been strong, especially when measures were phased progressively. Integrating initiatives into national frameworks (Jordan, Uganda), and leveraging existing programs to boost LFP (Pakistan, Uganda) was critical. Yet, lack of timely data complicated the monitoring of progress. Accounting for the heterogeneous impact of macroeconomic policies on the population ([IMF, 2022c](#)) can unlock further opportunities for improving the allocation of resources and supporting productivity and growth.

SECTION V. PUBLIC DEBT AND FORECAST ACCURACY²⁰

Public debt surged in many program countries during the pandemic. GRA borrowing countries have returned to pre-pandemic debt levels. PRGT borrowing countries continue to face high debt, but vulnerabilities were largely contained through strong post-pandemic fiscal efforts. Debt conditionality—particularly in PRGT programs—has supported efforts to address vulnerabilities, aided by strong implementation. Despite successive shocks during the 2026 ROC period, debt forecasting accuracy has improved, though some challenges persist.

A. Debt Trends and Risks

31. Public debt in program countries remains elevated relative to non-program peers but has broadly stabilized since its pandemic peak. In GRA borrowing countries, median debt has resumed a downward path, returning to pre-pandemic levels (Figure 13). By contrast, debt in PRGT borrowing countries remains above pre-pandemic levels despite stronger fiscal balances than their non-program counterparts—reflecting heterogeneity among PRGT countries, including fragile states and those with limited financial buffers.



32. High borrowing costs have added to debt burdens, driven by rising global interest rates and a shift in the financing mix. Funding costs have increased for all countries, particularly as global financial conditions began to tighten post-pandemic; program countries with high exposure to commercial creditors experienced the sharpest increase. Global financial tightening

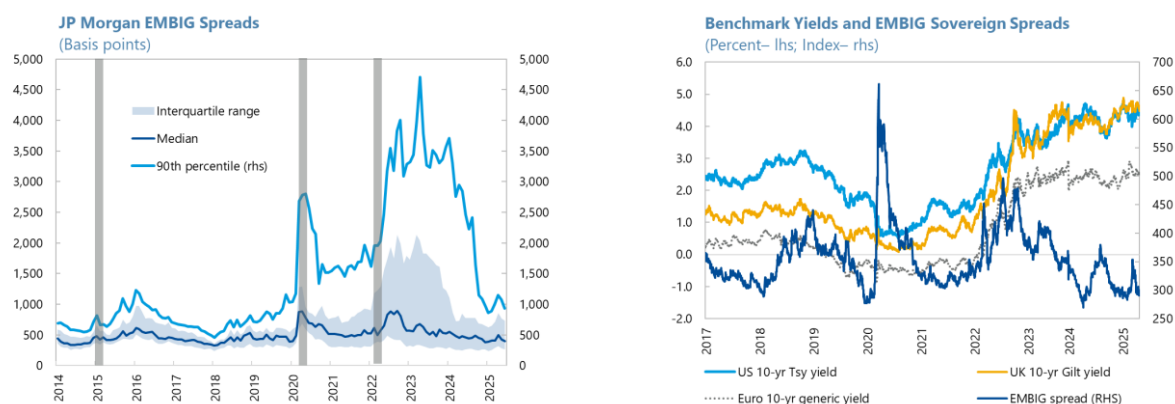
²⁰ Prepared by Anil Ari, Rimtautas Bartkus, Karina Garcia, Mahamoud Islam, and Agustin Velasquez, with contributions from Jocelyn Boussard and research assistance from Juliana Dueñas Villabona, Luisa Ballen, Chen Chen, and Joyce Saito.

prompted a flight to safety, leading to some countries losing access to international capital markets, also weighing on market re-access prospects in ongoing programs. As external financing conditions tightened, reliance on domestic financing increased significantly for both PRGT and GRA program countries. (Figure 14). In many cases, this increase was accompanied by higher costs, as domestic financing conditions also tightened in response to rising inflation and monetary policy tightening. This has frequently translated into higher interest rates, shorter maturities, and increased rollover risks, with banks absorbing a substantial share of government debt.

Figure 14. Financing Conditions and Debt Composition

Credit spreads spiked in 2022 but declined from 2024 in tandem....

... with increasing global interest rates.



Amid tighter external financing conditions, program countries turned increasingly to domestic markets.

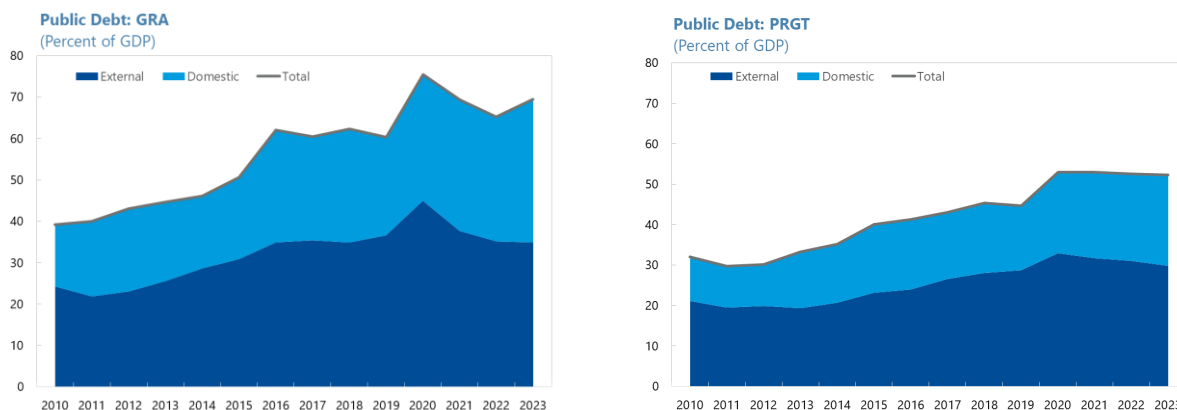
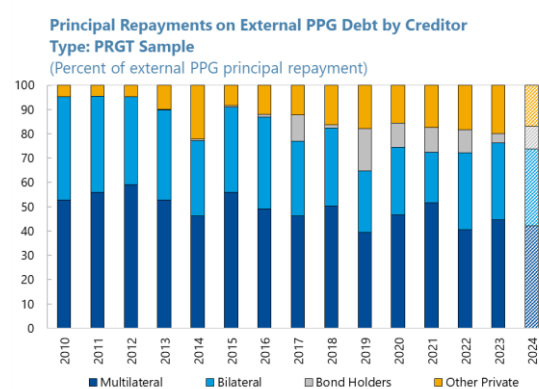
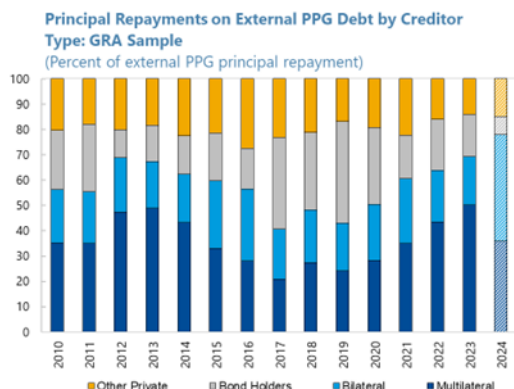


Figure 14. Financing Conditions and Debt Composition (concluded)

GRA borrowing countries have become more exposed to multilateral lenders...

...while PRGT borrowing countries' exposures to private creditors has been on the rise



Sources: Bloomberg and JP Morgan (EMBIG); IMF World Economic Outlook October 2024 and January 2025; World Bank International Debt Statistics (IDS); and IMF staff calculations.

Notes: Spreads data as of June 13, 2025. Domestic debt proxied as the difference between total and external public debt. Sample includes EMDE countries for which data is available for both total and external public debt. For the GRA sample: 12 countries out of 38 countries do not report data to IDS. 2024 onwards are IDS projections based on 2023 external debt stocks. For the PRGT sample: 12 countries out of 38 countries do not report data to IDS.

33. High debt burdens and elevated financing costs intensified debt service pressures, particularly among PRGT countries. External debt service obligations for the median PRGT country, while on the rise over the last decade, increased significantly in the post-pandemic years as has the share of obligations due to bilateral and commercial creditors. For GRA borrowing countries, refinancing risks have been much less acute than those of PRGT countries and have been partly mitigated by an increase in multilateral borrowing.

34. Debt vulnerabilities have gradually improved in PRGT programs countries and remained broadly unchanged in GRAs (Figure 15). In PRGT programs, the share of countries at high risk or in debt distress peaked during the pandemic but has declined since 2021, with the proportion at low or moderate risk returning to pre-pandemic levels. By contrast, sovereign risks in GRA programs have remained broadly stable, though nearly one third of the programs face high risk of sovereign stress. Within the program period, debt vulnerabilities evolved unevenly across both groups—improving in some cases (particularly where debt restructuring occurred) and worsening in others owing to structural weaknesses, shocks, or implementation setbacks:

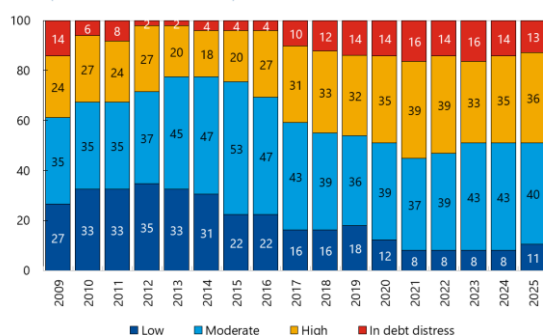
- **GRA programs:** Most saw no changes in debt vulnerabilities, with improvements in 7 percent of programs. About 60 percent started with debt assessed as "sustainable with high probability," with only a few (Ecuador 2019 EFF, Seychelles 2017 PCI, and Tunisia 2016 EFF) experiencing deterioration. One third entered a program with debt assessed as "sustainable but not with high probability", including three exceptional access programs (Argentina 2018 SBA, Argentina 2022 EFF, and Egypt 2020 SBA). Although their sustainability assessment did not worsen, several cases (Argentina, Egypt, Gabon, Ukraine) required follow-on arrangements within two years as new shocks, together with pre-existing vulnerabilities, insufficient reform momentum, and limited accumulation of buffers, contributed to slower progress.

- PRGT programs:** About 80 percent of cases saw no change in debt risk ratings; with nearly one-third remaining with elevated debt vulnerabilities throughout,²¹ including several fragile states (e.g., Afghanistan, Central African Republic, Liberia). Deterioration occurred in about 10 percent of programs, marked by program implementation delays (Ethiopia 2019 EFF-ECF, and Sierra Leone), unresolved external arrears (Sao Tome and Principe), or overreliance on domestic debt and worsening financing terms (Niger). Half of the cases where debt vulnerabilities improved—10 percent of programs—were linked to debt relief or restructuring (Somalia, Chad, Zambia).

Figure 15. Assessment of Risk of Debt Distress

Debt vulnerabilities have somewhat declined in PRGT program countries, especially post-pandemic...

Evolution of Risk of Debt Distress
(Percent of LICs with DSAs)



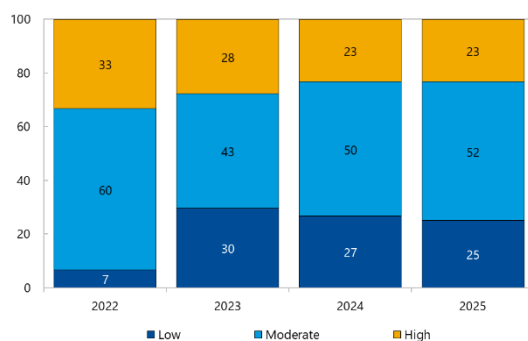
One-third of PRGT programs remained in or at high risk of debt distress throughout the program period.

PRGT ROC 2026: DSA Risk Assessment

		Total at program approval	End-program				Total at end-program
			In debt distress	High	Moderate	Low	
Program approval	In debt distress	8	4	3	1	0	6
	High	16	2	14	0	0	19
	Moderate	22	0	2	19	1	23
	Low	9	0	0	3	6	7

Debt vulnerabilities have remained broadly unchanged for GRA program countries...

Evolution of Risk of Debt Distress
(Percent of EMs with SRDSAs)



...as have debt sustainability assessments during the program period for most GRA programs.

GRA ROC 2026: Debt Sustainability Assessment

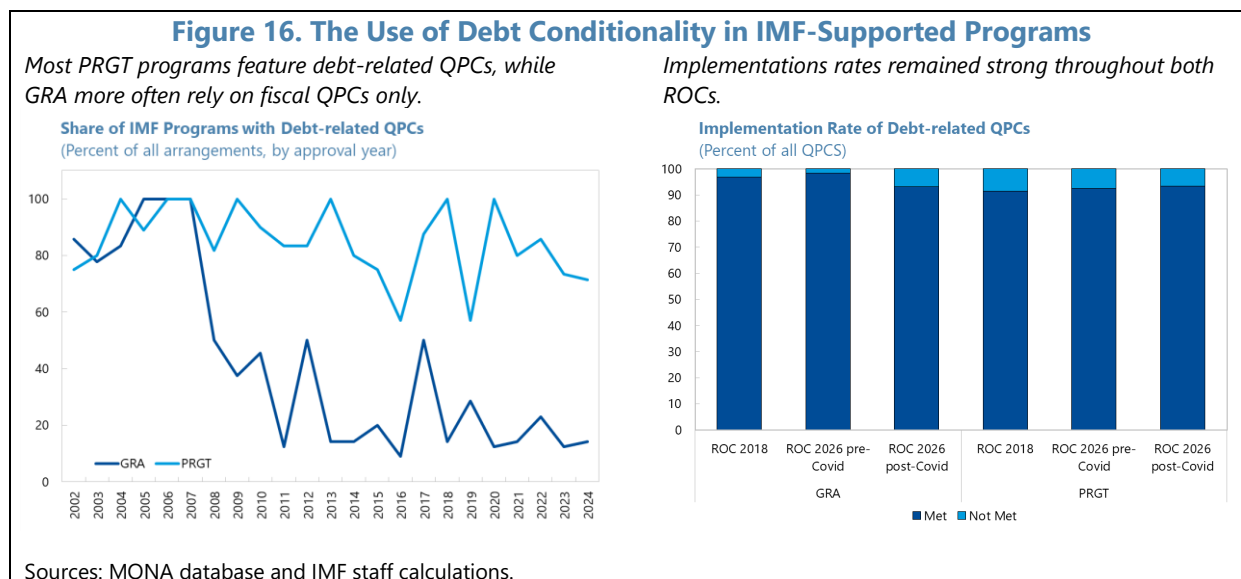
		Total at program approval	End-program			Total at program end
			Unsustainable	Sustainable but not with high probability	Sustainable with high probability	
Program approval	Unsustainable	2	0	2	0	0
	Sustainable but not with high probability	16	0	16	0	22
	Sustainable with high probability	29	0	4	25	25

Sources: LIC DSA and SRDSF database; IMF staff reports; IMF staff calculations; MONA database.

Notes: As of April 30, 2025. The analysis on GRA borrowing countries is based on the SRDSF methodology for DSA, first introduced in 2022. For 2022, the sample is composed of only 15 countries so the distribution for this year should be taken with caution. The matrices for PRGT are based on the LIC DSA overall risk of debt distress available at the beginning and at the end of each program. The matrices for GRA programs are based on sustainability assessments from SRDSFs available at the beginning and end of the program. When no SRDSF was available, sustainability assessments were taken from program documents. These exclude programs envisaged to end after 2025.

²¹ Assessed as remaining with an assessment of either in high risk or in debt distress.

35. Debt conditionality continues to support efforts to address debt vulnerabilities. QPCs on the stock or accumulation of debt—hereafter debt limits—are typically used when fiscal conditionality alone is insufficient to address existing vulnerabilities or when weaknesses in the quality and coverage of national fiscal statistics warrant their use to complement “above-the-line” fiscal conditionality. About 85 percent of PRGT programs have included debt limits, while GRA programs—particularly since the 2009 lending framework reforms²²—have typically relied on fiscal conditionality (Figure 16). Implementation remained strong throughout—above 90 percent—with increasing compliance in the PRGT and a slight dip post-pandemic in the GRA.



B. Debt Forecast Errors and Drivers

36. Debt forecast accuracy has improved over time, though large errors during the pandemic and persistent challenges—particularly in PRGT programs—remain.²³ Forecast errors were on a declining path prior to 2020, but surged during the pandemic as growth collapsed, and fiscal deficits widened. However, as economies recovered and inflation picked up, forecast errors narrowed in 2021–22 and turned slightly negative by 2023. Debt forecasts had a smaller upper tail in the ROC-2026 sample compared to ROC-2018, and in GRA programs relative to PRGT, with the latter likely reflecting data quality limitations.

²² The [2009 reform](#) allowed more flexibility in program design, and a presumption that non-LIC debt limits would not include concessionality requirements.

²³ The analysis in this section relies on a new database covering 144 IMF-supported programs (40 GRA and 104 PRGT) between 2008 and 2021 drawing from LIC, MAC DSA, and SRDSA templates. The analysis compares debt-to-GDP projections at program approval to actual outcomes three years later.

37. A regression-based approach is used to decompose forecast errors across a set of potential debt drivers.²⁴ The approach relies on the following cross-sectional regression specification:

$$D_i^{FE} = \alpha + \beta \Delta X_i^F + \gamma X_i^{FE} + \epsilon_i$$

where i denotes a cross-section of 88 programs,²⁵ D_i^{FE} represents debt forecast errors²⁶ and X_i represents a vector of debt drivers. The specification distinguishes between projected changes in debt drivers ΔX_i^F , and forecast errors X_i^{FE} in these.²⁷ A systematic link between ΔX_i^F and debt forecast errors D_i^{FE} (e.g., if an increase in primary deficit is empirically associated with higher than expected debt outturns) could be interpreted as a sign of miscalibration of the models staff rely on for debt forecasts (or “model error”) while an association between X_i^{FE} and D_i^{FE} would suggest debt forecast errors are in part due to unexpected outturns for underlying debt drivers. These are respectively captured by the estimates for the coefficients β and γ . The intercept α can then be interpreted as capturing systematic over- or under- prediction of debt-to-GDP ratios after controlling for debt drivers. Finally, ϵ_i represents the residuals.

38. The results point to systematic under-prediction of debt, driven in part by model limitations and persistent optimism. Selected debt drivers—such as structural primary fiscal balances,²⁸ the bilateral exchange rates with the U.S. dollar, trade partner real GDP growth rates, long-term real interest rates in advanced economies (AEs), and oil prices interacted with pre-program oil trade balances—are statistically significant with expected signs.²⁹ Notably, forecasts for

²⁴ The analysis focuses on empirical associations and their decomposition and should not be interpreted causally given likely endogeneity and omitted variable bias.

²⁵ From the dataset of 144 programs, 22 programs with a debt restructuring in the forecast period and 34 programs with missing data for debt drivers were excluded. Controlling debt restructurings with a binary variable was also considered, but ultimately not preferred as the sub-sample of debt restructuring cases contained several large positive and negative outliers in terms of debt forecast errors.

²⁶ D_i^{FE} is defined based on forecasts at program approval and at a 3-year horizon as described above, i.e., if a program i is approved at year T , then $D_i^{FE} = D_{i,T+2} - D_{i,T+2|T}^F$ where $D_{i,T+2|T}^F$ is the forecast at T for debt-to-GDP at $T + 2$.

²⁷ Similar to D_i^{FE} , ΔX_i^F and X_i^{FE} are defined with respect to program approval forecasts and at a 3-year horizon after program approval. For debt drivers that are flow variables (e.g., structural primary fiscal balances), ΔX_i^F is defined as the forecast average change in X_i relative to pre-program over the 3-year horizon, i.e., $\Delta X_i^F = \frac{X_{i,T|T} + X_{i,T+1|T} + X_{i,T+2|T}}{3} - X_{i,T-1}$ where T is the program approval year. For debt drivers that are not flow variables (e.g., exchange rates), ΔX_i^F is defined multiplicatively and at end-horizon, i.e., $\Delta X_i^F = 100 \frac{X_{i,T+2} - X_{i,T-1}}{X_{i,T-1}}$. Debt driver forecast errors are similarly defined additively and as an average over 3-years for flow variables $X_i^{FE} = \frac{X_{i,T} + X_{i,T+1} + X_{i,T+2}}{3} - (X_{i,T|T} + X_{i,T+1|T} + X_{i,T+2|T})$ and multiplicatively and at end-horizon for stock variables $X_i^{FE} = 100 \frac{X_{i,T+2} - X_{i,T-1}}{X_{i,T-1}}$.

²⁸ Structural primary fiscal balances are adjusted for the economic cycle and one-off (e.g., pandemic support) spending measures.

²⁹ Higher-than-expected structural primary fiscal balances and real interest rates in advanced economies are associated with higher debt forecast errors (i.e., under-prediction of debt-to-GDP ratios), while stronger-than-expected exchange rates and trade partner GDP growth (a proxy for exogenous export demand shocks) are associated with lower debt forecast errors. Higher-than-expected oil prices are associated with higher (lower) debt forecast errors for net importers (exporters) of oil. A range of other debt drivers (e.g., terms of trade, the VIX volatility index, and broad money growth rates) were considered but ultimately not selected due to weaker associations with debt forecast errors and/or less data availability than the selected drivers.

trade partner growth are negatively associated with debt forecast errors, implying that staff systematically under-estimate their impact on debt. Owing to these “model errors,” changes in trade partner growth are associated with debt forecast errors even when the changes are predicted correctly.³⁰ In addition, the regression intercept is positive and statistically significant, suggesting systematic under-prediction of debt-to-GDP ratios after controlling for debt drivers and forecast errors.

Table 1. Regression Results

Variables	Debt forecast error	
	Est. coefficient	P-value
Forecast change:		
Structural primary fiscal balance	0.66	0.29
Long-term real interest rates in advanced economies	3.51	0.11
Exchange rate with USD (+ = appreciation)	0.07	0.50
Trade partner GDP growth rate	-1.38***	0.00
Oil price x pre-program oil trade balance	-0.03	0.14
Forecast error:		
Structural primary fiscal balance	-2.10***	0.00
Long-term real interest rates in advanced economies	5.21***	0.00
Exchange rate with USD (+ = appreciation)	-0.35***	0.00
Trade partner GDP growth rate	-1.49**	0.05
Oil price x pre-program oil trade balance	-0.03*	0.05
Intercept	17.95***	0.00
Number of observations	88	
R-squared	0.39	

Sources: IMF WEO; GEE and GAS databases; IMF staff calculations.

Notes: OLS regression with robust standard errors. *** P<0.01, ** P<0.05, * P<0.1. Excludes programs in the database with missing data for debt drivers and with a restructuring in the forecast period.

39. The decomposition indicates that debt forecast errors are explained by shortcomings in (pre-program) surveillance, program design and implementation, and exogenous shocks that materialize during programs (Figure 17).

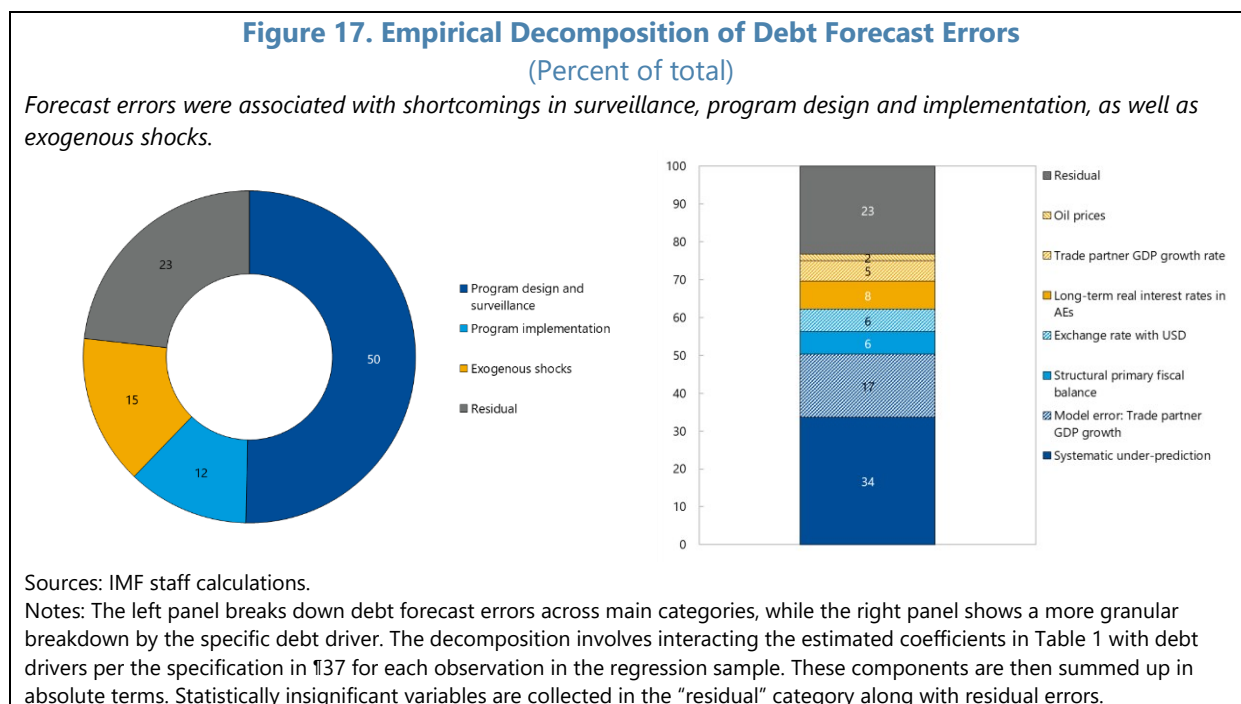
- **Surveillance and program design.** A third of debt forecast errors are explained by systemic under-prediction of debt, which could indicate optimism bias or a tendency for previously unrecognized debt obligations to be incorporated into debt outturns after program approval.³¹ About a fifth are explained by “model errors” associated with the systemic under-estimation of the impact of trade partner growth rates on debt. Together, these suggest that nearly half of debt forecast errors are related to systemic shortcomings in surveillance and program design.
- **Program implementation.** Forecast errors in structural primary fiscal balances and the exchange rate each explain about 6 percent of total forecast errors (i.e., 12 percent together)

³⁰ The forecast changes for trade partner growth are based on the IMF’s GEE database. Another possible interpretation for “model errors” is that staff’s debt forecasts do not incorporate these inputs, or are based on different, less accurate, forecasting assumptions.

³¹ Debt forecast data are unlikely to be driven by changes in the debt perimeter as they are adjusted for backward revisions in the data after program approval. Optimism bias could also arise due to ambitious reform assumptions in program forecasts. Consistent with this, [Ho and Mauro \(2016\)](#) find greater optimism bias in growth forecasts made prior to a country entering an IMF-supported program.

and could be interpreted as reflecting systemic shortfalls in implementing programmed adjustment policies.³²

- **Exogenous shocks.** About 15 percent of debt forecast errors can be attributed to shocks emanating from oil prices, long-term interest rates in AEs and trade partner GDP growth which are exogenous to macroeconomic conditions in the program country.
- **Residual.** About 25 percent of forecast errors are unexplained, possibly reflecting unobserved factors.



40. The decomposition results are largely robust across sub-samples (Table 2). There are two notable exceptions. First, the ROC 2026 sub-sample attributes a larger role to forecast errors in long-term real interest rates in AEs, driven by the COVID period. Second, model errors associated with underestimation of impacts from trade partner growth have a larger role in PRGT arrangements than in GRA.

³² Forecast errors in exchange rates can also be due to external shocks not fully captured by the trade partner growth, oil price and AE long-term interest rate variables in the regression.

Table 2. Decomposition for Different Sub-Samples

	Full sample	GRA	PRGT	FCS	FCS- PRGT	ROC 2018	ROC 2026	COVID (2020-23)	Excl. COVID
Program design and surveillance	50	48	52	52	51	47	53	53	52
Systematic under-prediction	34	34	33	34	34	32	35	34	34
Model error: Trade partner growth	17	14	18	18	17	15	18	18	19
Program implementation	12	12	12	12	11	13	11	11	11
Structural primary fiscal balance	6	7	5	5	6	7	5	5	5
Exchange rate with USD	6	5	6	6	5	5	6	6	6
Exogenous shocks	15	16	14	11	16	20	11	14	15
Long-term real interest rates in AEs	8	8	7	4	9	12	4	8	8
Trade partner growth	5	5	6	5	6	6	5	6	6
Oil prices	2	3	1	2	1	2	2	1	1
Residual	23	24	23	25	22	20	25	22	22
Observations	88	32	56	27	25	57	61	34	54

Sources: IMF staff calculations.

Notes: The decomposition involves interacting the estimated coefficients in Table 1 with debt drivers per the specification in ¶37 for each observation in the regression sample. These components are then summed up in absolute terms for the full sample and across a range of sub-samples. Statistically insignificant variables are collected in the “residual” category along with residual errors. The 2020-23 in COVID (2020-23) refers to forecast years. PRGT includes blenders.

SECTION VI. EVOLVING MONETARY POLICY CONDITIONALITY³³

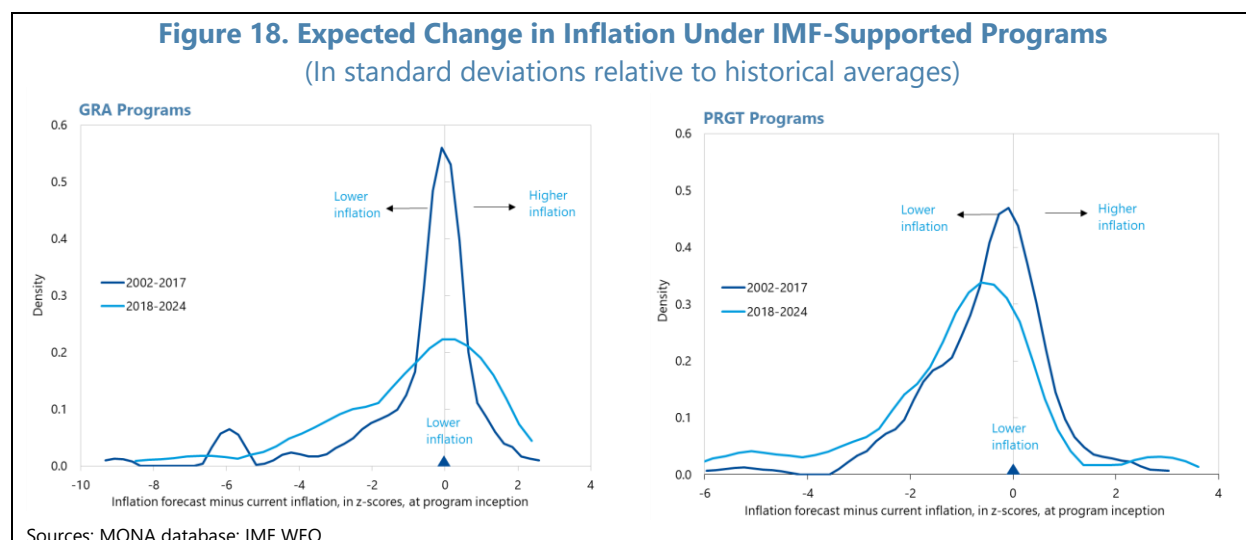
In this ROC period, IMF-supported programs faced a significantly more inflationary global backdrop relative to the previous two decades. A synthetic control-based analysis finds that most programs achieved lower medium-term inflation outcomes than “no-program” counterfactuals. However, results varied significantly across countries, partly reflecting differences in policy regimes and the strength of the nominal anchors, with exchange rate regimes being a key driver of heterogeneity. Drawing on newly collected data, this section also assesses the experience with review-based monetary policy conditionality (RBC). Well-calibrated RBC provides flexibility to countries with sufficiently mature policy frameworks, while institutional reforms to support RBC are critical for those with evolving frameworks.

A. Inflation Outcomes and Drivers

41. IMF-supported programs during the ROC 2026 period were implemented amid severe global shocks and high inflation volatility, prompting a greater focus on disinflation. With inflation well above historical norms in many countries, programs—particularly in PRGT cases—increasingly focused on disinflation, marking a departure from earlier periods (2002–2017) when programs more commonly focused on inflation stabilization (Figure 18). This shift reflected not only

³³ Prepared by Yazan Al-Karablieh, Thomas J. Carter, and Mahmoud Harb, with research assistance from Juliana Dueñas Villabona, Sumaiyah Mirza, Salvador Traettino, and Helen Zheng.

the scale of the inflation challenge but also the need to restore policy credibility and anchor expectations amid heightened global uncertainty.



42. Despite the difficult context, many programs achieved lower medium-term inflation outcomes than implied by “no-program” counterfactuals. Results from a synthetic control method (SCM) shows that, by the fifth year, inflation for the median ROC 2026 program was modestly below a “no-program” counterfactual, with a larger difference in PRGT countries (Figure 19).³⁴ However, outcomes were highly heterogeneous: interquartile ranges reveal that a substantial number of programs experienced medium-term inflation outcomes higher than suggested by a “no-program” benchmark.

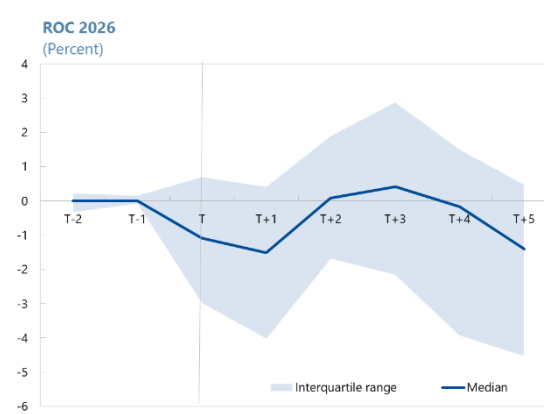
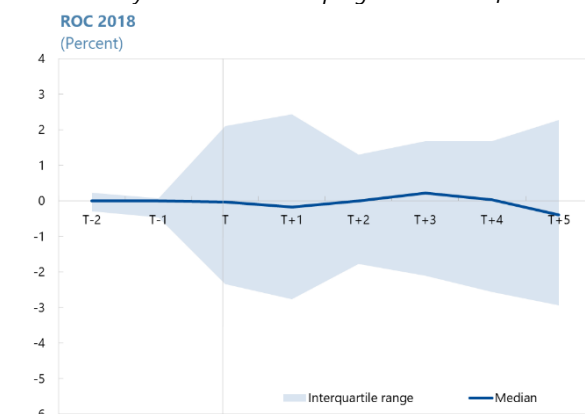
³⁴ Counterfactuals were constructed using a SCM, following [Newiak and Willems \(2017\)](#), [Duncan and others \(2022\)](#), [Kuruc \(2022\)](#), and [Bhalla and others \(2023\)](#). The SCM matches each program to a pool of non-program countries that: (i) did not enter programs within multi-year windows before or after the program under consideration, but (ii) share similar development levels, monetary and exchange rate policy frameworks, and trade profiles with the program country. A weighted combination of these matched countries was chosen to replicate the program country’s inflation path up to program start but is allowed to diverge thereafter. Since the SCM does not control for domestic shocks that may have contributed to program entry, early deviations between actual and counterfactual outcomes are difficult to interpret. The analysis is therefore focused on a five-year horizon over which program policies should have adapted to initial conditions.

**Figure 19. Inflation Outcomes Under IMF-Supported Programs,
Relative to “No-Program” Counterfactuals**

(Difference between actual and counterfactual outcomes for headline inflation)

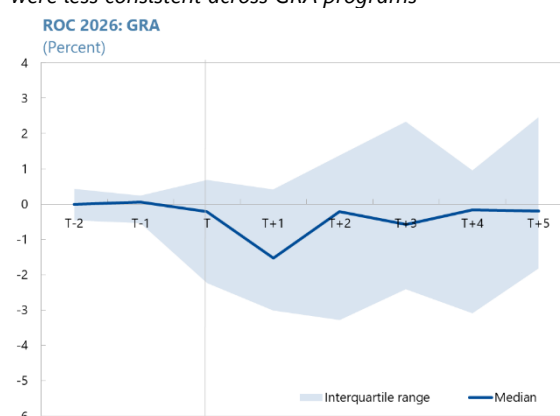
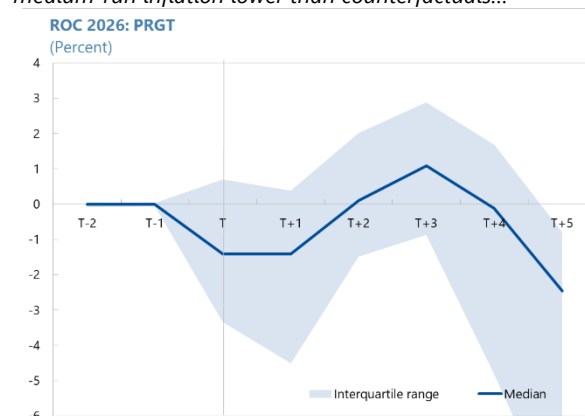
For the median ROC2018 program, medium-run inflation was modestly lower than a “no-program” counterfactual...

...and even lower for the median ROC2026 program.



Within ROC2026, most PRGT programs experienced medium-run inflation lower than counterfactuals...

...while medium-run outcomes relative to counterfactuals were less consistent across GRA programs



Sources: MONA database; IMF AREAER; IMF WEO; IMF staff calculations.

Notes: The results are derived from a synthetic control analysis comparing inflation outcomes under IMF-supported programs to counterfactual scenarios constructed from weighted combinations of non-program countries with similar pre-program characteristics. A positive value indicates that inflation was higher under a given program than in the corresponding “no-program” counterfactual, while a negative value implies lower inflation under a given program.

43. Exchange rate regimes and nominal anchors played key roles in explaining the heterogeneity in inflation outcomes (Figure 20). Analysis based on a pooled sample of ROC 2018 and ROC 2026 programs compares programs initiated under established floating regimes or pegs against ones initiated under “intermediate or transitional regimes” (i.e., regimes that fall between conventional pegs and floating arrangements, or were in flux prior to program start).³⁵ The results suggest that intermediate and transitional regimes tended to exhibit the widest dispersion, with

³⁵ The breakdown is based on countries’ de-facto AREAER classifications over the three years prior to program start, setting aside cases involving currency boards or no separate legal tender. “Intermediate” refers to arrangements that fall in between conventional pegs and floating arrangements, while “transitional” refers to cases where the exchange rate arrangement changed over the three-year window under consideration.

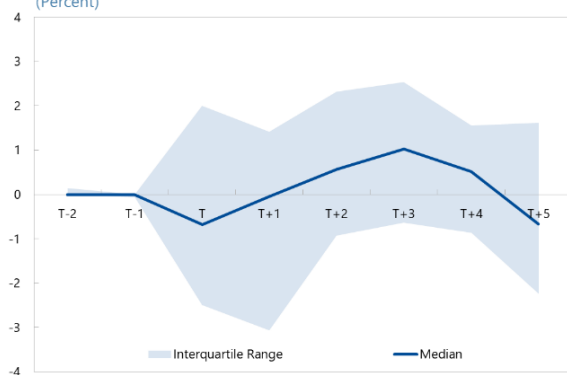
many experiencing substantially higher medium-term inflation relative to “no-program” counterfactuals. Conversely, conventional pegs and floating arrangements showed less dispersions, with the latter achieving a noticeably lower median inflation trajectory. Additionally, programs initiated amidst large depreciations displayed higher dispersion in their inflation outcomes relative to “no-program” counterfactuals.³⁶ Together, these results point to challenges that programs faced when containing the inflationary effects of large depreciations and stabilizing inflation in the absence of clear, established anchors.

Figure 20. Inflation Outcomes Under IMF-Supported Programs, Relative to “No-Program” Counterfactuals: Role of the Exchange Rate and Policy Regime

(Difference between actual and counterfactual outcomes for headline inflation)

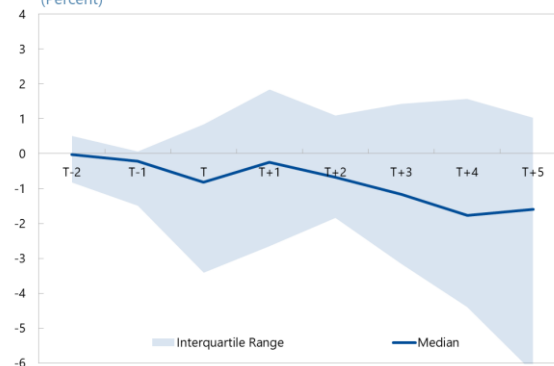
Medium-run outcomes were lower than “no-program” counterfactuals for programs starting from pegs...

ROC 2018 & ROC 2026: Conventional Pegs
(Percent)



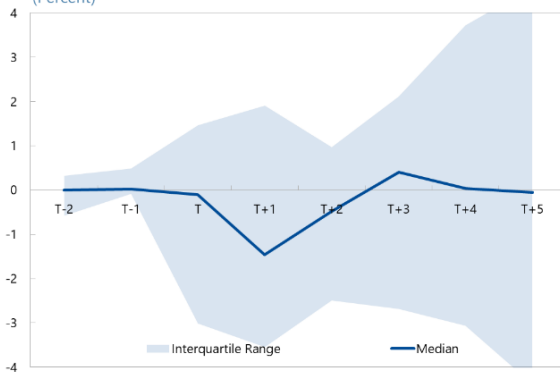
...and floating regimes...

ROC 2018 & ROC 2026: Floaters
(Percent)



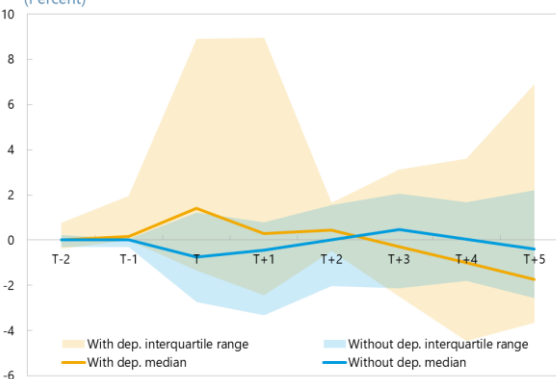
...but higher and more dispersed for programs starting from intermediate or transitional exchange rate regimes

ROC 2018 & ROC 2026: Intermediate and Transitional Regimes
(Percent)



Outcomes relative to counterfactuals were also dispersed for programs associated with sharp depreciations

ROC 2018 & ROC 2026: Programs With and Without Sharp Depreciations
(Percent)



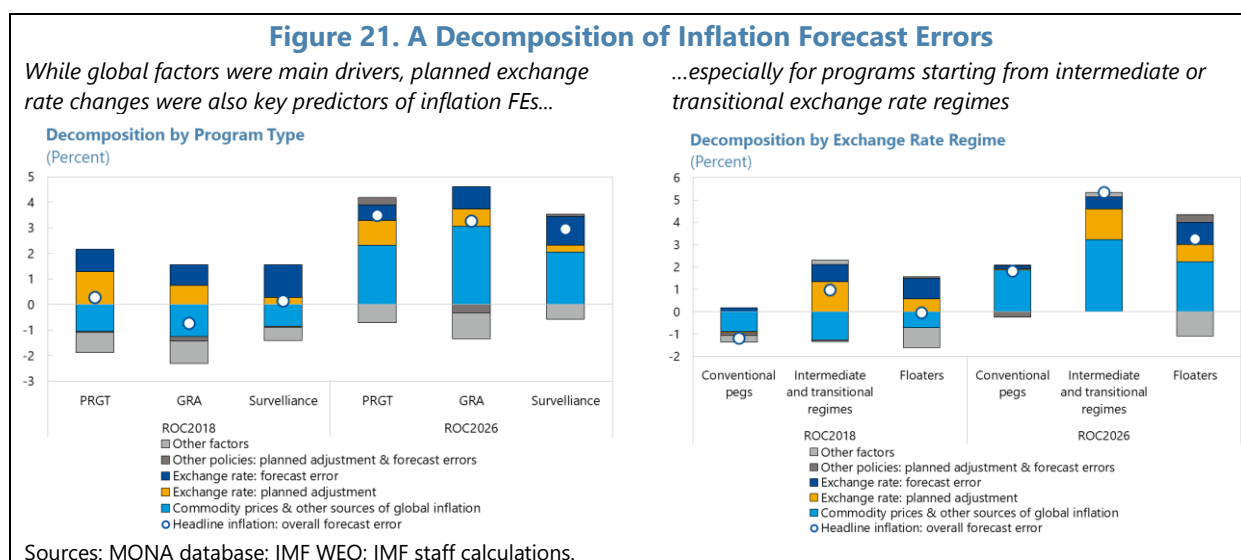
Sources: MONA database; IMF AREAER; IMF WEO; IMF staff calculations.

Notes: For details, see notes in Figure 19.

³⁶ Sharp depreciation episodes were identified using a methodology detailed in Section VII, and programs were then categorized depending on if an episode started within a plus-or-minus-two-year window relative to program start.

44. Regression analysis of two-year ahead inflation forecast errors (FEs) under IMF-supported programs underscores the central role of the exchange rate (Figure 21). Overall FEs, which were small during ROC 2018, but turned positive and sizeable during ROC 2026, averaging over 3 percentage points for both PRGT and GRA programs—modestly higher than a control group of surveillance cases.³⁷ A regression analysis was used to examine the extent to which these FEs can be attributed to: (i) *global factors* such as commodity prices; (ii) the *planned adjustments in the exchange rate and domestic fiscal and monetary policy* that were embedded in program design; and (iii) *ex-post deviations of the exchange rate and domestic policy variables* from their planned trajectories (which could reflect shocks and/or policy slippages).³⁸ Key results include the following:

- **Global factors**, particularly commodity prices and supply-chain pressures, were the main driver of the increased FEs during ROC 2026.
- **Planned exchange rate adjustments** were the next largest source of error for many programs, particularly PRGT and programs initiated in the context of “intermediate” or transitional exchange rate regimes, indicating that program design may have underestimated the potential for exchange rate pass-through in the absence of strong nominal anchors.
- **Unanticipated exchange rate movements**, especially in floating regimes and surveillance countries, contributed further to errors.
- **Planned and unanticipated changes in domestic policy variables** generally played smaller roles but were still notable contributors for PRGT programs, suggesting challenges in policy design and implementation in lower-capacity settings, especially amidst elevated uncertainty.



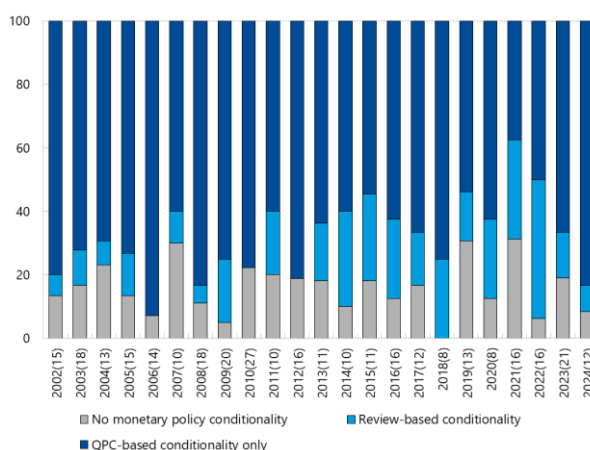
³⁷ These averages and all results in Figure 21 were computed after dropping hyperinflation episodes (i.e., annual inflation over 100 percent).

³⁸ The analysis builds on [Blanchard and Leigh \(2013\)](#), [IMF \(2019\)](#), and [Ismail and others \(2020\)](#) and uses money growth and the primary fiscal balance as a share of GDP as proxies for domestic monetary and fiscal policy, respectively.

B. Experience with Review-Based Monetary Policy Conditionality

45. In light of evolving inflation dynamics and the limitations of quantity-based targets, the Fund has introduced various forms of review-based monetary policy conditionality (RBC).³⁹ Historically, IMF-supported programs relied on ceilings on the central bank's net domestic assets (NDA) or base or reserve money (BRM) to support price stability. However, as domestic monetary policy frameworks matured and the link between monetary aggregates and inflation weakened (IMF 1999, 2014), the effectiveness of these targets diminished in many countries. In response, the Fund introduced RBC in 1999 through Inflation Consultation Clauses (ICCs) for inflation-targeting countries. RBC was then expanded in 2014 with Monetary Policy Consultation Clauses (MPCCs) for countries with evolving monetary policy frameworks that either have good track records, supported by technical and institutional development, or have committed to substantial strengthening the monetary policy framework. In such cases, RBC can enhance flexibility while maintaining accountability and shifting the focus of monetary policy conditionality toward key policy outcomes. Under RBC, conditionality is typically centered on inflation, framed by outer, and possibly inner, tolerance bands.⁴⁰ A breach of the outer band triggers a mandatory consultation with the IMF's Executive Board, interrupting access to Fund resources until the consultation has taken place. A breach of the inner bands triggers a consultation with IMF staff only. RBC is tailored to country-specific circumstances and is sometimes supplemented with QPCs or ITs on the central bank balance sheet, where relevant (IMF, 2024b).

Figure 22. Monetary Policy Conditionality under IMF-Supported Programs (Share of total annual program approvals)



Sources: MONA database; IMF staff reports.

Notes: The numbers in parentheses represent the number of programs approved each year.

46. The use of RBC increased following the introduction of MPCCs in 2014 and has averaged around 24 percent of approved programs during the past five years (Figure 22).

Since 1999, RBC has been integrated into 51 of the 376 IMF-supported programs across 29 countries, primarily middle-income economies. ICCs were most commonly applied to inflation targeters with floating exchange rate arrangements, while MPCCs were applied to a more diverse set of policy regimes (Figure 23). Adoption of RBCs in PRGT programs has been more limited, though some countries—such as Ghana, Rwanda, and Uganda—have benefited from successive arrangements that alternated between ICC and MPCC frameworks.

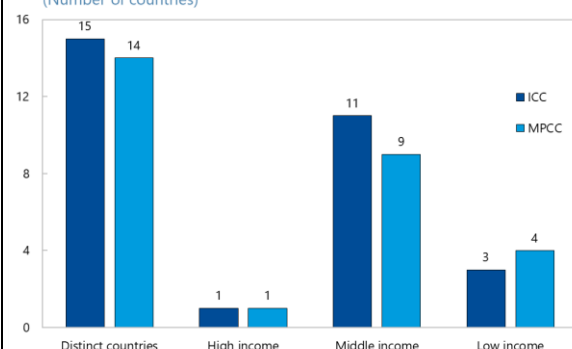
³⁹ This subsection is based on an in-depth analysis of 291 staff reports covering all reviews of 51 programs with review-based monetary policy conditionality until end 2024, and on granular data for monetary policy conditionality in 264 programs from the Monitoring of Fund Arrangements and the IMF's Finance Department databases.

⁴⁰ For MPCCs, bands can be specified either for the inflation rate or for a monetary aggregate.

Figure 23. Distribution of ICC and MPCC Programs by Income Level and Policy Regime

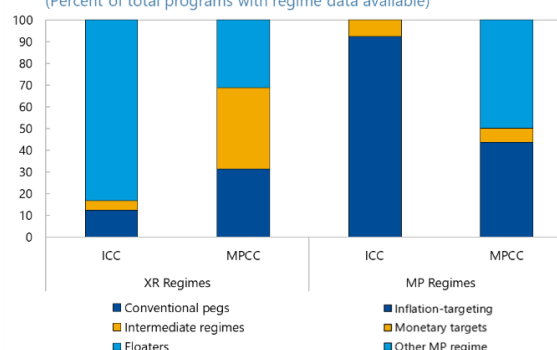
Use of RBC has been most common among middle income countries...

Use of RBCs in IMF-Supported Programs, by Country Income Level
(Number of countries)



...with MPCCs applied to a more diverse set of policy regimes, relative to ICCs

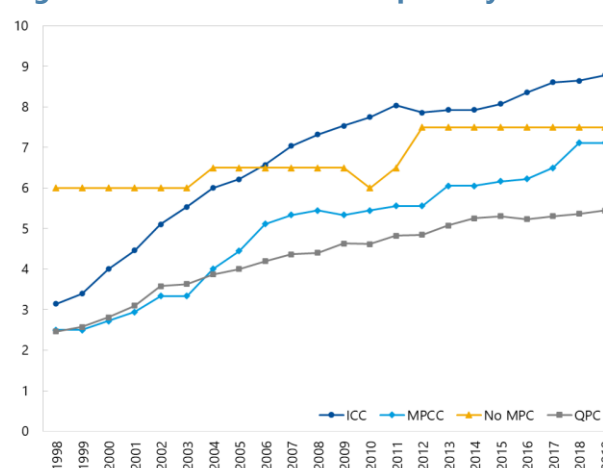
Monetary Policy Regimes at Approval of IMF-Supported Programs with ICC and MPCC
(Percent of total programs with regime data available)



Sources: MONA database; IMF AREAER database; IMF, WEO; and IMF staff calculations.

47. The design of RBC has varied across programs, reflecting country-specific circumstances. Nearly all RBC programs have used headline CPI as the anchor variable.⁴¹ About two-thirds of arrangements included both inner and outer consultation bands, enabling a graduated enforcement mechanism. Most of the remaining programs featured only outer bands, potentially limiting the scope for early warning, while a small number of programs only incorporated inner bands. Roughly one-third of programs also included supplementary QPCs.

48. Institutional reform efforts often preceded and accompanied the inclusion of RBCs in programs. Many countries that benefited from RBC arrangements, including Albania, Armenia, Egypt, Ghana and Sri Lanka, have made notable progress in various reforms, as reflected in rising scores on the Central Bank Transparency Index (Figure 24).⁴² These

Figure 24. Central Bank Transparency Index

Sources: Transparency Index from [Dincer and others \(2019\)](#); IMF staff calculations.

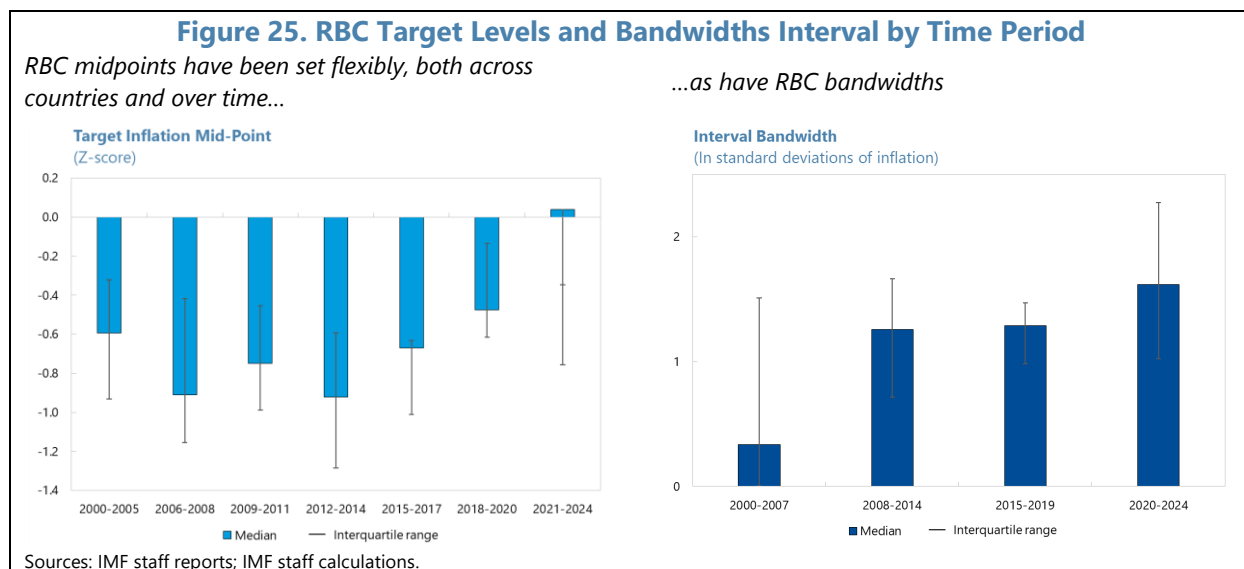
Notes: Series are calculated using simple averages of the indices for countries in each group. "ICC" and "MPCC" groups refer to countries with IMF-supported programs that included an Inflation Consultation Clause or Monetary Policy Consultation Clause, respectively, during the sample period. "No MPC" refers to countries without any monetary policy conditionality in the sample years. "QPC" refers to countries with programs that included monetary QPCs but neither an ICC nor an MPCC the sample period.

⁴¹ Dual bands were included in 35 out of 51 programs with review based monetary policy conditionality covered in this analysis.

⁴² A caveat, however, is that some countries—such as Albania—operated under MPCCs during periods of relatively low inflation, which may have eased implementation challenges.

institutional improvements helped lay the foundation for the effective use of RBC frameworks within IMF-supported programs.⁴³

49. The calibration of target levels and tolerance bands under RBC has evolved with inflation trends. Earlier programs emphasized disinflation, setting midpoints materially lower than the average levels of inflation prevailing prior to program approval. More recently, midpoints have moved closer to pre-program averages and tolerance bands have widened (Figure 25). These trends partly reflect decisions to provide countries with greater flexibility—especially over the post-pandemic period—combined with improvements in initial conditions. Programs launched during high inflation periods (e.g., Armenia’s 2022 SBA and Serbia’s 2022 SBA) typically set initial higher targets that were gradually reduced as conditions improved. By contrast, programs initiated during low-inflation periods (e.g., Rwanda’s 2019 PCI) tended to maintain stable targets. The width of outer tolerance bands also differed significantly across programs, from narrow ranges of around ± 2.5 – ± 4 percent (e.g., Armenia’s 2022 SBA, Costa Rica’s 2021 EFF, Serbia’s 2021 PCI) to much wider—and in some cases asymmetric—bands (e.g., Egypt’s 2022 EFF, with an upper band more than 20 percentage points above the program target). Inner tolerance bands also showed significant variation, ranging from ± 1 percent (Albania’s 2014 EFF, Romania’s 2013 SBA, Moldova 2016 ECF-EFF) to ± 3 percent (Rwanda’s 2022 PCI and Egypt 2020 SBA).



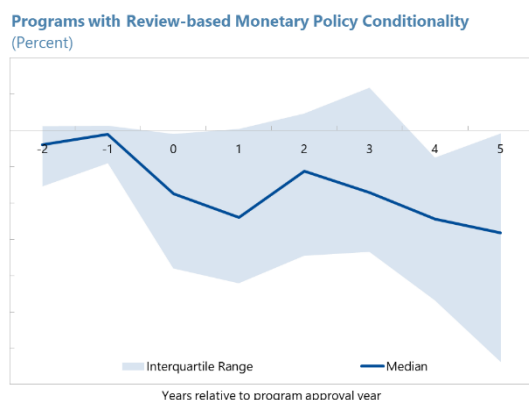
50. Synthetic control estimates suggest that most RBC programs achieved lower medium-term inflation than “no-program” counterfactuals (Figure 26). Five years after program start, median inflation under RBC programs was about two percentage points lower than a “no-program” counterfactual, with all cases in the interquartile range experiencing lower outcomes than their comparators. This suggests that RBC programs have been relatively consistent in achieving levels of

⁴³ Strong central bank transparency contributes to accountability and enhances policy effectiveness. The IMF has developed the Central Bank Transparency Code (CBT), a voluntary international standard designed to help central banks strengthen their transparency practices across five key pillars: governance, policies, operations, outcomes, and official relations with government and other agencies. Further information is available on the [dedicated CBT website](#).

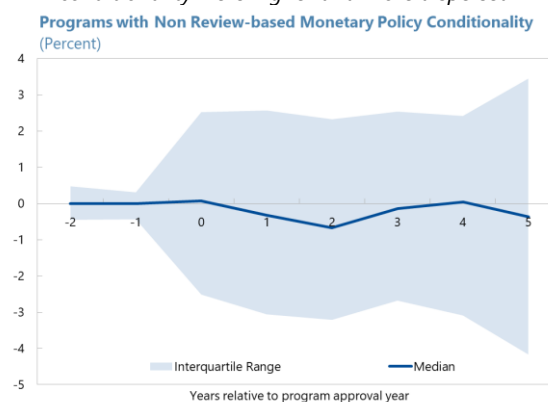
medium-term inflation lower than suggested by “no-program” benchmarks—especially in the case of ICCs, relative to MPCCs. As MPCCs have been applied to a wider range of policy frameworks—including many that could be characterized as “intermediate” or “transitional” in the sense of ¶43—the difference observed between the ICC and MPCC programs may reflect the greater challenges the latter face in designing and implementing conditionality within frameworks that lack clear or well-established nominal anchors. At the same time, programs relying exclusively on quantitative conditionality proved less consistent than either ICC and MPCC programs in avoiding higher medium-term inflation outcomes relative to “no-program” counterfactuals.⁴⁴ While QPC compliance was higher than for RBCs during ROC 2026, this did not systematically translate into lower realized inflation, possibly reflecting weak links between inflation and the central bank balance sheet or an overly lax calibration of some QPCs.

Figure 26. Inflation Outcomes Relative to “No-Program” Counterfactuals, by Form of Monetary Policy Conditionality

Most RBC programs experienced lower medium-run inflation, relative to “no-program” counterfactuals...



...while outcomes for programs with non-RBC-based conditionality were higher and more dispersed

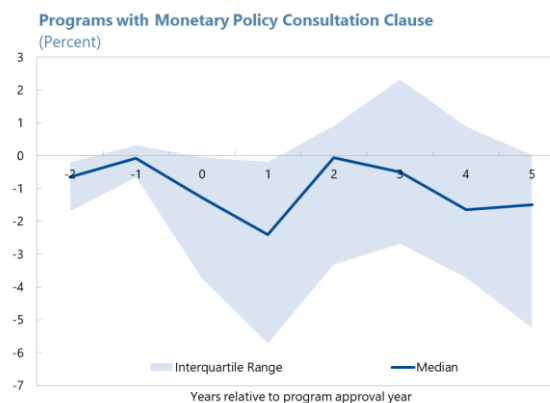
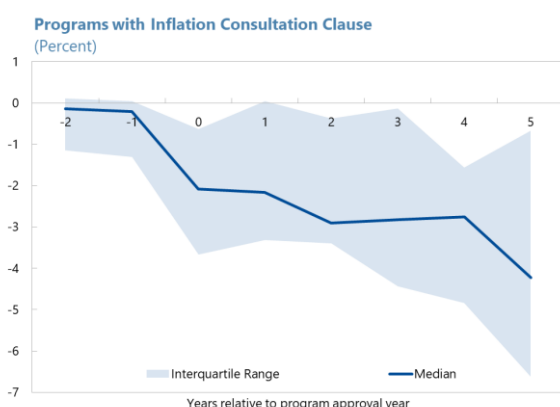


⁴⁴ Though not reported in Figure 26, a “deeper dive” into the results for programs relying solely on quantitative monetary policy conditionality reveals important differences between cases where the conditionality was based on ceilings for BRMs, versus ceilings for NDAs. For example, while the median NDA-based program achieved medium-run inflation outcomes broadly comparable to a “no-program” counterfactual, the median BRM-based program achieved notably lower outcomes—possibly reflecting the fact that BRMs target “liability-side” aggregates that, in some contexts, may be more closely linked with inflation. At the same time, BRM programs displayed greater dispersion in their performance relative to “no-program” counterfactuals, suggesting potential challenges in the calibration and implementation of BRM-based conditionality.

Figure 26. Inflation Outcomes Relative to “No-Program” Counterfactuals, by Form of Monetary Policy Conditionality (concluded)

Among RBC programs, differences relative to counterfactuals were especially pronounced for ICCs...

...and less pronounced for MPCCs.

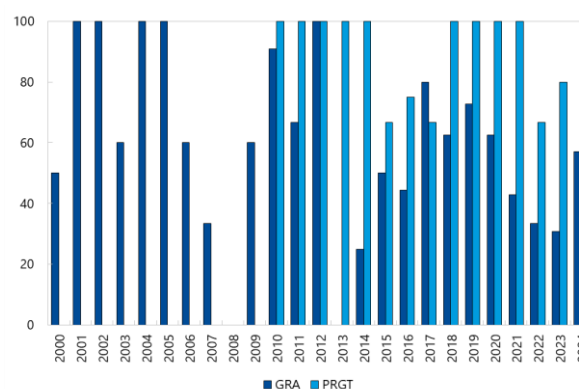


Sources: IMF AREAER; MONA database; IMF WEO; and IMF staff calculations.

51. RBC targets were often calibrated and adjusted to balance realism and credibility in the face of shocks. Sri Lanka’s 2023 EFF, for instance, initially set a high MPCC midpoint target of 52 percent to accommodate the first-round effects of energy price and exchange rate pressures, progressively lowering to 15 percent in subsequent reviews, as inflation fell rapidly amid fiscal tightening and currency appreciation. By contrast, Mozambique’s 2022 ECF raised its midpoint from 7 to 15 percent in end-2022 in response to commodity price shocks and flooding, later reducing it to 6 percent as inflation declined. While such adjustments provided flexibility in response to short-term shocks, they also underscored the importance of maintaining credibility by keeping conditionality aligned with achieving price stability over the medium run, supported by clear communication.

52. Since the introduction of RBC, 72 out of 291 reviews recorded breaches of inflation bands, with the frequency rising during the ROC 2026 period (Figure 27).

Figure 27. RBC Compliance Rates (Percent of reviews)



Sources: IMF staff reports; IMF staff calculations.

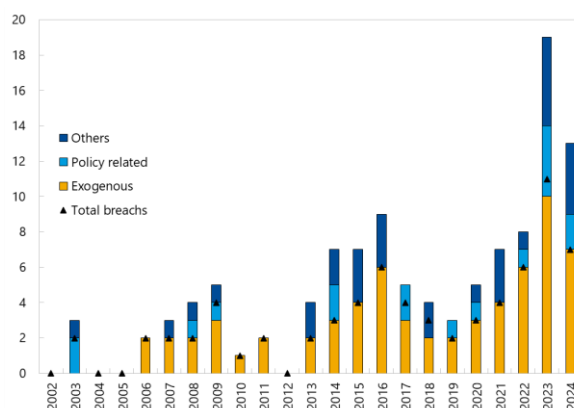
Breaches were evenly split between outer and inner bands. Half of the total reflected inflation overshoots, typically associated with outer-band breaches that triggered Board-level consultations, especially during the 2022 global inflation shock (for example Armenia’s 2019 SBA, Costa Rica’s 2021 EFF, Serbia’s 2021 PCI). Lower-band breaches were more prevalent during periods of low global inflation, such as the post-GFC (e.g., Albania’s 2014 EFF, Guatemala’s 2009 SBA, Hungary’s 2008 SBA, Jamaica’s 2016 SBA, Romania’s 2013 SBA, Serbia 2015 SBA) and the first part of the COVID-19 crisis

(e.g., Egypt’s 2020 SBA, Rwanda’s 2019 PCI), with few such breaches occurring during the 2022–23 inflation shock (e.g., Armenia 2022 SBA).

53. Most breaches were associated with exogenous shocks, with several cases linked to the fiscal policy stance (Figure 28).⁴⁵ RBC requires

the authorities to discuss the causes of breaches with staff—for inner bands—or the Board—for breaches outer bands—and to present remedial measures to keep the program on track.⁴⁶ Staff reports for program reviews commonly linked inflation deviations to global energy and food prices (55 reviews) domestic food prices (12 reviews), and administered price adjustments (11 reviews). Exchange rate depreciation was sometimes cited as a contributing factor (e.g., in Brazil’s 2002 SBA, or Türkiye’s 2005 SBA). Fiscal policy also played a role with both tightening (e.g., Armenia’s 2019 SBA, Sri Lanka’s 2016 EFF), and loosening (e.g., Argentina’s 2018 SBA, and Romania’s 2011 SBA) contributing to breaches.

Figure 28. Key Drivers of Breaches of Review Based Conditionality (Number of explanations of breaches)



Sources: IMF staff reports; IMF staff calculations.

54. Corrective responses typically involved monetary tightening and measures to reinforce central bank credibility. Such measures included raising policy interest rates, adjusting reserve requirements, and reaffirming policy commitments. Several programs also emphasized the importance of maintaining tight fiscal policies and advancing structural reforms designed to support disinflation efforts. Some countries took measures to strengthen their inflation targeting frameworks, improved forecasting capacity (e.g., Rwanda’s 2022 PCI), or introduced forward guidance to enhance transparency and policy effectiveness (e.g., Albania’s 2014 EFF). Authorities often, combined interest rate adjustments with broader structural reforms or public communications strategies intended to preserve policy credibility and support macroeconomic stability (e.g., Armenia’s 2019 SBA, Romania’s 2011 SBA, and Rwanda’s 2022 PCI).

55. In summary, RBC represents an important evolution in program design and can offer countries with appropriate policy frameworks greater flexibility compared to quantity-based approaches. RBC has allowed conditionality to be more closely tailored to country-specific objectives and circumstances, accommodating a range of policy frameworks and levels of institutional capacity. Its design has been flexible, with targets and consultation bands adjusted to evolving inflation conditions. This adaptability has supported policy responses to major global and domestic shocks. However, frequent breaches during global shocks highlight the need for carefully

⁴⁵ This analysis is based on a textual review of staff reports for programs with RBC. The causes of breaches are drawn from staff’s analysis and explanatory letters provided by the authorities and included in the staff reports.

⁴⁶ Out of 35 outer band breaches (up to end-2024 programs), all but one of the reviews with outer breaches included the associated Letter of Intent, while only 65.7 percent of them included a consultation attachment.

calibrated design, clear signaling, and strong institutions designed to maintain credibility and support timely, decisive responses to significant inflationary pressures.

SECTION VII. EXCHANGE RATE AND EXTERNAL ADJUSTMENT⁴⁷

This analysis examines the role of exchange rate adjustment in addressing external sector vulnerabilities in programs and the effectiveness of the External Balance Assessment (EBA) models in guiding exchange rate policies in such contexts. Significant real exchange rate adjustments were rare, while successful exchange system reforms depended on consistent macroeconomic and structural policies.

A. External Vulnerabilities and Exchange Rate Adjustment

56. Many programs in the 2026 ROC started with significant external vulnerabilities. The IMF's external sector assessments (ESA) indicated that nearly half of GRA and 60 percent of PRGT programs began with external positions moderately to substantially weaker than warranted by fundamentals and desirable policies, while only nine programs were assessed as moderately to substantially stronger than warranted.^{48 49} Even among those broadly in line, vulnerabilities such as low international reserves coverage or elevated external debt burdens were common (Figure 29).⁵⁰

57. Real exchange rate adjustment was rarely the main tool for rebalancing. About 24 percent of programs with their own currencies experienced nominal effective exchange rate (NEER) depreciation of at least 15 percent—mostly in GRA programs and floating exchange rate regimes (table in Figure 29).⁵¹ Yet, only half achieved comparable real effective exchange rate (REER) depreciation, with none observed in countries lacking their own currency. Similar patterns are observed for GRA and PRGT programs as well as among countries with floating and non-floating exchange rate regimes, underscoring the inflation's role in eroding nominal gains.

58. Case studies underscore the importance of consistent macroeconomic and structural policies to achieve durable REER adjustment. Programs that combined nominal exchange rate adjustment with tight fiscal and monetary policies and coordinated structural reforms were better positioned to eliminate real exchange rate overvaluation. For example, Rwanda's 2016 SCF achieved REER adjustment through credible macroeconomic policies and supporting structural reforms.

⁴⁷ Prepared by David Florian, Jaime Guajardo, Mahmoud Harb, Tobias Krahne, Hyunmin Park, and Faezeh Raei, with contributions from Yulia Ustyugova and Alexander Zaborovskiy.

⁴⁸ The discussion in this paragraph refers to the results of ESAs reported at program request or, if not available, the latest available ESA before the program request is used.

⁴⁹ Programs with moderately to substantially stronger ESAs included one PLL and three PCIs which did not involve an immediate BOP need. Arrangements without ex-post conditionality (FCL, SLL) are not included.

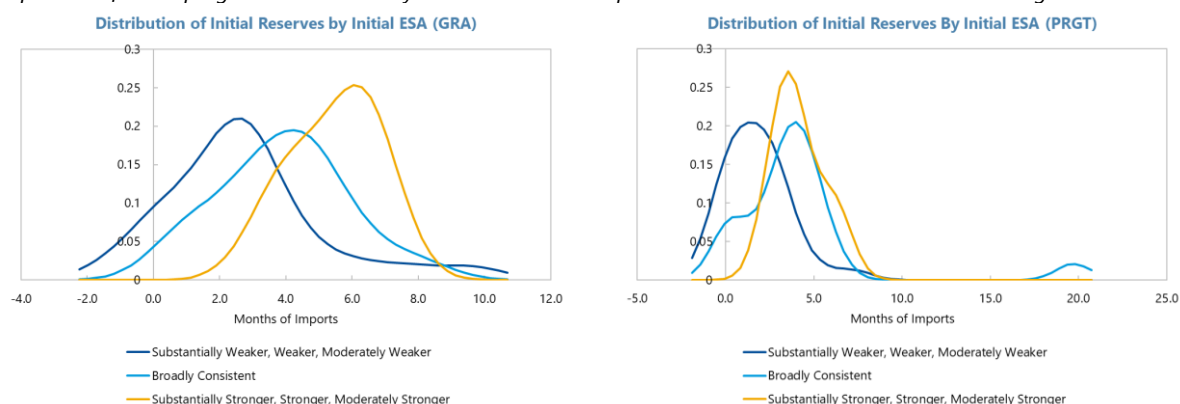
⁵⁰ Among countries with initial external position broadly in line, 57 percent of GRA and 11 percent of PRGT programs had external debt higher than 65 percent of GDP.

⁵¹ Countries without their own currencies include those with no separate legal tender based on the AREAER classification and WAEMU and CEMAC members. These countries cannot individually adjust their nominal exchange rates. But they can achieve real depreciation by maintaining inflation rates below those of their trading partners.

Ethiopia's 2024 ECF more than doubled the official exchange rate and significantly reduced the parallel market premium without triggering major inflation, supported by tight monetary and fiscal policies. In Mauritania (2017, 2023 ECFs), sustained and well-sequenced reforms to bolster central bank independence and deepen the FX market helped reduce central bank FX interventions and set the ground for higher exchange rate flexibility. By contrast, Guinea (2017 ECF) faced persistent REER overvaluation amid slippages in policy adjustment and slow progress in structural reforms, while Malawi's 2018 and 2023 ECFs illustrate how inadequate buffers, fiscal dominance, and delayed reforms undermined fixed exchange rate regimes and led to widening external imbalances.

Figure 29. Exchange Rate and External Adjustment

Programs with stronger initial ESA tended to have higher initial reserve coverage. However, 33 percent of GRA programs and 37 percent of PRGT programs with "broadly consistent" external positions still had low initial reserve coverage.



Sustained real effective exchange rate (REER) depreciation remained elusive.

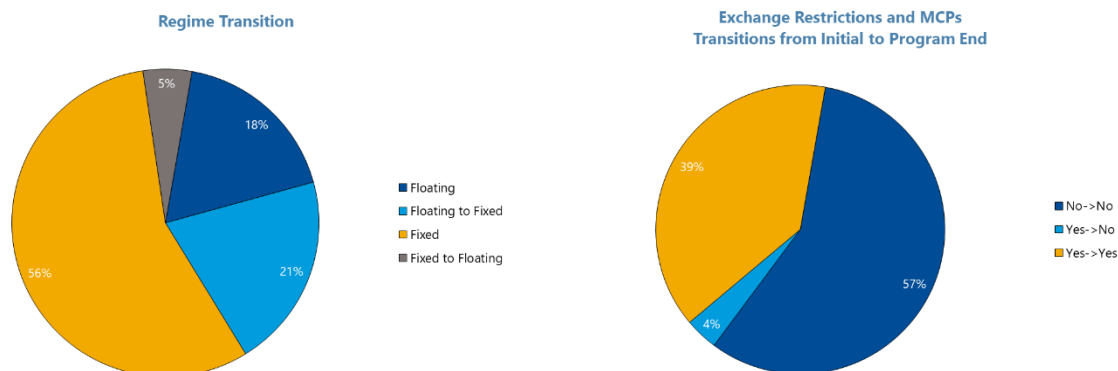
Share of IMF-supported Programs Achieving Exchange Rate Adjustment
(Percent)

	NEER Adjustment of at least 15 percent	Sustained REER Adjustment of at least 15 percent, conditional on NEER adjustment of similar magnitude
All programs in countries with their own currency	24	12
o/w Initial currency overvaluation	22	6
GRA programs	28	12
PRGT programs	19	12
Floating exchange rate regimes	50	20
Non-floating exchange rate regimes	17	10

Figure 29. Exchange Rate and External Adjustment (concluded)

A small fraction (5 percent) of programs transitioned from a non-floating regime to a floating regime, while a larger share (21 percent) moved in the opposite direction...

...and only a small fraction of programs fully removed exchange restrictions and MCPs.



Sources: International Financial Statistics; MONA database; IMF WEO; IMF External Sector Assessment database; Culiuc and Park (forthcoming); IMF staff calculations.

Notes: The initial reserves are net of IMF credit. Sustained real depreciation is assessed by comparing the average REER between months 19 and 24 after program approval with the average REER observed seven to twelve months prior to approval. The chart showing exchange regime transition excludes countries that are in a currency union or that don't have a separate legal tender according to 2023 AREAER.

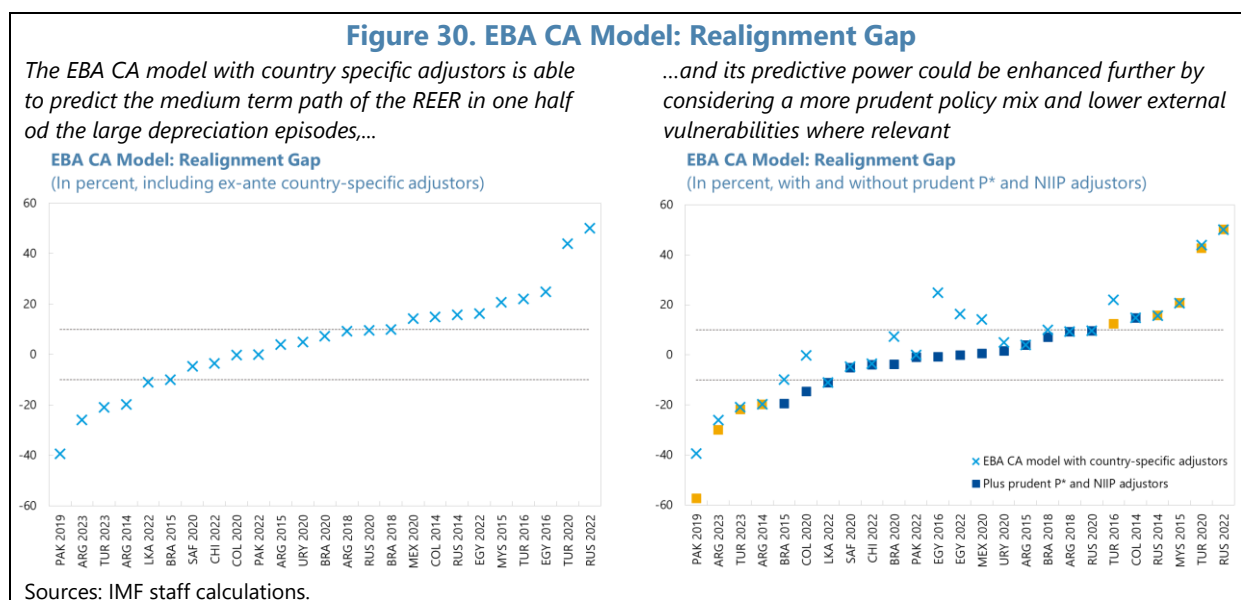
B. Large Exchange Rate Depreciation and the External Balance Assessment

59. The EBA Current Account Model (CA) shows moderate accuracy in predicting exchange rate misalignment. The EBA CA model is not well suited to predict the near-term REER movements, but it should shed light on the medium-term REER path given the model's focus on medium term fundamentals and desirable policies, especially when complemented with staff judgement and the external sustainability model. An evaluation of 25 large depreciation episodes in EMs between 2014 and 2024 finds that the CA model—when complemented with ex-ante country-specific adjusters—reasonably predicts the medium-term REER path in half of the episodes.⁵² The model's performance is assessed using the realignment gap, defined as the difference between the medium-term REER depreciation and the initial REER gap estimate from the CA model. This gap falls within the ± 10 percent range considered as an acceptable margin in 12 of the 25 cases (Figure 30), though the model substantially over- and under-predicts in the remainder.

60. Model accuracy improves when inputs are adjusted to reflect a more prudent policy mix and when complemented by the external sustainability (ES) model. Modifying the EBA CA model's assumptions on medium-term desirable policies (P^*) toward a tighter fiscal stance and higher reserve accumulation, and adding the ES model's REER depreciation needed to increase the net international investment position (NIIP) in countries where it is highly negative boost the

⁵² These depreciation episodes were identified based on monthly end-of-period bilateral exchange rates against the U.S. dollar following Culiuc and Park (forthcoming).

model's performance.⁵³ Under this combined approach, the model accurately predicts the medium-term REER path in two thirds of the sample (blue squares in Figure 30).



61. Nonetheless, the model faces notable limitations to predict REER movements in cases marked by high inflation and unresolved external imbalances (yellow squares in Figure 30). The model tends to overpredict the REER depreciation in high inflation environments where unanchored inflation expectations and fiscal dominance offset nominal depreciation, and underpredicts it in cases facing successive shocks (e.g., owing to political uncertainty or unresolved external imbalances). These results reflect the model's limitation in capturing near-term instability, as it is primarily designed to assess medium-term equilibrium outcomes.

SECTION VIII. ASSESSING STRUCTURAL CONDITIONALITY⁵⁴

Structural conditionality (SC) is central to IMF-supported programs, anchoring reforms aimed at supporting stabilization, resilience building, and growth. Despite repeated calls for greater parsimony, the volume and scope of SC has expanded, and reform depth has not improved. At the same time, SC has adapted to evolving needs and implementation rates remained high. While strong SC is shown to

⁵³ The fiscal balance P* is calibrated to lower public debt to 60 percent of GDP or by 12 percent of GDP in five years. No adjustment was made if public debt was lower than 60 percent of GDP or if the original P* was larger than the one estimated by staff. The reserve P* was adjusted to bring reserves to 100 percent of the ARA metric in five years. No adjustment was made if reserves were higher than 100 percent of the ARA metric or if the original P* was larger than the one estimated by staff. The ES model's REER depreciation is calibrated to raise the NIIP to -40 percent of GDP or by 10 percent of GDP in five years, which is added to the REER gap from the EBA CA model.

⁵⁴ Prepared by Anja Baum, Jocelyn Boussard, Liam O'Sullivan, with research assistance from Helen Zheng, Max Kushnir, and Holt Williamson.

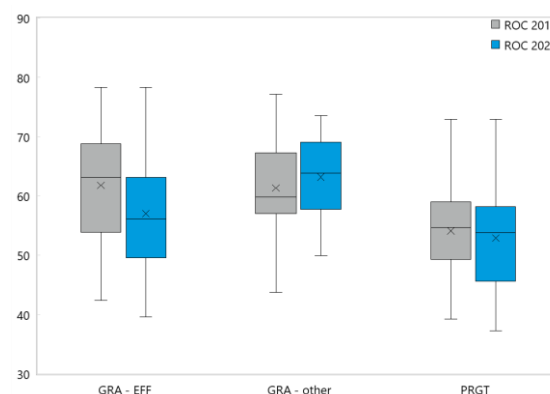
improve institutional quality, growth benefits often materialize after programs expire. An early focus on high-impact reforms, supported by capacity development, is therefore critical.

A. Overview

62. Structural conditionality (SC) is central to Fund-supported programs, anchoring reforms critical for macroeconomic stabilization, resilience building, and growth. As all conditionality, SC should include measures that assist member in resolving their BOP problems while meeting the criticality test under the Conditionality Guidelines. As programs confront deep-rooted structural challenges amid a complex global environment, the need for effective SC has grown. Reflecting this shift, EFFs are now more commonly deployed in GRA borrowing countries with weaker institutional capacity, while structural issues have also intensified in PRGT cases (Figure 31). At the same time, growing demands on SC highlight the importance of

parsimony ([IMF, 2019a](#)) and adaptability, enabling reforms to adapt to shocks while preserving strategic clarity and ambition.

Figure 31. Institutional Quality at Program Approval (Percentage points)



Sources: WDI; Fraser Institute; MONA database; IMF staff calculations.

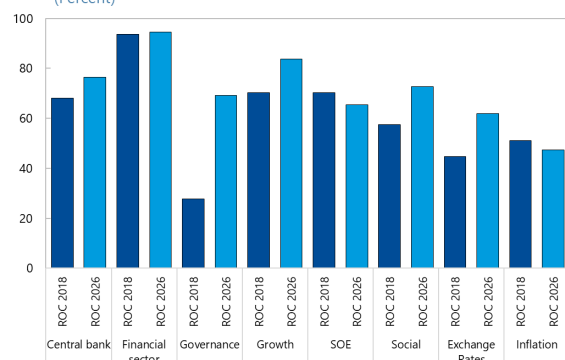
63. Programs have broadened their objectives over time, leading to an increased reliance on SC (Figure 32). While fiscal adjustment remains a core objective in nearly all programs, governance has gained prominence in both GRA and PRGT arrangements, with GRA programs also placing greater emphasis on growth, social, and exchange rate issues, despite the share of programs with SC on the external sector being limited. Adjustment-oriented fiscal conditionality still accounts for nearly 40 percent of all SC, but its share has declined—particularly in PRGT programs—as priorities shifted toward governance.⁵⁵ Despite the growing emphasis on growth-oriented reforms in the PRGT, only half of all programs include SOE reforms as an objective, even though these reforms play a vital role in strengthening public sector efficiency. Financial sector SC has also decreased, while social sector SC has gained traction, particularly in GRA programs. During COVID-19, GRA programs prioritized resilience-oriented SC, but this emphasis has since declined.

⁵⁵ This increase follows recommendations and subsequent increased focus on the macro-criticality of governance weaknesses discussed in the [2018 Review of 1997 Guidance Note on Governance](#).

Figure 32. Structural Focus of IMF-Supported Programs and Related Conditionality

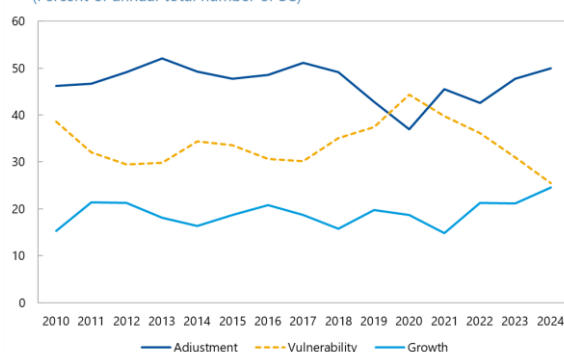
GRA programs saw a broadening set of strategic objectives across various categories...

GRA - Share of Programs with Respective Strategic Objectives
(Percent)



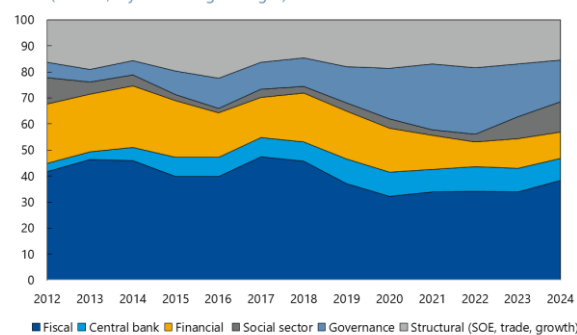
GRA programs saw a spike of SC focused on resilience during the pandemic period...

SC Orientation - GRA
(Percent of annual total number of SC)



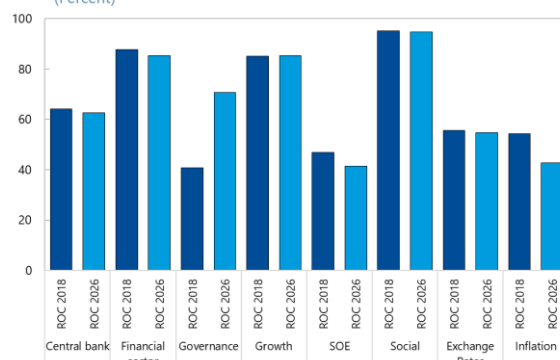
Priorities shifted towards governance and more recently social SC in the GRA, with financial SC declining steadily.

GRA, ROC 2026 - Sectoral composition of SBs across programs
(Percent, 2 year moving averages)



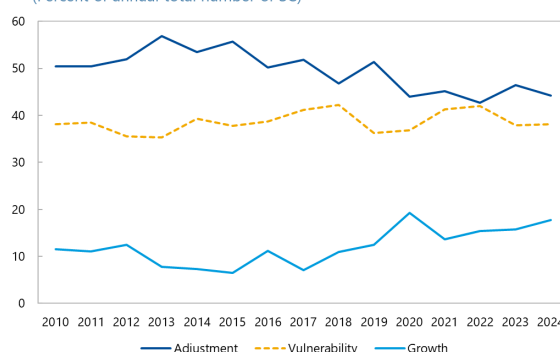
... while PRGT objectives remained largely stable, except for an increased focus on governance issues.

PRGT - Share of Programs with Respective Strategic Objectives
(Percent)



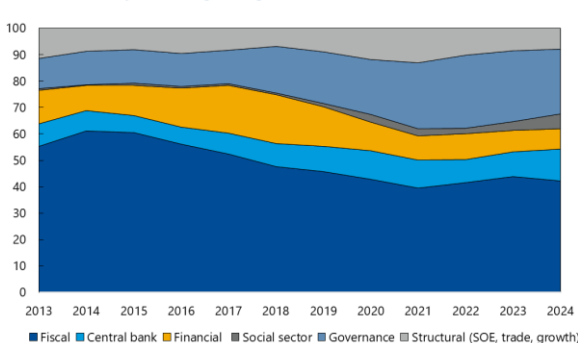
... while PRGT programs have become more growth-oriented, albeit from low levels.

SC Orientation - PRGT
(Percent of annual total number of SC)



Fiscal and financial SC declined in the PRGT, where an increased focus on governance aims at tackling core structural issues.

PRGT, ROC 2026 - Sectoral composition of SBs across programs
(Percent, 2 year moving averages)



Sources: MONA database; IMF staff calculations.

Notes: Frequency of fiscal objectives in top charts is not shown, as they are close to 100 percent with no change over time. SC orientation is categorized as follows: "Adjustment"—measures supporting stabilization (e.g., fiscal, monetary, FX, SOE subsidies, financial restructuring); "Vulnerability"—governance and institutional reforms (e.g., debt management, transparency, statistics, anti-corruption); "Growth"—private-sector and development reforms (e.g., labor/product markets, pensions, SOEs, trade, poverty reduction). The strength of SC is measured using the structural conditionality implementation and depth (SCID) score, based on the methodology in IEO (2021). Each SC is scored by implementation—1 (met or met with delay), 0.5 (partially met) or 0 (waived or not met)—and depth—1 (high), 2/3 (medium), and 1/3 (low). The SCID score is the sum of implementation scores weighted by depth.

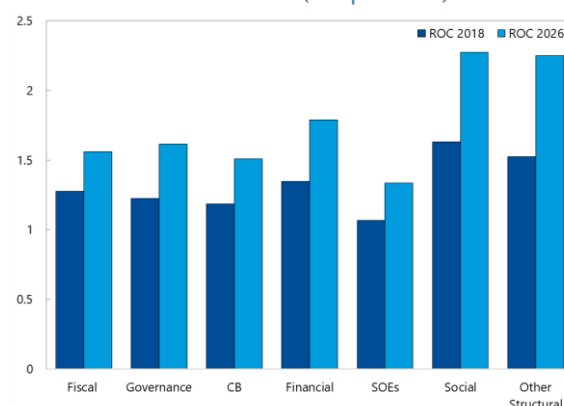
64. Overall, the number of SCs in Fund-supported programs remains at levels comparable to the ROC 2018, despite earlier calls for greater parsimony, with variation in sequencing and follow up (Figure 35). SCs continue to prioritize revenue and governance, with limited focus on social sectors. Sequencing varies: fiscal and governance SCs are often spread across reviews, while growth-related SCs—such as SOEs and trade—are typically one-off with little follow-up. Social objectives are mostly delivered through ITs, and social SCs, when used, are rarely sustained beyond two reforms, especially in PRGT programs. Limited continuity in growth and social reforms may weaken SC's long-term impact.

65. SC has also demonstrated adaptability to evolving needs. During the pandemic, the number of sectors covered declined from an average of 7.5 to 6 per program year but has since rebounded. Adaptability is also evident in longer implementation timelines—especially in growth-focused areas—reflecting the complexity of reforms and the need for broader engagement (Figure 33). However, this may have delayed progress on social and growth-oriented agendas.

66. Implementation of SC remains moderately strong, but reform depth has declined, with limited evidence of deeper sequencing. Over 80 percent of SC are either met or

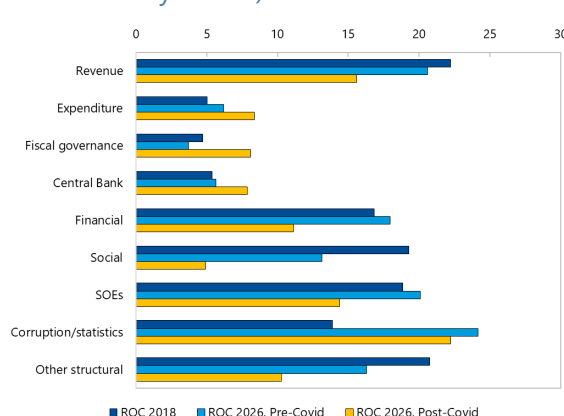
implemented with delay, but the share of unmet conditions has risen since the last ROC—particularly in GRA programs following the pandemic (Figure 35). The 2018 ROC emphasized the importance of high-depth conditionality for lasting impact, yet recent trends show a decline in depth, especially in GRA programs. Depth has fallen across many categories, while increases were seen in expenditure and fiscal governance, central bank, and corruption-related SCs (Figure 34). The greatest concentration of high-depth conditionality is now in anti-corruption measures, reflecting the increased emphasis on governance.⁵⁶ Importantly, depth does not appear to influence overall implementation, as roughly 80 percent of both high- and low-depth SC is met or implemented with delay. However, the highest shares of unmet SCs are found in SOE reforms, corruption and other private sector-related

Figure 33. SC – Time to Implementation from Set to Test Date (In quarters)



Sources: MONA database; IMF staff calculations. "Governance" includes fiscal governance, as well as anti-corruption measures.

Figure 34. Average High Structural Depth by Sector (Percent of total structural conditions by sector)



Sources: MONA database; IMF staff calculations.

⁵⁶ As noted in the 2023 review of the governance framework, durable progress against corruption requires a sequenced and tailored approach that is mindful of both implementation capacity and political economy realities.

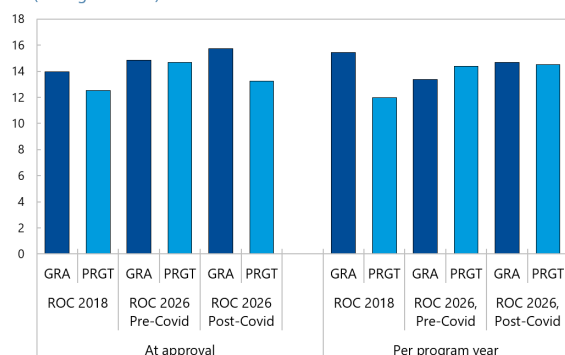
areas—sectors that also tend to carry a greater share of high-depth conditions. Only 27 percent of sequenced reforms show increased depth over time, underscoring the lack of progress toward deeper structural reforms in later program stages.

67. Modifications of structural benchmarks (SBs) remain frequent, while prior actions (PAs) are increasingly used to secure reform implementation. About 20 percent of all SBs are modified, mostly through deadline extensions. At the same time, the use of PAs has increased—especially post-pandemic—across both GRA and PRGT programs, and notably in financial, social, other structural and corruption-oriented areas. This may reflect challenges in implementing high-depth SC in these sectors, often with limited capacity development. Ownership may also be a factor, particularly for reforms requiring broad stakeholder buy-in, such as anti-corruption legislation.

Figure 35. Design and Implementation of Structural Conditionality

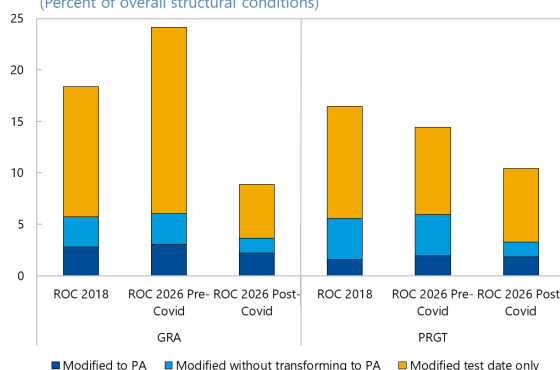
The volume of SC remains high...

Quantity of Structural Conditionality
(Average number)



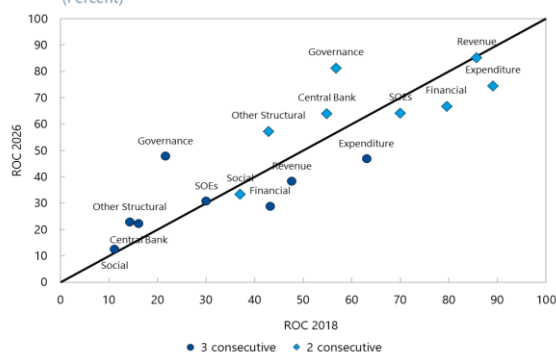
... and deadline extensions remain frequent, with PAs increasingly used to secure reform implementation.

Share of Modified Conditions: GRA Versus PRGT
(Percent of overall structural conditions)



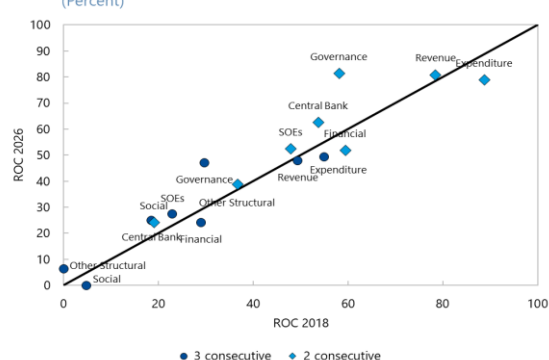
SC is sequenced (multiple reforms in the same sector across reviews) little beyond fiscal, where it has declined in the GRA...

Sectoral Sequencing in GRA Programs by ROC Sample
(Percent)



... while social, corruption and other growth-oriented SC are rarely continued beyond two reviews in either facility.

Sectoral Sequencing in PRGT Programs by ROC Sample
(Percent)



Implementation rates remained high in both accounts...

... while depth of SC declined slightly, especially in the GRA.

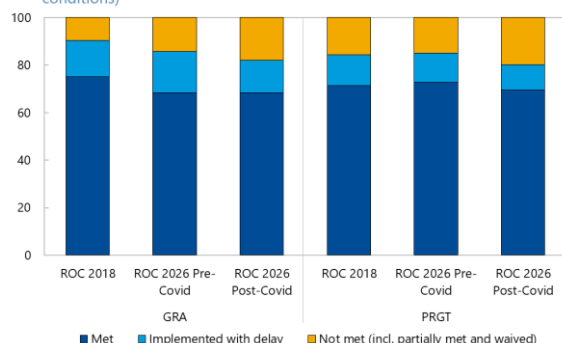
Figure 35. Design and Implementation of Structural Conditionality (concluded)

Implementation rates remained high in both accounts...

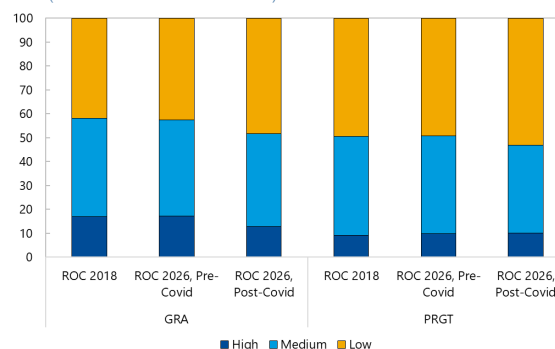
... while depth of SC declined slightly, especially in the GRA.

Implementation: GRA vs PRGT

(Percent of total met, implemented with delay, and not met structural conditions)

**Average Structural Depth by Account**

(Percent of structural conditions)



Sources: MONA database; IMF staff calculations.

Notes: In the middle two charts, dots represent the ratio of programs with SC of a given sector that was set for either two or three (and more) consecutive reviews, as a share of programs that had any SC in that given sector. "Consecutive" here allows for one review gaps to account for longer reform implementation timelines. The 45-degree black line marks equal sequencing prevalence from ROC 2018 to ROC 2026. A red dot lying below this line would indicate that sequencing in the respective sector declined over time (e.g., financial and expenditure SC), for dots lying above it sequencing increased, particularly visible for governance. "Governance" includes fiscal governance as well as anti-corruption measures.

B. Impact of Structural Conditionality on Growth

68. Two empirical strategies are employed to assess the impact of SC on growth. A key challenge in such analysis is establishing a credible counterfactual, as Fund programs are typically implemented in countries experiencing weak growth or crisis conditions. The *first* approach employs OLS to compare outcomes across programs, allowing for an assessment of how design features—such as the inclusion of SC—affect growth relative to the average Fund programs. The *second*, more robust approach constructs a synthetic counterfactual representing a scenario without Fund engagement, controlling for other determinants of growth before, during, and after the program period.⁵⁷

69. The strength and orientation of SC are assessed using the Structural Conditionality Implementation and Depth (SCID) index, following the methodology developed in the [IEO \(2021\) report](#). SCID captures both the implementation status and the depth of the SCs.⁵⁸ It is further classified into three broad categories based on its intended orientation:

- *Enhancing (inclusive) growth and economic efficiency.* This includes private-sector-related reforms in labor and product markets; pension, and civil employment reforms; SOE reforms

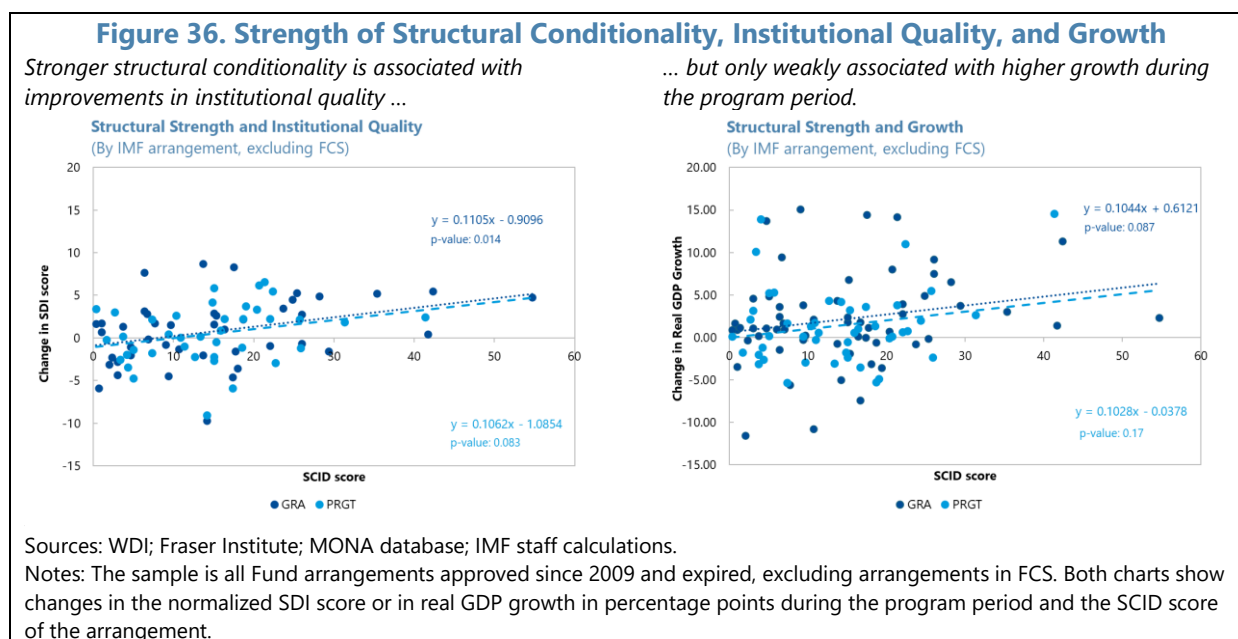
⁵⁷ An extensive literature has relied on various empirical methods to correct for the bias of selection into Fund-supported programs and finds mixed results, although studies based on more recent data suggest a more positive short-run impact on growth ([IEO, 2021](#); [Kim and Lee, 2021](#)), with limited evidence on medium to long-term effects.

⁵⁸ Specifically, each SC is scored by implementation—1 (met or met with delay), 0.5 (partially met) or 0 (waived or not met)—and depth—1 (high), 2/3 (medium), and 1/3 (low). The SCID score is the sum of implementation scores weighted by depth. Depth reflects the expected structural impact and durability of reforms, including their effectiveness in facilitating follow-up reforms ([IMF, 2024b](#)). Depth ratings are assigned using text analysis algorithms, which provide a systematic and reliable classification.

other than pricing; measures related to trade policy or natural resources; and measures related to the implementation of poverty reduction and development strategies.

- *Managing vulnerabilities and building resilience.* This category encompasses reforms aimed at strengthening fiscal, monetary, and financial governance or controls (such as debt management, fiscal transparency, intergovernmental fiscal relations, central bank auditing transparency and financial controls, economic statistics, and anti-corruption).
- *Facilitating adjustment.* Measures in this category support stabilization efforts such as fiscal revenue and expenditure measures, reforms that enhance monetary transmission mechanisms and foreign exchange market functioning, SOE subsidy reforms, and financial sector restructuring.⁵⁹

70. Stronger SC is associated with improvements in institutional quality, but growth effects are more elusive during the program period. Programs with higher SCID scores achieve larger improvements in institutional quality as measured by the Structural Development Indicator (SDI), especially in GRA arrangements (Figure 36).⁶⁰ The evidence linking improvements in institutional quality and growth during the program period is weaker: the relationship is either not statistically significant (PRGT arrangements) or appears driven by a few outliers (GRA programs).⁶¹



⁵⁹ The SCID-growth (SCIDG), SCID-vulnerability (SCIDV), and SCID-adjustment (SCIDA) scores attribute a weight of 1 to reforms that are intended primarily for growth, vulnerabilities, and adjustment, respectively, and 0 to all other reforms. Several reforms can have a dual orientation, i.e., fall into two categories, such as restructuring and privatization of financial institutions (adjustment and vulnerability), and anti-corruption legislations/policy (vulnerability and growth).

⁶⁰ Following [IMF \(2023b\)](#) and using Worldwide Governance Indicators (WGI) and indicators from the Fraser Institute, the SDI indicator comprises 5 main areas (governance, external sector, credit market regulation, labor market regulation, and business regulation) which all result from averages of multiple sub-indicators.

⁶¹ This absence of growth benefits during the program period is consistent with prior analyses ([IEO, 2021](#); [Kim and Lee, 2021](#)), that find a positive effect of IMF-supported programs on growth in the very short-run (first year).

71. The counterfactual analysis estimates the growth impact of Fund-supported programs by comparing trajectories of countries in similar circumstances. Following [Lang \(2021\)](#), [Krahnke \(2023\)](#), and [He and others \(2024\)](#), the identification strategy applies a difference-in-differences framework, leveraging the timing of program entry as an instrumented variable. Entry into an arrangement is instrumented using a *shift-share* variable that captures exposure to external shocks and historical program participation. The growth effect at horizon h is estimated using two-stage least square (2SLS), where the coefficient of interest β_2 captures the association between Fund arrangements and growth outcomes:

$$Y_{i,t+h} - Y_{i,t-1} = \beta_2(IMF_{i,t} = \beta_1 IV_{i,t} + \gamma_1 X_{i,t} + \delta_{1,i} + \eta_{1,t} + \delta_{1,i}t) + \gamma_2 X_{i,t} + \delta_{2,i} + \delta_{2,i}t + \eta_{2,t} + \epsilon_{i,t} \quad (1),$$

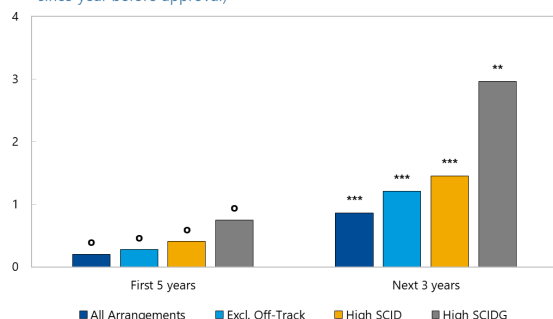
where $Y_{i,t+h} - Y_{i,t-1}$ is the cumulative change in the logarithm of real GDP per capita in country i between year t and year $t + h$, and $IMF_{i,t}$ is a binary variable indicating that country i entered into an arrangement with the Fund between the last quarter of year $t - 1$ and the first three quarters of year t . The vector X of covariates includes determinants of growth and determinants of entry into IMF arrangements identified by the literature, while δ and η are full sets of country and year fixed effects.⁶² Entry into program is instrumented in the first stage by $IV_{i,t}$, the product of the country's historical probability of requesting an arrangement (*share*), and the temporal variation in the IMF's liquid resources (*shift*).

72. The counterfactual analysis suggests that well-designed and well-implemented programs can yield sizable growth benefits, albeit with a lag. Annual real GDP per capita growth is estimated to be between 0.2 and 0.75 percentage points higher during the five years following approval of a Fund-supported program, relative to countries with similar initial conditions that do not enter into an arrangement with the Fund (Figure 37). While stronger and more growth-oriented programs tend to yield larger growth effects, these estimates are insignificant, potentially owing to considerable heterogeneity in outcomes. The impact, and its significance, becomes more pronounced over time: Between six and eight years after the start of the program, estimated growth gains rise to between 0.8 and 3 percentage points. For GRA programs, long-term growth effects range from 0.7 to 2 percentage points and are statistically significant. By contrast, effects for programs under the PRGT range between 0.6 and 3 percentage points but lack statistical significance.

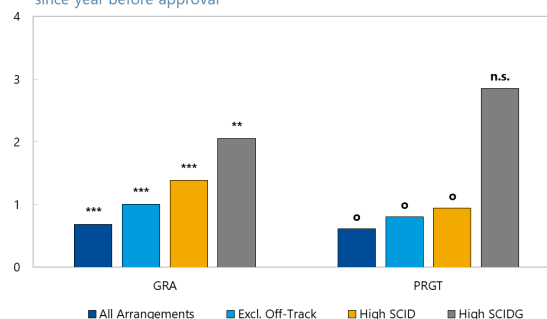
⁶² Included covariates are the primary fiscal balance, log international reserves, external debt, current account balance (all in percent of GDP), the percent change in terms-of-trade, log and squared log of GNI per capita (Atlas Method), a binary variable for the occurrence of a systemic banking, sovereign, or currency crisis ([Laeven and Valencia, 2020](#), [Nguyen and others, 2022](#)), voting similarity to the United States at United Nations General Assembly ([Bailey and others, 2017](#)), and the historical probability of entering into a Fund arrangement, standalone and interact with world GDP growth.

Figure 37. Counterfactual Analysis of the Effect of IMF Arrangements on Growth*Growth dividends take time to materialize*

Effect of IMF Arrangements on Growth Over Time
(Annualized cumulative percent change in GDP per capita since year before approval)

*Payoffs from PRGT arrangements are of similar magnitude but more uncertain than for GRA arrangements*

Long-Run Effect of GRA and PRGT Arrangements on Growth
(Annualized cumulative percent change in GDP per capita since year before approval)



Sources: IMF WEO; MONA database; IMF staff calculations.

Notes: The panel comprises all IMF members over 2000–2024, excluding (limited to) PRGT-eligible countries in GRA estimates (PRGT estimates). The charts report the average annualized coefficient β_2 from separate estimates of Equation (1) for each horizon h , from zero to five years (left panel), and from six to eight years (left and right panel). Statistical significance is indicated at the 99 (***) , 95 (**), 90 (*), 80 (^) or 70 (o) percent level, or below (n.s.). High SCID and SCIDG scores programs are programs with above median scores.

73. Despite important caveats, this analysis offers valuable insights. First, the sample size is constrained by the need to observe medium- to long-term growth outcomes, covering only 10 GRA and 23 PRGT programs from the 2026 ROC sample. Second, while the instrumental variable (IV) approach effectively predicts program engagement, it is less suited to capturing specific program design features—such as the strength of SC or growth orientation—potentially biasing results upward. Third, the IV methodology assumes homogeneous treatment effects, a more restrictive assumption than that required by causal difference-in-differences techniques. Nonetheless, the findings consistently support the conclusion that stronger and more growth-oriented SC is associated with improved long-term growth performance. This underscores the importance of early, high-impact reforms and the tailoring of conditionality to country-specific circumstances in the design of Fund-supported programs.

C. Role of Capacity Development

74. Strategically targeted Capacity Development (CD) is essential to support reform implementation and deepen SC. Since ROC 2018, Fund CD has aligned closely with program priorities—especially in fiscal and financial sectors—while relying on development partners to cover growth-related areas (Box 3). Greater integration is needed in macro-structural and SOE reforms, where growth payoffs are high. Improved coordination and tracking across CD providers will help avoid duplication and address gaps.

Box 3. Trends and Integration of Capacity Development in IMF-Supported Programs

Capacity development (CD) has expanded significantly over the past two decades, becoming increasingly field-based and more concentrated in fiscal areas and low-income and African countries.

Direct single-country and agency CD spending on programs rose from USD 60.5 million in FY23 to an estimated USD 81.5 million in FY25, following a sharp decline in 2020–21 due to pandemic-related travel restrictions.¹ Field-delivery has also grown: by 2023, CD intensity reached an average of 1.1 full-time equivalent (FTE) staff per program, compared to 0.47 FTEs in the post-GFC period (Figure 38). Direct CD spending on programs is now, on average, roughly 2½ times higher than on non-program countries, though regional variations exist.

The integration of CD into programs is assessed through a sectoral concordance analysis linking SC and CD. The methodology matches sectoral classification in MONA, CDMAP, and TIMS to evaluate alignment. Several key patterns emerge:

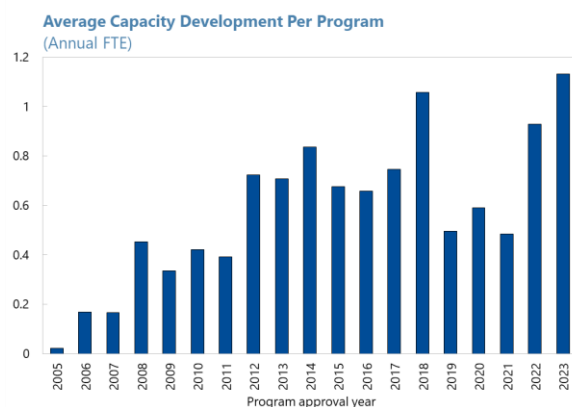
- **Sectoral focus:** CD support remains concentrated in the fiscal sector—especially PFM and debt management—though this emphasis is not fully mirrored in SC. At the same time, governance-related CD is declining, despite continued relevance in SC, and integration remains limited in macro-structural and SOE reforms. This, however, could be partly due to shared CD delivery with other IFIs.
- **GRA vs. PRGT:** There are some differences in emphasis, with PRGT programs showing stronger support on revenue and statistics, and both GRA and PRGT programs increasingly prioritizing PFM and debt management.
- **FCS/Recurrent programs:** The distribution of CD resources for FCS and follow-on programs largely mirror overall patterns, though recent efforts have increased CD support for debt management among prolonged users.
- **High Depth Structural Reforms:** Deeper SC is often backed by stronger CD, especially in fiscal and statistics, but integration remains uneven across other areas.
- **Domestic Revenue Mobilization (DRM):** Within DRM, support for tax administration far exceeds that for tax policy.

Cooperation between the Fund and other IFIs in CD has strengthened in recent years, but opportunities remain to deepen collaboration. The Fund's recent [CD strategy review](#) underscored the value of leveraging existing good practices and innovative communication to effectively address new priorities internally and with development partners. In core areas such as fiscal or monetary/financial policy, multiple coordination mechanisms are already in place—including joint missions with the World Bank and the European Commission, country-level coordination involving headquarters and field offices, and steering committees of Regional Capacity Development Centers and Global Thematic Funds (including the Debt Management Facility). Nonetheless, data to track shared conditionality and CD delivery remain scarce.

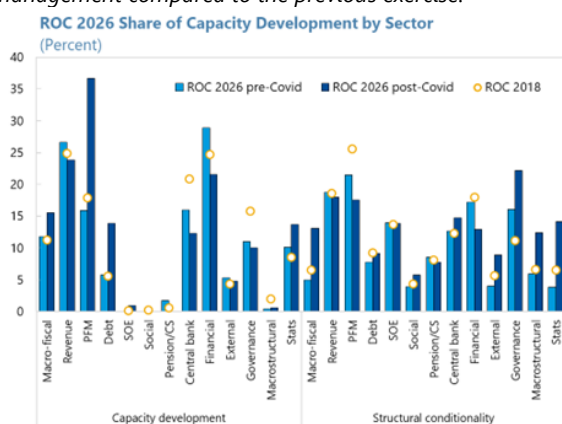
¹ Overall CD spending consists of direct and indirect subcomponents. Direct spending on CD reflects the cost of personnel, travel, and other costs of direct delivery, analytics, administration and CD management. Indirect spending reflects (i) management and administrative spending in delivering departments; and (ii) CD's share in corporate functions, including the Executive Board, Fund buildings and facilities. Direct spending accounted for an estimated USD 285 million in FY23 with indirect spending amounting to a further USD 190 million.

Figure 38. CD Integration into Fund-supported Programs

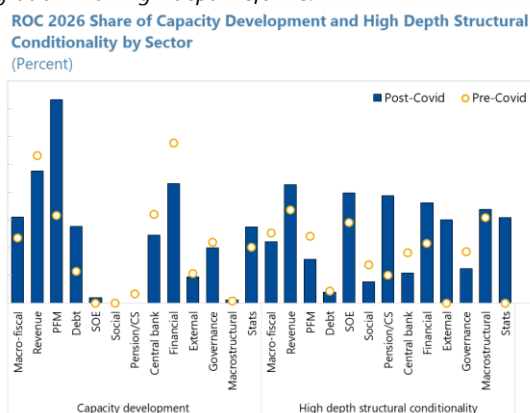
FTEs per program increased under the current ROC following a decline in the pandemic years.



At the overall level, the 2026 ROC sample places additional emphasis on CD support to PFM and debt management compared to the previous exercise.

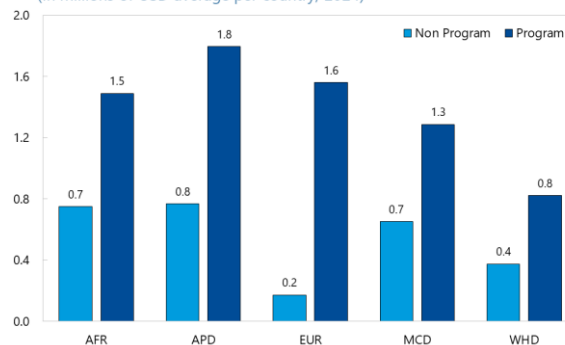


Further CD effort may be required to achieve greater integration with high-depth reforms.

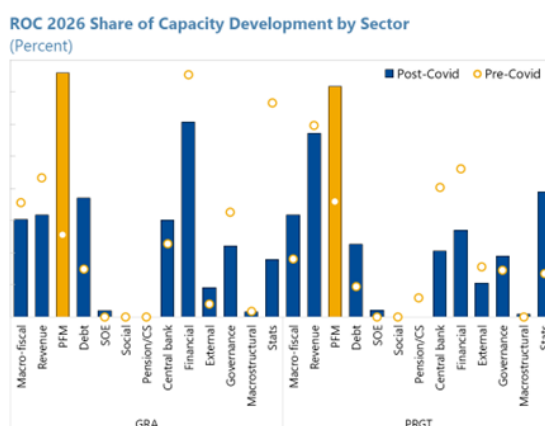


Recent data suggest that program spending on CD is on average more than double that of non-program spending, although there are regional variations.

Average Direct Spending: Program vs. Non-Program Countries
(In millions of USD average per country, 2024)

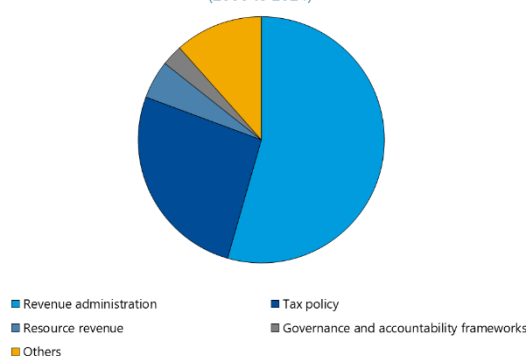


CD coverage of debt management and PFM have accelerated rapidly in the 2026 ROC sample for both GRA and PRGT programs.



Resource management analysis indicates that tax administration CD greatly outweighs support to tax policy.

DRM CD Activity Categories
(2006 to 2024)



Sources: CDMAP; TIMS; MONA database; IMF WEO; IMF staff calculations.

SECTION IX. FRAGILE, CONFLICT-AFFECTED STATES AND SMALL DEVELOPING STATES⁶³

Fund-supported programs in FCS and SDS have been increasingly realistic, adaptable, and tailored to country specificities. Consequently, success, completion rates, and reform implementation are broadly comparable to other programs, and the quality of adjustment has improved relative to the previous ROC. Increased capacity development and conditionality focused on governance, public finance management, and institutional reforms have contributed to the relatively good program performance.

75. FCS and SDS face structural constraints that set them apart from other EMDEs. FCS typically have limited fiscal space, weak institutional capacity, and a high dependence on external grants. Exchange rate and monetary policy frameworks are also typically weaker, contributing to elevated inflation and volatility. SDS, by contrast, feature limited economies of scale, scarce natural resources, and heavy reliance on trade, tourism, and large public sectors. These features heighten vulnerabilities to shocks and call for realistic, tailored, and adaptable program design. As many countries manifest elements of fragility and vulnerabilities to shocks, lessons from programs in FCS and SDS have broader relevance.

76. Program success rates in FCS and SDS are broadly in line with other programs, alongside improved performance and quality of adjustment (Figure 39). Arrangements with FCS and SDS show similar probabilities of success as other groups. In the PRGT space, those that protect or expand social spending and explicitly support social safety nets are marked by more sustained implementation and higher degrees of success. Structural conditionality (SC) implementation remains higher in FCS than in SDS and other programs, despite a slight post-pandemic rise in unmet SCs. Implementation rate of quantitative conditionality also remained high in FCS/SDS programs at around 90 percent. While PRGT-FCS have often adjusted through capital spending cuts, this has eased in the 2026 ROC, despite persistent revenue challenges.

77. Program design in FCS and SDS should be grounded in realistic forecasts, tailored to country-specific circumstances, and supported by targeted capacity development (CD). As emphasized in the IMF's [FCS strategy](#) and the guidance on the [IMF engagement with SDS](#), realistic forecasts, contingency planning and adaptability in response to shocks are key to robust program design in these contexts. Overly complex or not appropriately prioritized conditionality can undermine ownership, while clear, well-sequenced reforms aligned with core state functions foster stronger traction. Parsimony in the areas of focus of conditionality remains important—even with a higher volume of SCs—particularly in FCS, where conditionality can increase clarity and buy-in from key stakeholders, as opposed to MEFP commitments. Close coordination with development partners and targeted CD can enhance implementation prospects. Going forward, IMF engagement should continue leveraging available flexibility in program design and shifting toward fragility-informed design, better integrating political, social, and institutional diagnostics.

⁶³ Prepared by Jocelyn Boussard and Alexei Miksjuk, with research assistance from Lavinia Zhao.

78. Programs in FCS/SDS have increasingly adapted to elevated shock risks, but persistent optimism remains a challenge. Macroeconomic frameworks have become more realistic, with more conservative assumptions and great use of explicit buffers and contingency plans, but growth optimism remains a challenge especially in FCS (Figure 40). Structural reform implementation remains adaptable, as shown by the high share of SBs with revised implementation dates. Authorities also demonstrated stronger ownership of reforms, with fewer SCs converted to prior actions. Like other programs, FCS and SDS programs adjusted quantitative targets more readily in response to the COVID shock than others, with a similar share of non-met QPCs.

79. The 2026 ROC programs show somewhat more tailored conditionality and elevated CD support. Structural conditionality in FCS and SDS span a mix of high- and low-depth measures, similar to other programs. The use of conditionality related to governance, anti-corruption, off-budget spending and monetary financing has increased to address key vulnerabilities. Structural conditionality also increasingly covered central bank operations in post-pandemic programs. With increased sectoral coverage, however, the volume of structural conditions has remained high. At the same time, while FCS received increased CD resources, the distribution by sector for FCS and SDS is not significantly different from the broad overall distribution pattern.

Figure 39. Program Outcome in FCS and SDS

Program success in FCS and SDS is on par with other programs....

...and program review completion has caught up.

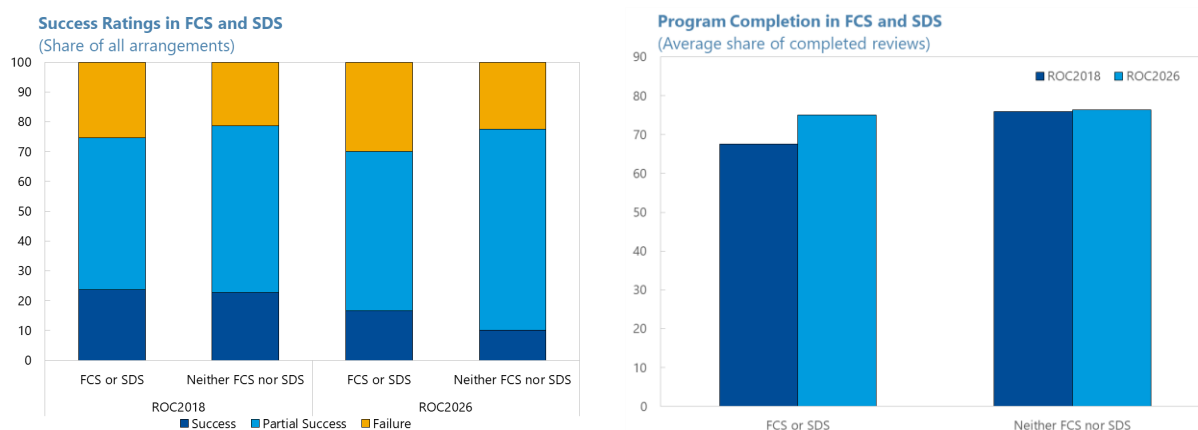
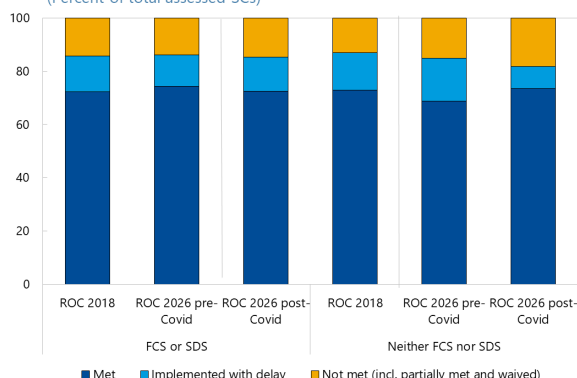
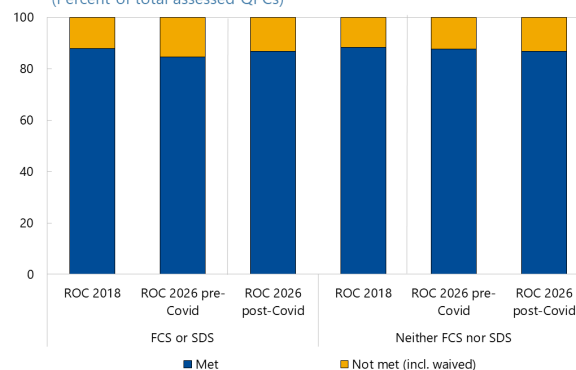
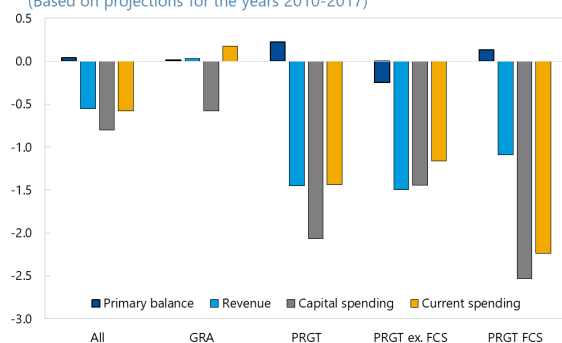
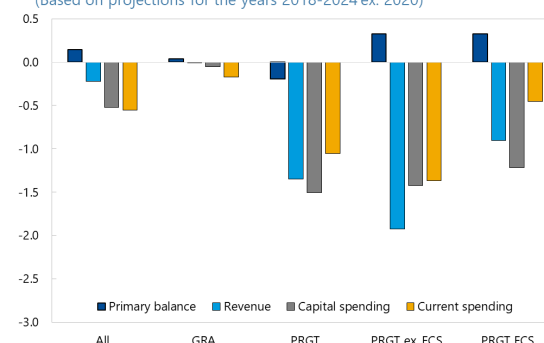


Figure 39. Program Outcome in FCS and SDS (concluded)*Structural implementation remains high...***FCS and SDS : Structural Implementation**
(Percent of total assessed SCs)*... as does quantitative implementation.***FCS and SDS : Quantitative Implementation**
(Percent of total assessed QPCs)*PRGT programs often saw revenue underperformance compensated by spending cuts, in the 2018 ROC sample...***Median 1-Year Ahead Forecast Errors**
(Based on projections for the years 2010-2017)*...and the 2026 ROC sample. Undershooting by FCS countries appears to have diminished in recent years.***Median 1-Year Ahead Forecast Errors**
(Based on projections for the years 2018-2024 ex. 2020)

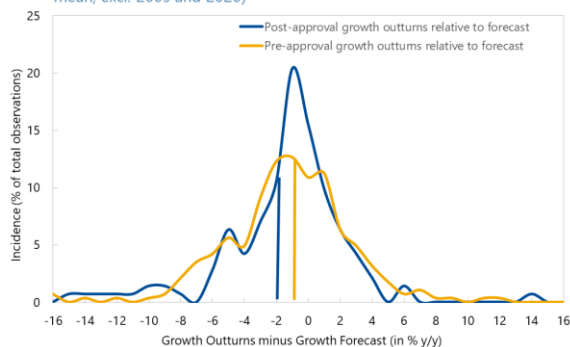
Sources: MONA database; IMF WEO; IMF staff calculations.

Figure 40. Program Design in FCS and SDS

2018 ROC: For FCS, historical growth outturns were below forecasts, on average. Realized growth turned out even weaker than the historical average.

FCS Programs Approved Between 2011 and 2017

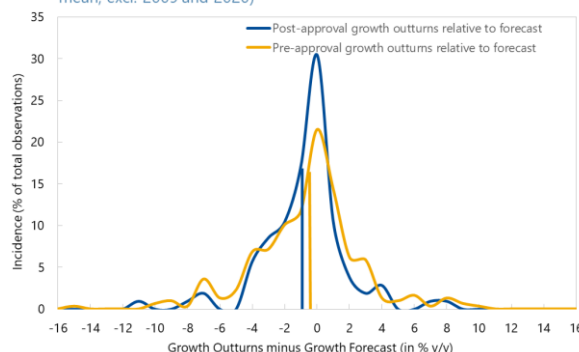
(Growth outturns minus 3-year ahead average forecast, vertical line = mean; excl. 2009 and 2020)



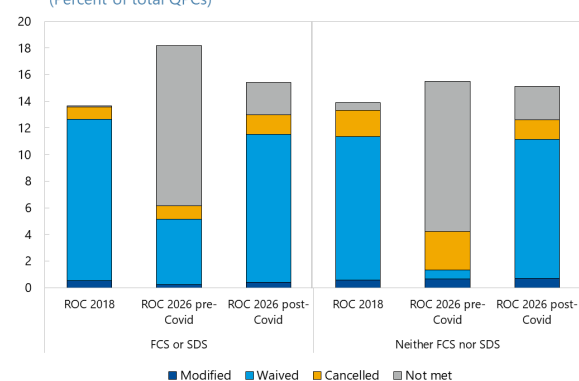
2026 ROC: Growth optimism for FCS has declined but remains substantial.

FCS Programs Approved Since 2018

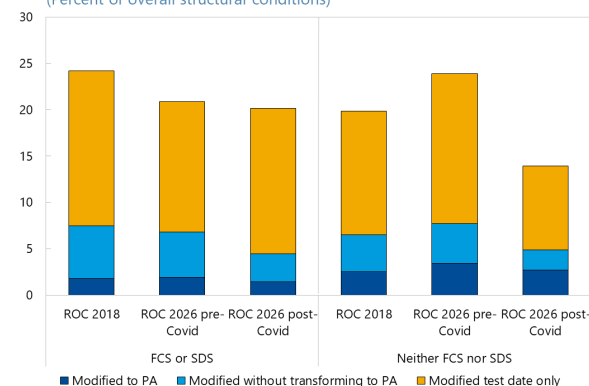
(Growth outturns minus 3-year ahead average forecast, vertical line = mean; excl. 2009 and 2020)



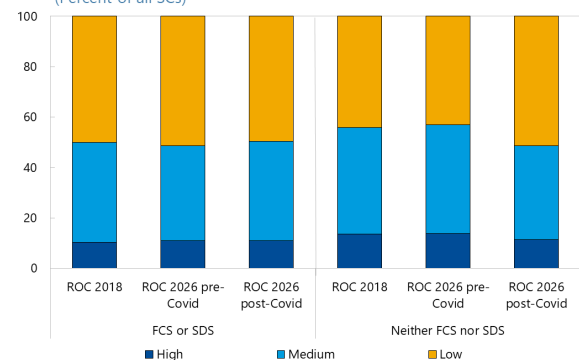
Program targets in FCS and SDS adapted to shocks as in other programs...

Share of Modified Quantitative Performance Criteria in FCS and SDS
(Percent of total QPCs)


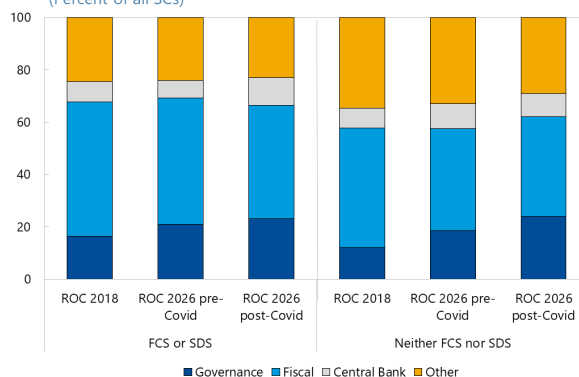
...and while timelines for SC implementation are routinely extended, a lower share of SBs were modified to PA

Share of Modified Structural Conditions in FCS and SDS
(Percent of overall structural conditions)


SC in FCS and SDS has similar depth to other programs...

FCS and SDS : Structural Depth
(Percent of all SCs)


...with a greater sectoral focus on fiscal and governance.

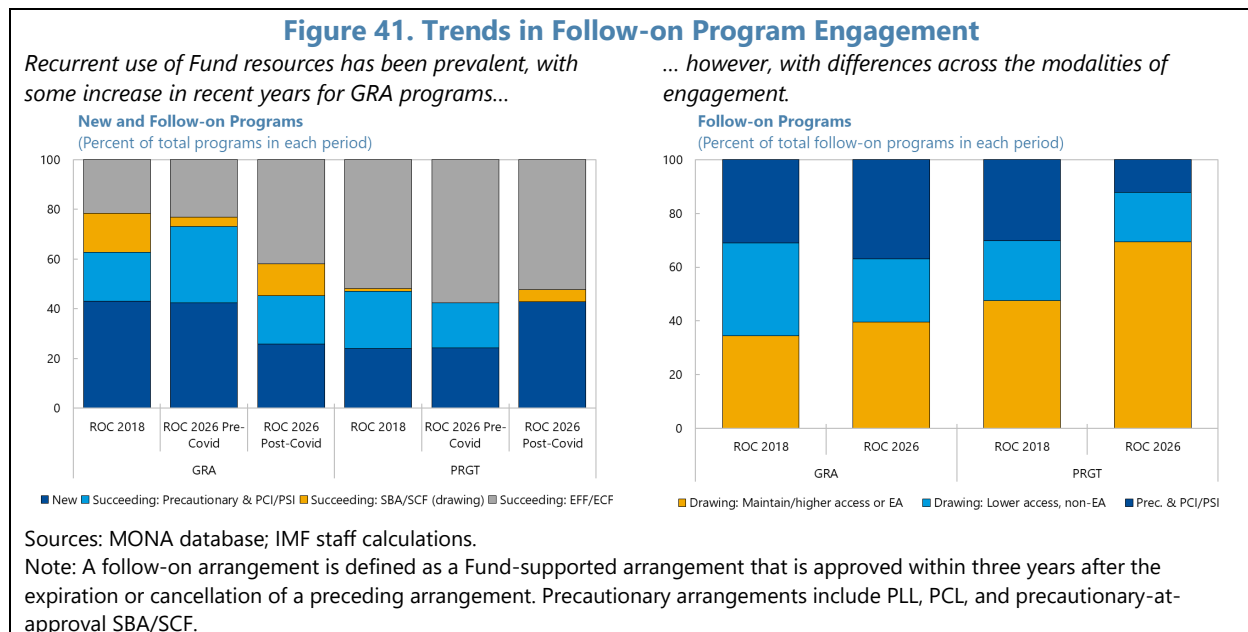
FCS and SDS : Sectoral Composition
(Percent of all SCs)


Sources: MONA database; IMF WEO; IMF staff calculations.

SECTION X. FOLLOW-ON PROGRAM ENGAGEMENT⁶⁴

Follow-on program engagements have been on the rise and raise important considerations regarding their effectiveness in resolving BOP problems and restoring medium-term external viability. The evidence points to substantial variation across such cases, with some follow-on programs building on earlier achievements and supporting continued reform efforts.

80. The 2026 ROC sample shows notable shifts in the modalities and access levels of follow-on programs (Figure 41).⁶⁵ Among follow-on GRA arrangements, which make up 70 percent of GRA programs in this ROC, there has been a shift away from SBAs/SCFs, which have more than halved, mainly towards EFFs—that tackle serious medium-term BOP problems (i.e., of a structural nature or because of slow growth and an inherently weak BOP position)—and signaling and precautionary instruments that now account for about one-third of follow-on GRA cases. About 65 percent of PRGT programs are follow-on arrangements, with ECFs remaining the dominant modality. The share of precautionary follow-on arrangements has declined, while ECF usage has risen to about 85 percent. The share of follow-on PRGT arrangements with the same or higher access than the preceding program has increased significantly—from about 40 to 60 percent—reflecting persistent financing needs, whereas the increase under the GRA has been more modest.



⁶⁴ Prepared by Fah Jirasavetakul, Gaelle Pierre, Faezeh Raei, and Yulia Ustyugova, with contributions from Alex Grohovsky, Yuko Hashimoto, and Hyunmin Park, and research assistance from Paula Arias and Yiran Chen.

⁶⁵ Recurrent program engagement with the Fund can be assessed in various ways. In this analysis, recurrent program engagement is defined as a program succeeded by a follow-on program within three years after its expiration or cancellation. Another metric used is the Long-Term Program Engagement (LTPE), which refers to a member that spent at least 7 of the past 10 years under Fund-supported disbursing programs. The LTPE brings a longer-term view, thereby enriching the analysis for PRGT-eligible countries because their protracted BOP problems often require sustained, sequenced support, but excludes transitions to non-disbursing/signaling programs.

81. The rise in follow-on arrangements can reflect a range of underlying factors, including recurrent shocks and the scale and persistence of external and structural challenges. In a low-growth and shock-prone environment, members may face new BOP needs from new shocks, or they may seek continued engagement to safeguard earlier gains including through precautionary or lower access arrangements, or they may need to complete adjustment and reform plans, or it may take more time for measures to yield expected benefits. In the PRGT, the concept of a “protracted BOP problem” is embedded in the design of the ECF, recognizing that many LICs face deep-rooted structural issues, requiring longer time to resolve. In practice, successive arrangements are often used. The need for follow-on arrangements ex post can therefore arise even when meaningful progress has been made under preceding programs.

82. To assess the nature and effectiveness of recurrent engagement, follow-on programs can be grouped for analytical purposes based on the performance of the preceding program. In GRA programs, the access level of the follow-on arrangement can serve as an additional indicator since sustained progress should improve market access, thereby lowering the need for Fund financing in subsequent programs. As illustrated in Figure 42, follow-on programs are classified as:

- *“Successor programs” that follow preceding programs, which are classified as successful or partially successful.⁶⁶ They typically build on improved conditions and enhanced resilience, reflected in stronger policy buffers. In the GRA context, successor programs are also expected to involve lower access levels or take the form of precautionary arrangements, signaling reduced financing needs, owing to improved market access and greater policy credibility.*
- *“Repeated programs” are programs that follow unsuccessful preceding programs. In GRA cases, such repeated programs involve requests for the same or higher access, potentially suggesting limited traction in addressing underlying vulnerabilities or implementing durable reforms under previous arrangements.*

Figure 42. Distinguishing Follow-on Program Engagement

GRA programs		PRGT programs	
		Progress in preceding program	
		Some success	Unsuccess
Followed by follow-on program with	lower access or precautionary	Successor	Repeated
	higher access	Repeated	Repeated
		Some success	Unsuccess
		Successor	Repeated

Sources: IMF staff estimations

Notes: See Background Paper, Section I on the program success definition

⁶⁶ For GRA programs, the success rating precludes categorizing a program as fully successful if it is followed by another disbursing arrangement, even if vulnerabilities decline.

83. Successor programs represent most of PRGT follow-on programs, while their share has declined in GRA cases. In the 2026 ROC sample, successor programs represented 78 percent of PRGT follow-ons—slightly down from 85 percent in the 2018 ROC.⁶⁷ For GRA programs, the share of successor programs fell from 38 percent in 2018 ROC to 30 percent in 2026 ROC, despite over half of preceding programs being classified as full or partial success. The declining share reflects the higher proportion of follow-on programs maintaining or increasing access levels, often in response to elevated shocks and financing needs.

84. Successor programs in the 2026 ROC built on progress in their preceding programs through stronger policy adjustments and structural reforms (Figure 43 and 44).

- In GRA programs, successor arrangements followed preceding programs with a more prudent fiscal stance, smaller increases in public debt, stronger external adjustment, and greater reserve accumulation—despite global shocks. The preceding programs also showed stronger implementation of structural conditions (SCs). Conversely, repeated programs were preceded by weaker fiscal performance and deteriorating external positions. Fund exposures among successor GRA programs remained contained, but repeated GRA programs involved repeated access and rising exposure.
- In PRGT programs, successor arrangements built on preceding programs with stronger fiscal and external adjustments, lower inflation, and more reserve accumulation compared to repeated programs. Preceding programs associated with successor arrangements also featured stronger implementation of SCs, which may reflect more sustained reform efforts that supported macroeconomic outcomes. Despite progress, public debt increased across both groups of programs, reflecting financing needs related to the COVID-19 pandemic and other shocks. PRGT credit outstanding rose more broadly, reflecting large BOP needs in the shock-prone and uncertain environment.⁶⁸

85. Meanwhile, longer-term program engagement (LTPE) among PRGT-eligible countries has declined significantly and become concentrated among the poorest members.⁶⁹ The number of PRGT-eligible countries with LTPE status fell from nearly 30 in early 2000 to just 8 in 2024, assisted by milestones achieved under the debt relief initiatives and the graduation of some countries from PRGT eligibility. The PRGT-eligible LTPE countries are primarily the poorest and most vulnerable members, often operating in fragile contexts.⁷⁰ LTPE status is associated with larger initial macroeconomic imbalances, fragility, and weaker government effectiveness. Compared to other PRGT programs, countries with LTPE status in 2019–24 show a higher incidence of program unsuccess—by about 10 percentage points—linked to lower program completion rates, often associated with fragile contexts. While these countries performed relatively well in domestic revenue mobilization and social spending—key in fragile settings—they struggled to boost growth,

⁶⁷ Preceding programs without success ratings are dropped from this analysis.

⁶⁸ For a detailed discussion of Fund exposure to PRGT-eligible countries ([IMF, 2024e](#)).

⁶⁹ See Footnote 65.

⁷⁰ The vast majority of these countries have GNI per capita at or below the IDA operational threshold (US\$1,335 in FY25), and are either currently classified as fragile and conflict-affected states or have a recent history of fragility.

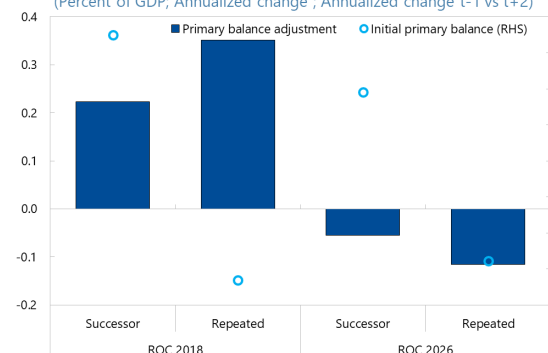
complicating macroeconomic adjustment. There is no evidence that PRGT programs with LTPE featured greater adaptability in conditionality, as measured by the share of modified SCs.

Figure 43. GRA Successor vs Repeated Programs

Relative to repeated programs, GRA successor programs often built on programs with stronger fiscal discipline, ...

GRA. Fiscal Adjustment of Preceding Programs

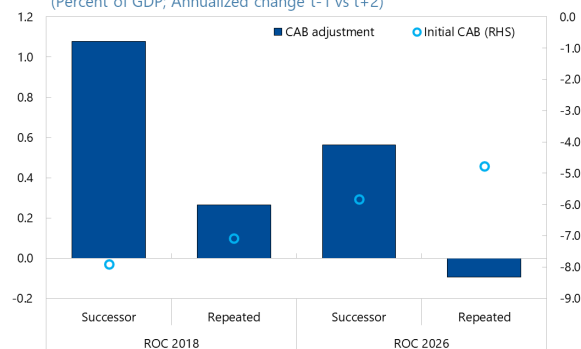
(Percent of GDP; Annualized change; Annualized change t-1 vs t+2)



..., more robust external adjustments, ...

GRA. External Adjustment of Preceding Programs

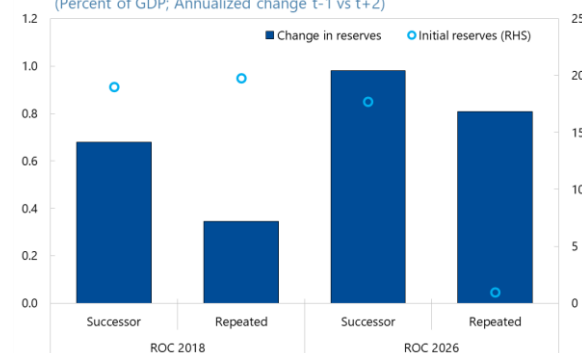
(Percent of GDP; Annualized change t-1 vs t+2)



... greater reserves accumulation, ...

GRA. Change in Reserves of Preceding Programs

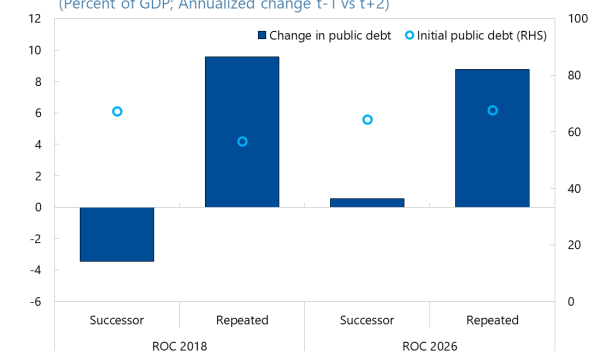
(Percent of GDP; Annualized change t-1 vs t+2)



... and enhanced resilience against debt increase during the period of shocks, ...

GRA. Change in Public Debt of Preceding Programs

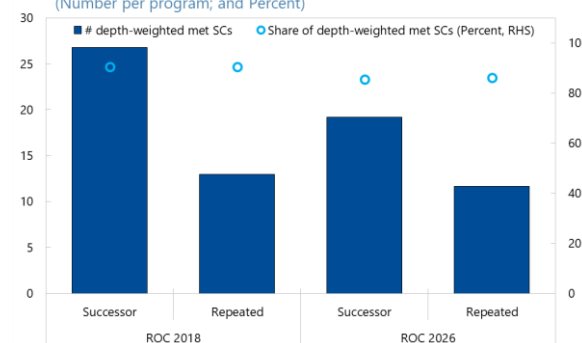
(Percent of GDP; Annualized change t-1 vs t+2)



... along with robust implementation of structural conditionality.

GRA. Structural Conditionality of Preceding Programs

(Number per program; and Percent)



GRA successor programs saw declining Fund exposure in the previous ROC period and a more moderate increase in the current ROC period than repeated programs.

GRA. Change in Credit Outstanding of Successor Programs

(Percent of quota)

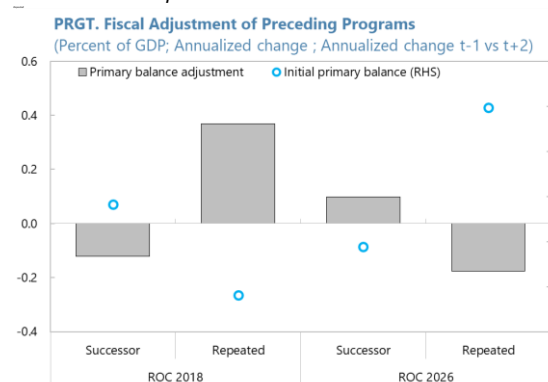


Sources: MONA database; IMF WEO; IMF staff calculation.

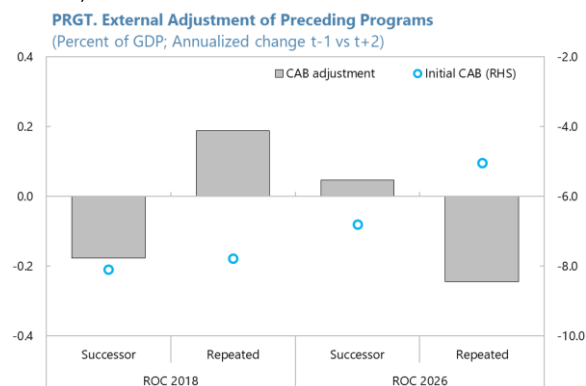
Note: A preceding program has a follow-on arrangement within three years after the program ends or expires. Excluding more recent programs in the 2026 ROC sample with no information to assess their success ratings.

Figure 44. PRGT Successor vs Repeated Programs

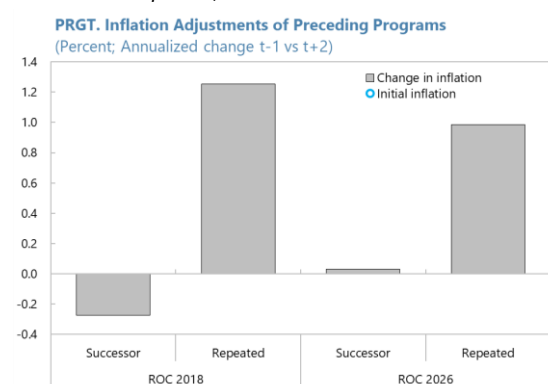
Successor programs in 2026 ROC built on programs that delivered some fiscal...



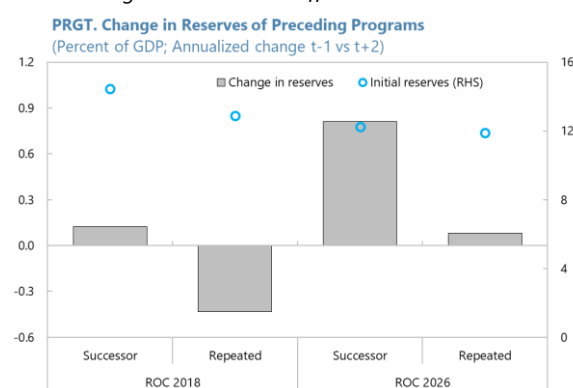
... and external adjustment despite challenging external context...



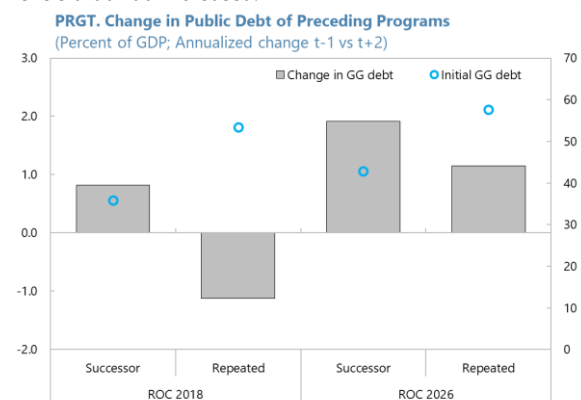
...contained inflation...



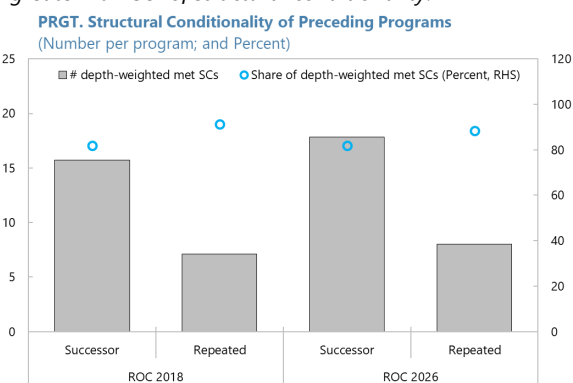
... and strengthened reserve buffers.



Both successor and repeated programs faced public debt levels that had increased.



Successor arrangements built on programs that met a greater number of structural conditionality.



Sources: MONA database; IMF WEO; IMF staff calculation.

Notes: A preceding program has a follow-on arrangement within three years after the program ends or expires. Excluding more recent programs in 2026 ROC sample with no information to assess their success ratings.

SECTION XI. ENTERPRISE RISK ASSESSMENTS IN UFR⁷¹

Experience has shown that a sound Enterprise Risk Assessment facilitates a discussion of major and critical program-related risks with the Executive Board that might otherwise be too sensitive to cover in official contingency planning. Summarizing the risks and their mitigants in a designated section of staff reports can bring clarity, focus, a basis for comparison, and a point of accountability to risk-based deliberations.

86. Enterprise Risk Assessments (ERA) have become a key tool for supporting risk-based decision making in IMF lending arrangements.⁷² Designed for internal audiences, including the Executive Board, they allow for a candid evaluation of risks early in program discussions without compromising the catalytic role of Fund engagement. Their confidential nature encourages frank analysis, particularly important amid frequent shocks and heightened uncertainty. ERAs use a standardized framework, often supported by the Document Risk Self-Assessment (DRSA) tool, based on the enterprise risk management (ERM) taxonomy and ratings scales. Reporting has evolved, to include stand-alone ERA supplements for high-risk programs where public disclosure could undermine program credibility. As of end-2024, ERAs have been conducted for about one-quarter of countries with active IMF programs.

87. The experience with ERAs highlights their value in Use of Fund resources (UFR) cases involving major enterprise risks:

- **Clarity and structure.** ERAs consolidate key enterprise risk considerations—often scattered across documents—into a single, organized assessment. This includes risks not typically covered in staff reports, such as reputational or model risks and evenhandedness.
- **Clear bottom lines.** Risk ratings and structured analysis support consistent, transparent assessments across cases and over time.
- **Accountability.** ERAs enhance internal decision-making by clearly documenting risks for Management and the Board, reinforcing accountability within the Fund’s governance framework.
- **Portfolio risk monitoring.** Standardized taxonomy and ratings enable aggregation across programs, improving institutional tracking and informing Risk Reports.

88. ERAs for UFR cases can help identify key risks and mitigation strategies relevant to Board decisions on new program requests and reviews. While early use focused on whether to proceed with program requests (the focus of this analysis), ERM is increasingly informing other decisions, such as adjustment of program modalities. Importantly, ERAs highlight key risk trade-offs

⁷¹ Prepared by Maxym Kryshko, Juliana Araujo, and Marina Rousset.

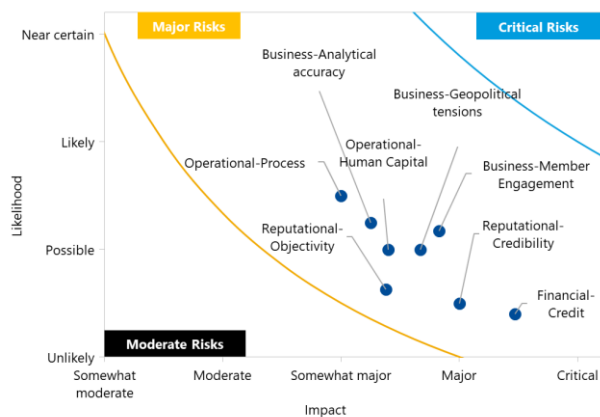
⁷² In December 2022, the Board adopted a new policy on Enterprise Risk Management (ERM) to manage enterprise-wide risks and enhance decision-making for the Fund (SM/22/206, Rev 1, 11/16/2022). The ERM policy includes assessments of risks related to lending, surveillance, and internal processes. The approach to ERAs is discussed in the ERM guidance (SM/25/117).

and intertemporal dynamics; for instance, measures that mitigate member engagement risks in the short term may inadvertently increase financial risks in the medium term.

89. The risks associated with having a new arrangement or proceeding with a review typically included (Figure 45):

- **Strategic risks.** Examples include risks to Fund’s PRGT finances, failure to restore debt sustainability, perception of evergreening, and strains on the Fund’s lending framework owing to frontloading of access. While infrequent, their impact can be significant. Mitigants include strong policy track record, sound program design, and adequate financing assurances.
- **Business risks.** These stem from weak program design or implementation challenges. Programs may fall short in resolving external imbalances within the envisioned timeframe, or may go off-track due to partial or incomplete implementation of policies, policy reversals, or socio-political instability. These risks are partly mitigated by the review-based nature of Fund lending with risk-based periodicity, strong program ownership, and upfront policy actions. Additional safeguards include a robust and targeted communication strategy to build public and stakeholder buy-in, broad political support, measures to protect the vulnerable, and contingency planning.
- **Financial risks.** In the short term, these include high Fund credit concentration and potential arrears to the Fund if programs go off track with thin reserves. Medium term risks arise from the prolonged use of Fund resources and the program’s failure to restore medium-term viability, either due to policy slippages or exogenous factors. Mitigants include targeted macro-critical conditionality, phased access, burden sharing, adequate financing assurances, reserve buildup strategy, and commitment to service debt.
- **Reputational risks.** Engagement with prolonged users or countries facing non-economic shocks may raise concerns about evenhandedness. Reviews completed despite limited implementation of reforms, or program collapses amid social unrest, can also harm the Fund’s credibility. Effective communication, demonstrated ownership, calibrated conditionality, and targeted CD to address capacity gaps can help manage these risks.

Figure 45. Enterprise Risk Assessments for Having a New Arrangement or Proceeding with a Review (GRA Lending Portfolio)
(Based on impact and likelihood score)



Sources: Enterprise Risk Assessments as of Mid-Year Risk Update 2025 for GRA.

Notes: Excluding risks with a sample size of less than 2. The area above the yellow line captures major risks. The area above the blue line captures critical risks.

- **Operational risks.** These involve process and data issues (lags and variability in data sources, data limitations and quality concerns, delayed or partial sharing of data, data misreporting) and risks to staff safety and well-being during missions.

90. Not entering into a Fund arrangement or delaying a review can also carry significant risks:⁷³

- **Financial risks.** For countries with prolonged Fund engagement, program interruptions can leave underlying imbalances unresolved, which would raise the risk of arrears to the Fund, particularly when reserve buffers are low.
- **Business risks.** Delayed engagement can mean missing the window to provide timely and adequate financial assistance and policy advice to a member facing balance of payments difficulties, potentially worsening the crisis and triggering regional spillovers.
- **Strategic risks.** A failure to act may lead to disengagement by the member, aggravating stigma and pushing the country to alternative, possibly costlier financing sources. This could exacerbate geopolitical tensions or threaten to reverse initial macroeconomic gains (e.g., via monetary financing).
- **Reputational risks.** Not proceeding with the program or review may raise concerns about evenhandedness, especially if the member had a strong track record. Inaction could damage the Fund's credibility and erode confidence in its ability to deliver on its mandate.

91. The value of ERAs in informing the Board discussion of program risks depends on how the risk management questions and issues are framed and answered. Based on experience with pilots, ERAs have informed the Board discussion of the risks of proceeding, or not, or with delay, at program approval or in reviews.

⁷³ The Fund's decision to proceed with an arrangement or with the completion of a review requires that all the legal requirements and applicable Fund policies, including requirements for adequate safeguards for the temporary use of the Fund resources, are met. Where these requirements are not met, the Fund cannot proceed with the arrangement or completion of the review.

SECTION XII. STAKEHOLDER SURVEY: PRELIMINARY VIEWS⁷⁴

Staff sought high-level views on IMF program design and conditionality through a targeted stakeholder survey. Based on responses received to date, stakeholders generally considered that Fund-supported programs improved economic and financial conditions and supported countries affected by external shocks. Program conditionality was broadly seen as focused, evenhanded, and appropriately tailored to country-specific circumstances, with the IMF widely perceived as a trusted adviser to country authorities and official creditors.

92. The survey centered on the design and conditionality of IMF-supported programs, timely and effective support for countries hit by shocks, and the IMF's role as a trusted advisor (Figure 46). Staff surveyed country authorities, Executive Directors, and IMF mission chiefs and resident representatives for all IMF member countries. Selected academic experts, private sector representatives, and civil society organizations were also invited to participate. As of August 12, staff received responses from 100 country authorities, 16 Executive Directors, 65 mission chiefs, and 25 resident representatives. Responses received so far represent views from 86 out of 191 IMF member countries. The survey will remain open through September to assemble the most complete picture of stakeholder views possible.

A. Effective and Timely Support

93. Stakeholders overwhelmingly agreed that Fund-supported programs were effective in improving economic and financial prospects. Stakeholders credited IMF programs with contributing to improved outcomes in debt sustainability, financial sector reforms, monetary policy and reserves targets and to a lesser extent, structural reforms. Notably, 9 in 10 country authorities so far agreed or strongly agreed that IMF programs were effective in improving economic and financial prospects. In addition, 87 percent of respondents agreed that IMF-supported programs improved fiscal policy outcomes.

94. Fund-supported programs were credited with providing timely and effective support to member countries during external shocks. The Fund was viewed most favorably for its response during recent health emergencies such as the Ebola outbreak and the COVID-19 pandemic, with 82 percent of respondents deeming its support timely and effective.⁷⁵ A majority of stakeholders also held a positive view of IMF support following natural disasters and episodes of global financing tightening, with weaker perception on timeliness and effectiveness when

⁷⁴ Prepared by Rimtautas Bartkus and Irene Yackovlev, with research assistance from Helen Zheng, and ITD support from Kevin Bailey.

⁷⁵ This result is consistent with the findings of the IEO Evaluation on the [IMF's Emergency Response to the COVID-19 Pandemic](#).

confronting geopolitical shocks or trade policy uncertainty. About 70 percent of respondents viewed Fund program design as providing the ability to adjust to unforeseen shocks.

B. Focused, Evenhanded, and Tailored Conditionality

95. Program conditionality was considered focused and evenhanded. Over three-quarters of respondents agreed or strongly agreed that program conditionality was focused on areas necessary for program success. Based on the responses to date, concerns about uniformity of treatment among some stakeholders remain, with about one-tenth of respondents in the 2026 ROC survey disagreeing or strongly disagreeing that lending policies ensure similar conditionality across countries with similar characteristics. However, sentiment among country authorities improved from 39 percent positive responses in the 2018 survey to 57 percent in 2026.⁷⁶

96. Fund-supported programs were largely viewed as appropriately tailored to country-specific circumstances. About 94 percent of respondents agreed that that Fund-supported programs appropriately considered the country-specific nature of economic and financial vulnerabilities in their design. IMF-supported programs were also perceived by significant majorities as considering country-specific needs for policy adjustment, debt sustainability, and domestic priorities and needs. However, the perception of consideration of reform willingness and implementation capacity, and social and political conditions in program countries was weaker. Stakeholders largely viewed the balance between financing needs and policy adjustments in the conditionality of Fund-supported programs as appropriate.

97. Perceptions of structural reforms in the context of IMF-supported programs were mixed. Areas of concern included the consideration of the distributional impact of these measures and their number. This is consistent with anecdotal evidence that in some program cases observers have called out structural reform commitments as too onerous and insufficiently calibrated to capacity constraints.

C. Trusted Advisor

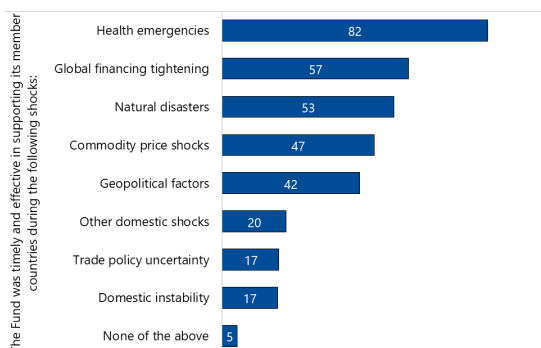
98. Stakeholders decisively viewed the IMF as a trusted advisor to country authorities and official creditors. Over 90 percent of respondents agreed or strongly agreed that the IMF was a trusted advisor to its member countries and official creditors. However, the lower levels of trust perceived by leading thinkers and academics as well as by civil society organizations also suggest that, especially in complex program cases, there is scope for improving outreach and communication with these groups.

⁷⁶ The results are not fully comparable as all country authorities were surveyed while in 2018 only program countries were surveyed. Still, responses gathered so far suggest that restricting the analysis to just program country authorities would not change the broad results.

Figure 46. IMF Survey of Stakeholders: Views on Outcomes, Programs, and Trust

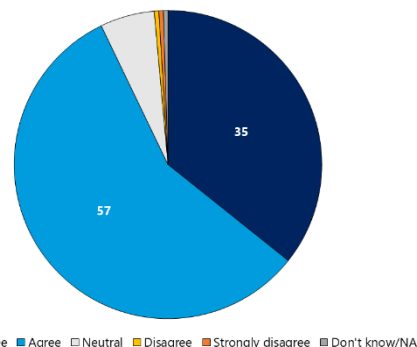
Stakeholders recognized timely and effective support during shocks, especially health emergencies....

Capacity to Absorb Shocks
(Percent of responses)



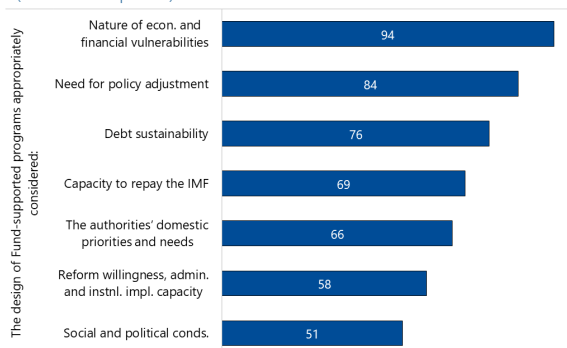
...and overwhelmingly shared a positive perception of the IMF as a trusted advisor to country authorities.

IMF as a Trusted Adviser to Country Authorities
(Percent of responses)



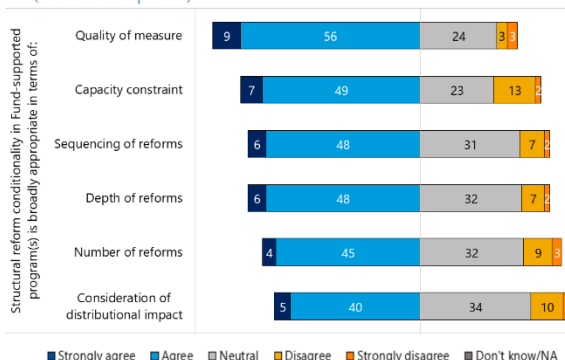
Stakeholders viewed programs as especially tailored to the nature of vulnerabilities and need for policy adjustment...

Tailoring to Country Circumstances
(Percent of responses)



...at the same time, views were mixed on whether structural conditionality was appropriate.

Structural Conditionality
(Percent of responses)



Source: IMF staff calculation based on 2026 ROC Stakeholder Survey responses received as of August 12, 2025.

Notes: As of August 12, 2025, inputs to survey questions were received from authorities from 86 IMF member countries, 16 Executive Directors, 65 Mission Chiefs, 25 Resident Representatives, 4 Academic Experts, and 8 Civil Society Organizations.

References

- Bailey, M. A., Strezhnev, A., & Voeten, E., 2017. [Estimating dynamic state preferences from United Nations voting data](#). *Journal of Conflict Resolution*, 61(2), 430–456.
- Bhalla, S., Bhasin, K., & Loungani, P., 2023. [Macro effects of formal adoption of inflation targeting](#). (IMF Working Paper No. 2023/007)., International Monetary Fund.
- Bird, G., Qayum, F., & Rowlands, D., 2021. [The effects of IMF programs on poverty, income inequality and social expenditure in low income countries: an empirical analysis](#). *Journal of Economic Policy Reform*.
- Blanchard, O. J., & Leigh, D., 2013. [Growth forecast errors and fiscal multipliers](#). *American Economic Review*, 103(3), 117–120.
- Cengiz, D., Dube, A., Lindner, A., & Zipperer, B., 2019. [The Effect of Minimum Wages on Low-Wage Jobs](#). *The Quarterly Journal of Economics*, 134(3), 1405–1454.
- Clements, B., Gupta, S., & Nozaki, M., 2013. [What happens to social spending in IMF-supported programmes?](#) *Applied Economics*, 45(28), 4022–4033.
- Culiuc, A., & Park (forthcoming). *Currency Crises in the Post-Bretton Woods Era: A New Dataset of Large Depreciations*. IMF Working Paper.
- Dincer, N., Eichengreen, B., & Geraats, P., 2019. [Transparency of Monetary Policy in the Post Crisis World](#). Oxford University Press.
- Duncan, R., Martínez García, E., & Toledo, P., 2022. [Just Do IT? An assessment of inflation targeting in a global comparative case study \(Globalization Institute Working Paper No. 418\)](#). Federal Reserve Bank of Dallas. (Dallas)
- Eicher, T. S., Eskimez, R. K., & Newiak, M., 2024. [Effects of IMF-supported programs on gender inequality \(IMF Working Paper No. 2024/247\)](#). International Monetary Fund.
- Hastie, T., Tibshirani, R., & Friedman, J., 2009. [The Elements of Statistical Learning: Data Mining, Inference, and Prediction \(2nd ed.\)](#). Springer.
- He, D., Kim, J., & Kyobe, A., 2024. [The catalytic impact of IMF lending on Official Development Assistance \(IMF Working Paper No. 2024/134\)](#). International Monetary Fund.
- Ho, G., & Mauro, P., 2016. [Growth—Now and Forever? IMF Economic Review](#), 64(3), 526–547.
- Independent Evaluation Office, 2008. [An IEO Evaluation of Structural Conditionality in IMF-Supported Programs](#). (Washington).

———, 2021. [Growth and adjustment in IMF-supported programs \(Evaluation Report\)](#). (Washington).

———, 2023. [The IMF's Emergency Response to the COVID-19 Pandemic](#). (Washington).

International Monetary Fund, 1984. Decision No. 7842-(84/165): [Misreporting and noncomplying purchases in the General Resources Account—Guidelines on corrective action](#). In Selected Decisions and Selected Documents of the International Monetary Fund, Forty-Third Issue. (Washington).

———, 1987. Decision No. 8759-(87/176) ESAF: [Instrument to Establish the Poverty Reduction and Growth Trust](#). In Selected Decisions and Selected Documents of the International Monetary Fund, Forty-Third Issue. (Washington).

———, 1999. [Inflation targeting – Implications for IMF conditionality](#) (Policy Paper No. PP/99/78). International Monetary Fund. (Washington)

———, 2009. [Debt limits in fund-supported programs — Proposed new guidelines: Supplementary information and proposed decision](#) (Policy Paper No. 2009/08). (Washington).

———, 2012a. [2011 Review of Conditionality—Overview Paper](#). (Washington).

———, 2012b. [2011 Review of Conditionality Background Paper 3: Outcomes of Fund-Supported Program](#). (Washington).

———, 2014. [Conditionality in evolving monetary policy regimes](#). (Policy Paper No. PP/14/3.). (Washington).

———, 2018. [Review of 1997 guidance note on governance: A proposed framework for enhanced Fund engagement](#). (Policy Paper No. PP/03/0918). (Washington).

———, 2019a. [2018 Review of Program Design and Conditionality](#). (Policy Paper No. 19/012). (Washington).

———, 2019b. [A strategy for IMF engagement on social spending](#). (Policy Paper No. 2019/016). (Washington)

———, 2021. [How to Assess Country Risk: The Vulnerability Exercise Approach Using Machine Learning](#). (Technical Notes and Manuals No. 2021/003). (Washington)

———, 2022a. [IMF Engagement on Pension Issues in Surveillance and Program Work](#). (Technical Notes and Manuals No. 2022/004). (Washington).

———, 2022b. [IMF Engagement on Social Safety Net Issues in Surveillance and Program Work](#). (Technical Notes and Manuals No. 2022/007). (Washington).

- , 2022c. [IMF Strategy Toward Mainstreaming Gender](#) (Policy Paper No. 2022/037). (Washington)
- , 2023a. [IMF Engagement on Health Spending Issues in Surveillance and Program Work](#). (Technical Notes and Manuals No. 2023/008). (Washington).
- , 2023b. [Structural reforms to accelerate growth, ease policy trade-offs, and support the green transition in emerging market and developing economies](#). (Staff Discussion Note No. 2023/007). (Washington).
- , 2023c. [Staff Guidance Note on The Implementation of The IMF Strategy for Fragile and Conflict-Affected States \(FCS\)](#). (Policy Paper No. 2023/010). (Washington).
- , 2024a. [Operational Guidance Note for IMF Engagement on Social Spending Issues](#). (Policy Paper No. 2024/022). (Washington).
- , 2024b. [Operational Guidance Note on Program Design and Conditionality](#). (Policy Paper No. 2024/004). (Washington).
- , 2024c. [Review of the Fund’s capacity development strategy—Towards a more flexible, integrated, and tailored model](#). (Policy Paper No. 2024/014). (Washington).
- , 2024d. [Staff Guidance Note On The IMF’s Engagement With Small Developing States](#). (Policy Paper No. 2024/035). (Washington).
- , 2024e. [2024 Review Of The Poverty Reduction And Growth Trust Facilities And Financing — Reform Proposals](#). (Policy Paper No. 2024/047). (Washington).
- , 2025a. [IMF Engagement on Education Spending in Surveillance and Program Work](#). (Technical Notes and Manuals No. 2025/007). (Washington).
- , 2025b. [Seychelles: Fourth Reviews Under the Arrangement Under the Extended Fund Facility and the Arrangement Under the Resilience and Sustainability Facility; Request for Modification of Performance Criteria; and Request for Modification of Reform Measures Under the Resilience and Sustainability Facility](#). (IMF Country Report No. 25/147). (Washington.)
- Ismail, K., Perrelli, R., & Yang, J., 2020. [Optimism bias in growth forecasts—The role of planned policy adjustments](#) (IMF Working Paper No. 2020/229). International Monetary Fund.
- Kim, J. Y., & Lee, K. Y., 2021. [Structural Conditions, Structural Reforms and Growth in IMF Supported Programs](#). (IEO Background Paper No. BP/21 01/04). International Monetary Fund. Retrieved from

- Krahnke, T., 2023. [Doing more with less: The catalytic function of IMF lending and the role of program size](#). *Journal of International Money and Finance*, 135, Article 102856.
- Kuruc, K., 2022. [Are IMF rescue packages effective? A synthetic control analysis of macroeconomic crises](#). *Journal of Monetary Economics*, 127(2), 38–53.
- Laeven, L., & Valencia, F., 2020. [Systemic Banking Crises Database II](#). *IMF Economic Review*, 68, 307–361.
- Lang, V. F., 2021. [The economics of the democratic deficit: The effect of IMF programs on inequality](#). *The Review of International Organizations*, 16(3), 599–623.
- Mourmouras, A., Ivanova, A., Anayiotos, G. C., & Mayer, W., 2003. [What Determines the Implementation of IMF-Supported Programs?](#). (IMF Working Paper No. 03/008). International Monetary Fund.
- Newiak, M., & Willems, T., 2017. [Evaluating the impact of non-financial IMF programs using the synthetic control method](#) (IMF Working Paper No. 17/109). International Monetary Fund.
- Nguyen, T. C., Castro, V., & Wood, J. A., 2022. [A new comprehensive database of financial crises: Identification, frequency, and duration](#). *Economic Modelling*, 108, Article 105770.
- Taylor, A. M., & Jordà, Ó., 2013. [The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy](#). (CEPR Discussion Paper No. 9646). Centre for Economic Policy Research.