IEO

Independent Evaluation Office
of the International Monetary Fund

GROWTH AND ADJUSTMENT IN IMF-SUPPORTED PROGRAMS

EVALUATION REPORT **2021**

ABOUT THE IEO

Established in 2001, the Independent Evaluation Office (IEO) conducts independent and objective evaluations of the IMF's policies, activities, and products. In accordance with its terms of reference, it pursues three interrelated objectives.

▶ To support the Executive Board's institutional governance and oversight responsibilities, thus contributing to accountability.

▶ To enhance the learning culture within the Fund by increasing the ability to draw lessons from experience.

▶ To strengthen the Fund's external credibility by enhancing transparency and improving understanding of the work of the IMF.

Independence is the fundamental anchor of the IEO's work. It is completely independent of the IMF's management team and staff, and operates at arm's length from the Executive Board. Its budget is separate from the Fund's (it accounts for about 0.5 percent of the institution's total budget), but subject to the same control procedures. The IEO is entitled to access any internal information and documents and does with very limited exceptions. The office's work is evaluated periodically by external experts.

For further information on the IEO and its ongoing and completed evaluations, please see our website IEO.IMF.org or contact the IEO at +(1) 202.623.8623 or at IEO@IMF.org.

IEO

Independent Evaluation Office
of the International Monetary Fund

GROWTH AND ADJUSTMENT IN IMF-SUPPORTED PROGRAMS

EVALUATION REPORT **2021**

CONTENTS

The following Background Papers are available on the IEO website at IEO.IMF.org:

BP/21-01/01. Cross-Country Analysis of Program Design and Growth Outcomes:
            2008–19

BP/21-01/02. Initiating Growth Surges: The Role of IMF-Supported Programs

BP/21-01/03. Fiscal Adjustment and Growth in IMF-Supported Programs

BP/21-01/04. Structural Conditionality, Structural Reforms and Growth in
            IMF-Supported Programs

BP/21-01/05. Exchange Rate Adjustment and Growth in IMF-Supported Programs

BP/21-01/06. Market Debt Operations and Growth in IMF-Supported Programs

BP/21-01/07. Growth and Adjustment in IMF-Supported Programs for Africa

BP/21-01/08. Growth and Adjustment in IMF-Supported Programs for Asia
            and Pacific

BP/21-01/09. Growth and Adjustment in IMF-Supported Programs for Europe

BP/21-01/10. Growth and Adjustment in IMF-Supported Programs for Middle East
            and Central Asia

BP/21-01/11. Growth and Adjustment in IMF-Supported Programs for
            Western Hemisphere

The following conventions are used in this publication:

▶ An en-dash (–) between years or months (e.g., 2021–22 or January–June) indicates the years or months covered, including the beginning or ending years or months.

▶ A slash (/) between years or months (e.g., 2021/22) indicates a fiscal or financial year, as does the abbreviation FY (e.g., FY2021).

▶ "Billion" means a thousand million; "trillion" means a thousand billion.

Some of the documents cited and referenced in this report were not available to the public at the time of publication of this report. Under the current policy on public access to the IMF's archives, some of these will become available three or five years after their issuance. They may be referenced as EBS/YY/NN and SM/YY/NN, where EBS and SM indicate the series and YY indicates the year of issue. Certain other types of documents may become available 20 years after their issuance. For further information, see IMF.org/external/np/arc/eng/archive.htm.

FOREWORD

This evaluation examines a critical and long-standing issue for IMF lending: how well have IMF-supported programs been able to sustain economic activity while delivering adjustment needed for external viability? Lessons from this evaluation are particularly relevant as many countries seeking IMF support in response to the COVID-19 pandemic face strong headwinds to growth.

After careful empirical analysis of IMF financing arrangements over the period 2008–19, the evaluation does not find evidence of a consistent bias towards excessive austerity in IMF-supported programs. Indeed, it finds that IMF-supported programs have yielded growth benefits during the program relative to a counterfactual of no Fund engagement and have boosted post-program growth outcomes.

Notwithstanding these positive findings, program growth outcomes consistently fell short of program projections. Greater scrutiny of the realism of program projections would certainly help to mitigate growth optimism, but even more important would be to achieve stronger and better growth outcomes by paying greater attention in program design and implementation to growth-friendly policies, including social and distributional consequences.

To shed light on how to meet this challenge, the evaluation assesses the extent to which different policy instruments were used to support program growth objectives. It finds that growth-friendly fiscal policies typically had only mixed success, including in protecting low-income and vulnerable groups. Structural conditionalities were of low depth and their potential growth benefits were not fully realized, suggesting a need to promote deeper, more growth-oriented reforms supported by more effective capacity development and stronger collaboration with the World Bank and other relevant partners. Use of the exchange rate instrument was relatively limited, while market debt operations were sometimes too little and too late.

Based on these findings, the report proposes three recommendations aimed at strengthening attention to growth implications of IMF-supported programs, including the social and distributional consequences. I am pleased that all three were endorsed broadly by the Managing Director and by the Executive Board during the Board discussion of the report. I look forward to more detailed decisions to move this agenda forward.

Charles Collyns
Director, Independent Evaluation Office

CONTRIBUTORS

This report was prepared by an IEO team led by Jun Il Kim, and including Jung Yeon Kim, Reginald Darius, Yasemin Bal Gündüz, Joshua Wojnilower, Jean-Marc Bedhat Atsebi, Kwang Yeon Lee, Jiakun Li, and Hasan Toprak.

The external consultants were Peter Allum, Ajai Chopra, Russell Kincaid, Mauro Mecagni, and Marco Pinon-Farah for the country case studies, and Aitor Erce and Sanjeev Gupta for the thematic papers.

The team is grateful to Annette Canizares, Arun Bhatnagar, Andrea Nicole Tumbaco, and Amy Gamulo for administrative assistance.

The evaluation benefited from discussions with participants at several workshops and many interviews with country officials and Fund staff. However, the final judgments are the responsibility of the IEO alone.

The report was approved by Charles Collyns.

ABBREVIATIONS

AREAER	Annual Report on Exchange Arrangements and Exchange Restrictions
ASCD	Average Depth Score of Structural Conditions
ASCG	Average Growth-Orientation Score of Structural Conditions
ASCI	Average Implementation Score of Structural Conditions
ASCID	Average Composite of Implementation and Depth Scores of Structural Conditions
ATE	Average Treatment Effect
BOP	Balance of Payments
CA	Current Account
CD	Capacity Development
CEMAC	Central African Economic and Monetary Community
DLP	Debt Limits Policy
DSA	Debt Sustainability Analysis
DSF	Debt Sustainability Framework
EA	Exceptional Access Arrangement
ECF	Extended Credit Facility
EDY	External Debt-to-GDP Ratio
EFF	Extended Fund Facility
ERPT	Exchange Rate Pass-Through
ERR	Exchange Rate Regime
ESAF	Enhanced Structural Adjustment Facility
ESF	Exogenous Shocks Facility
FAFE	Fiscal Adjustment Forecast Error
FSCI	Aggregate Implementation Score of Fiscal SCs
FSCID	Aggregate Implementation and Depth Score of Fiscal SCs
FSCIDG	Aggregate Implementation, Depth, and Growth-Orientation Score of Fiscal SCs
FSL	Fiscal Stability Law
FTE	Full-Time Equivalent
GFE	Growth Forecast Error
GEI	Government Effectiveness Index
GFC	Global Financial Crisis
GRA	General Resources Account
IDI	International Development Institution
LIA	Lending into Arrears

LIC	Low-Income Country
LIC-DSF	Low-Income Country Debt Sustainability Framework
LIOA	Lending into Official Arrears
MAC	Market Access Country
MAC DSA	Market Access Country Debt Sustainability Analysis
MONA	Monitoring of Fund Arrangements
NEER	Nominal Effective Exchange Rate
NER	Nominal Exchange Rate
NPL	Non-Performing Loan
NPV	Net Present Value
OGNC	Operational Guidance Note on Conditionality
PBY	Fiscal Primary Balance (as a share of GDP)
PDY	Public Debt-to-GDP Ratio
PFM	Public Financial Management
PRGF	Poverty Reduction and Growth Facility
PRGT	Poverty Reduction and Growth Trust
PSI	Policy Support Instrument
PUBINVY	Public Investment (as a share of GDP)
REER	Real Effective Exchange Rate
ROC	Review of Program Design and Conditionality
SBA	Stand-By Arrangement
SC	Structural Condition
SCF	Stand-By Credit Facility
SCI	Aggregate Implementation Score of SCs
SCID	Aggregate Implementation and Depth Score of SCs
SCIDG	Aggregate Implementation, Depth, and Growth-Orientation Score of SCs
SOCIALY	Social Spending (as a share of GDP)
SRI	Structural Reform Index
TA	Technical Assistance
VAT	Value-Added Tax
WAEMU	West African Economic and Monetary Union
WEO	World Economic Outlook
YGAP	Output Gap (in percent of trend GDP)

EXECUTIVE SUMMARY

This evaluation assesses how well IMF-supported programs have helped to sustain economic growth while delivering adjustment needed for external viability. It focuses on IMF financing arrangements over the period 2008–19, under both the General Resources Account (GRA) and the Poverty Reduction and Growth Trust (PRGT).[1] While the evaluation does not assess the experience during the COVID-19 pandemic, its lessons have become even more relevant as many countries now face strong headwinds to growth as they seek IMF-supported programs for achieving durable recoveries.

FINDINGS AND LESSONS

Under the Articles of Agreement, the IMF lends to countries to help correct balance of payments (BOP) problems without resorting to measures destructive of national prosperity. While IMF-supported programs give primary place to restoring external viability, attention to supporting activity during a program and fostering medium-term growth has increased over time, and particularly since the Global Financial Crisis (GFC) in 2008.

This increasing attention to the growth consequences of IMF-supported programs seems to have delivered some positive results. The evaluation does not find evidence of a consistent bias towards excessive austerity in IMF-supported programs during the evaluation period. Except in the crisis context, IMF-supported programs were in most cases able to sustain output broadly in line with historical norms while still delivering needed adjustment. Indeed, cross-country evidence suggests that programs have yielded growth benefits relative to a counterfactual of no Fund engagement and that stabilization and reforms implemented in the program context boosted post-program growth performance. Historical data over a longer time horizon suggest a positive role of IMF-supported programs at initiating sustained growth surges.

---

[1] Programs under the GRA provide non-concessional lending support, while programs under the PRGT provide lending on concessional terms to low-income countries.

Notwithstanding these positive findings, program growth outcomes consistently fell short of program projections, more so in GRA programs than in PRGT programs, broadly consistent with findings of the 2018 Review of Program Design and Conditionality (ROC). Of the programs covered in the evaluation, around one-half experienced an average growth shortfall (relative to initial program projections) during the program period of ½ percentage points or more, while one-fourth had a growth shortfall of over 1½ percentage points. Macro modeling errors, particularly those related to fiscal multiplier assumptions, seem to have been a source of such growth optimism, especially in GRA programs outside of a crisis context. While fiscal multiplier assumptions were broadly in line with the "bucket approach" suggested by guidance given to staff, they were not discussed widely in program documents and their adaptation to country circumstances was limited. Case study evidence suggests that political economy factors in program negotiations also played a significant role in motivating ambitious growth projections and there was limited attention in the program design stage to contingencies to respond to possible growth shortfalls.

Persistent growth optimism raises serious concerns because growth outcomes below program projections in the macroeconomic framework imply slower than intended progress in increasing incomes and strengthening the public balance sheet, undercut program ownership, and fuel rising adjustment fatigue and public opposition to reforms. While greater scrutiny of the realism of program projections as recommended by the 2018 ROC could help to reduce growth optimism, it seems even more relevant to consider whether IMF-supported programs can achieve stronger growth outcomes more in line with the program's macroeconomic framework by paying greater attention to growth-friendly policies in program design and implementation.

To shed light on this question, the evaluation examined to what extent different policy instruments were used to support program growth objectives during the evaluation period and how they could be applied to foster stronger growth outcomes.

Fiscal policies typically incorporated growth-friendly measures, but with mixed success. Tax mobilization improved in PRGT programs allowing higher capital

spending than otherwise, while GRA programs were able to make the tax structure more growth promoting in the post-program period. However, GRA programs relied heavily on spending cuts to achieve adjustment; efforts to protect low-income and vulnerable groups often fell short of their goals; health and education spending did not increase significantly in either PRGT or GRA programs; and a number of case studies raised concerns that growth benefits of higher public investment were limited by poor project selection and wasteful implementation. To help address such concerns, more attention is needed to building better public financial management and governance, building on recent initiatives in this area, and to strengthen monitoring and reporting of the social and distributional impact.

Structural conditionality included in programs generally played a positive role in promoting reforms and growth, but the potential growth benefits of structural reforms were not fully realized. Implementation of structural conditions (SCs) was positively associated with independently measured progress in structural reforms and helped to boost growth within and after the program, with a stronger growth impact for SCs with higher depth and growth orientation. However, the bulk of SCs was oriented to stabilization rather than promoting growth and the average depth and growth orientation of SCs was relatively low. While capacity development (CD) assistance was provided to support SC implementation and was broadly appreciated by country officials, it does not seem to have been delivered more to countries with weaker capacity nor been consistently effective in strengthening SC implementation. Moreover, implementation was significantly weaker for SCs in areas outside of Fund expertise and where collaboration with partners was sought. Country officials felt that Fund teams sometimes paid too little attention to growth-oriented reforms, relying too heavily on partner institutions, even for reforms crucial to program success. Overall, this evidence suggests the need to increase the focus on promoting deeper, more growth-supporting reforms, supported by steps to more closely integrate program and CD work and to strengthen collaboration with partners.

Use of the exchange rate as a policy tool to support growth and external adjustment during programs was quite limited. Exchange rate regime transition was infrequent and more often toward greater fixity, reflecting in part fear

of floating. Efforts were typically made to correct significant pre-program overvaluation, although more generally the impact of nominal exchange rate movements on the real effective exchange rate (REER) was partially muted by pass-through to prices. There was also a tendency towards a loss of competitiveness in PRGT programs relying on the exchange rate as a nominal anchor. Where it did occur, significant REER depreciation seems to have supported external adjustment and growth particularly in PRGT programs. This experience suggests that there could be greater scope to use the exchange rate as a policy tool in program design subject to the principle that the exchange rate regime choice is ultimately the authorities' decision and to members' obligations under Article IV to avoid manipulating exchange rates to prevent effective BOP adjustment or to gain an unfair competitive advantage. However, doing so successfully would require early attention to providing a supporting policy framework, including to securely anchor inflation expectations, to develop a deep foreign exchange market, and to alleviate supply-side impediments to exports.

In a number of cases, market debt operations were useful to restore debt sustainability and provide the basis for renewed market access, supporting a return to growth. However, debt operations were sometimes too little and too late, and thus had only mixed success in strengthening debt sustainability and improving the balance of payments position. Debt operations with principal haircuts and upfront fiscal adjustment were more successful than those with just debt reprofiling and lower coupons. This experience suggests that while respecting the neutrality principle, the IMF should consistently seek to ensure ambitious debt operations upfront to address debt sustainability concerns to qualify for access to Fund financing, based on careful application of the recently modified debt sustainability analysis (DSA) frameworks.

Three more general lessons are worth emphasizing.

First, the diverse experience in the case studies underlines that there is no simple recipe for delivering better growth outcomes in IMF-supported programs given the variety in country circumstances and preferences, the underlying causes and contexts of the BOP problems, and the potential scope for policy action. The need for careful tailoring for country conditions is underlined by the case study experience showing the importance of ensuring that the adjustment and growth strategy is fully owned by the government and broadly supported.

Second, the groundwork for a successful policy response to cushion the output consequences of an exogenous shock should ideally be laid well in advance through surveillance and CD work. The case studies repeatedly show that meaningful reforms to strengthen such growth resilience take many years to put in place and become effective, even with strong efforts to provide CD support.

Third, growth and reform strategies envisaged in program design should pay adequate attention to social and distributional consequences. While the focus in this evaluation has been largely on aggregate outcomes, fair distribution of the burden of adjustment and the rewards of recovery are of prime importance, both in their own right to meet national goals and to ensure continued public support for program implementation.

RECOMMENDATIONS

Notwithstanding the IMF's increasing attention to growth in program design and the generally positive role played by IMF-supported programs in promoting growth, the IMF should consider a number of actions to further enhance program countries' capacity to sustain activity while undertaking needed adjustment during the program period and to enhance growth prospects beyond the program.

Recommendation 1–Attention to growth implications of IMF-supported programs should become more thorough, systematic, realistic and sensitive to social and distributional consequences.

▶ Board papers supporting GRA as well as PRGT programs should clearly explain the program's growth implications both during the program and over the medium term. They should discuss how program design reflects the country's growth strategy and how growth considerations have been taken into account in the macroeconomic framework, ideally based on a well-calibrated country-specific model. Documents should provide more systematic coverage of the quality dimensions

of growth, including the distributional consequences of adjustment and reform policies such as how low-income and vulnerable groups are affected during the program period and how they would share in growth over time.

▶ In discussing the program's macroeconomic framework, particular attention should be paid to discussion of fiscal multiplier assumptions, especially where available country-specific modeling is limited, which should be further fine-tuned to country circumstances based on available evidence and informed judgement.

▶ Program design should pay more consistent attention to contingencies for growth shortfalls, based on scenario analysis, which should help fend off negative perceptions of the Fund's austerity bias.

▶ Efforts to pay greater attention to distributional aspects may require more granular approaches to conditionality and monitoring. Subject to data availability, strengthened monitoring of key social and distributional metrics would help to measure progress and signal emerging issues for program reviews.

▶ Revisions to the 2002 Guidelines on Conditionality and the 2014 Operational Guidance Note on Conditionality should be considered to give further guidance on the role of Fund-supported programs in fostering favorable growth outcomes. The update to the Guidance Note now under way can provide an opportunity to advance this work, while revisions to the Guidelines on Conditionality could be considered in the next Review of Program Design and Conditionality.

Recommendation 2—IMF-supported programs should pay greater attention to supporting deep, more growth-oriented structural reforms with more effective capacity development support and more effective collaboration with partners in areas outside the Fund's core mandate and expertise.

▶ The structural reform strategy should be geared to what is important and not what is most easy to agree

on or monitor or where the IMF has core expertise, subject to careful consideration of the country's implementation capacity and the program's goals. Structural conditionality should be parsimonious enough to avoid overtaxing country capacity but also more focused on correcting underlying distortions and removing structural impediments critical to achieving sustained and inclusive growth.

▶ The Fund should seek ways to strengthen collaboration with the World Bank and other relevant partners in design and implementation of structural reforms in shared and non-core areas. A useful step could be preparation of a Board paper reviewing experience with Bank-Fund collaboration in Fund-supported programs.

▶ The Fund should revisit how CD support is integrated with program design and implementation aimed at promoting deeper and more successful reform efforts in the program context.

Recommendation 3—The Fund should continue to invest in building a toolkit of models and monitors that can be applied as a basis for analysis of the adjustment-growth relationship and assessing growth-related developments in the program context.

▶ Functional departments could continue to take the lead in developing a suite of models suitable for analyzing the adjustment-growth relationship that are tractable and easily accessible for use by country desks to calibrate and apply in their country context. Particular attention should be paid to developing small-scale, easy-to-adapt macro/growth models for low-income countries where data are limited.

▶ Country teams should be encouraged to apply the models now being developed to achieve greater realism in program projections, to explore trade-offs between alternative policy mixes, and explain baseline projections and associated risks to authorities, which should help promote country ownership and mitigate the tendency towards growth optimism. Teams would determine case by case the models best suited to country circumstances

and needs. Area departments could also contribute by undertaking in-depth case studies on program successes and failures.

▶ The Fund should increase efforts to keep track of whether structural reforms were sustained after the program concludes, for example by investing more in the Research Department's new structural reform database.

▶ Further attention should be given to developing and deploying monitors to help support country desks' capacity to track developments in key distributional indicators and to gauge program impact on key social distributional dimensions of growth, in close collaboration with the World Bank and other agencies.

BUDGETARY IMPLICATIONS

It should be recognized that full implementation of many of these recommendations would have significant resource costs. Most significantly, the recommendations to take a fuller and more rigorous approach to analyzing and supporting program growth strategies with greater attention in program documents could add considerably to the time needed for program work (including for effective collaboration with the World Bank and other partners). At the same time, much of this work is already well underway or at least anticipated in the Fund's work program as part of the follow-up to the 2018 ROC. Taking on the additional commitments required would depend on a broader strategic decision to increase attention in the program context to ensure that IMF-supported programs not only deliver sufficient adjustment but also contribute in a more thorough way to sustained and inclusive growth.

1 INTRODUCTION

This evaluation assesses how well IMF-supported programs have helped to sustain economic growth while delivering necessary adjustment for external viability over the period 2008–19. While the evaluation does not assess the experience during the COVID-19 pandemic, its lessons have become even more relevant as countries seeking IMF support now face particularly strong headwinds to growth.

According to the Articles of Agreement, one of the fundamental purposes of the IMF is to make its resources temporarily available to members to help solve balance of payments (BOP) problems without resorting to "measures destructive of national or international prosperity." This purpose is echoed in the Fund's 2002 Guidelines on Conditionality, which stipulate that IMF-supported programs should be primarily directed to solving the member's BOP problems and achieving medium-term external viability while "fostering sustainable economic growth." As indicated in a 2013 Board paper on Jobs and Growth, IMF-supported programs should therefore "help maintain and strengthen growth as much as possible," while ensuring that programs meet their primary external goals (IMF, 2013a).

Supporting growth has been recognized as important not just for its own sake but also as a key ingredient to achieving economic and financial stability. The close linkage between debt sustainability and growth is well known. Weak growth outcomes render targeted balance sheet repairs or correction of stock imbalances more difficult to achieve and may undercut political support for adjustment and reform. However, achieving an appropriate and realistic combination of adjustment and growth has always been a challenging task. The relationship between adjustment and growth is likely to be nonlinear and uncertain, particularly if the economy is under financial stress or operating outside normal macro-economic conditions. It is also likely to differ depending on the horizon considered—that is, adjustment may depress growth in the short run but support growth in the medium and longer run.

In practice, the Fund's attention to growth in the context of financing arrangements has increased over time. The introduction of the Extended Fund Facility (EFF) back in the 1970s was intended to provide support for comprehensive programs over an extended period to correct payments imbalances because of structural impediments or slow growth. Greater emphasis on growth and poverty reduction for low-income countries (LICs) was reflected in the introduction of the Enhanced Structural Adjustment Facility (ESAF) in 1987 and the Poverty Reduction and Growth Trust (PRGT) in 2000. In the aftermath of the Global Financial Crisis (GFC), the Fund paid greater attention to growth in program design out of concern for the contractionary effects of adjustment on already weak economies. Greater emphasis on growth—both during the program period and afterwards—was also reflected in the 2014 update of the Operational Guidance Note on Conditionality and in revisions to the framework for debt sustainability assessment and the debt limits policy. These changes have fostered increasing attention to growth-friendly policies in program

design such as protection of public investment and social spending, growth-enhancing structural reforms, and debt operations to alleviate the extent of fiscal adjustment needed to achieve debt sustainability.

Despite increased attention, growth outcomes have tended to fall short of growth projections included in the program's macroeconomic framework. The 2018 Review of Program Design and Conditionality (ROC) found that while IMF-supported programs were often quite successful in solving the member's BOP problems, program growth projections were generally too optimistic. Growth short-falls implied less progress in reducing stock imbalances or ensuring debt sustainability than intended, while disappointed growth expectations could adversely affect domestic support for program implementation.

Seemingly lackluster growth outcomes under IMF-supported programs have often been criticized as indicative of an excessive austerity bias and continued lack of attention to growth consequences of IMF-supported programs.[1] In addition to raising questions about the benefits and costs of Fund support for the recipient countries, such concerns have resulted in a perceived stigma more broadly, potentially discouraging use of IMF financing and challenging the Fund's reputation.

Concerns about the growth impact of IMF-supported programs have fostered an extensive academic liter-ature. Empirical findings vary substantially depending on the sample periods and countries covered as well as empirical approaches used. In broad terms, the literature is inconclusive about the growth impact of IMF-supported programs, reflecting in part significant empirical challenges involved in identifying appropriate counterfactuals and isolating the impact of programs on growth from influences of other factors. Some recent academic studies have found positive evidence on the growth benefit of IMF-supported programs, but this remains an area of continuing research.

This evaluation aims to contribute to the continuing discussion over whether the Fund pays sufficient attention to growth concerns in the context of IMF-supported programs by assessing experience with adjustment and growth in program design and outcomes over the period 2008–19 and seeks to draw lessons for the Fund's lending framework. The evaluation recognizes that growth outcomes during IMF-supported programs should be assessed taking due account of the difficult circumstances faced by program countries and the substantial adjustment often needed to restore external viability. Thus, stabili-zation programs typically involve restraints on aggregate demand to close the gap between income and absorption. As a result, growth would normally be expected to fall short of historical trend performance during the program itself, although the additional external financing provided by the Fund and other sources could alleviate this impact. Programs can also help to raise growth performance after the program once adjustment is complete and the benefits of reforms supported by the program start to grow. Thus, the evaluation assesses whether programs helped countries to achieve higher growth than otherwise, distinguishing between the short run (i.e., during the program) and the medium run (i.e., after the program).

The evaluation builds on the findings of the recently completed 2018 ROC and other relevant studies, including earlier IEO evaluations. The findings of the evaluation are based on extensive empirical work using a large data set of program design and macroeconomic outcomes, a range of detailed country case studies, and six thematic background papers exploring growth-supporting strategies considered in IMF-supported programs (Box 1). The case studies cover 17 countries that accessed IMF support under the GRA and the PRGT to examine country-specific aspects of program design and outcomes to complement the findings of cross-country analysis at the aggregate level. The evalu-ation sample for the empirical work consists of 131 IMF financing arrangements with conditionality (including arrangements treated as precautionary) approved and completed between September 2008 and March 2020.[2] Data used in the analysis are taken mostly from the World Economic Outlook (WEO) database and the Monitoring of Fund Arrangements (MONA) database of the IMF. The latest actual data on economic outcomes are taken from the 2020 January vintage of the WEO database, while program projections and real time data are taken from various

---

[1] See Przeworski and Vreeland (2000), Dreher (2006), Van Waeyenberge and others (2010), and Ghosh (2019).

[2] See Annex for the full list and the composition of programs included in the evaluation sample as well as the data conventions used to determine program duration.

vintages depending on the specific window of time that the analysis is focusing on.

It is important to highlight up front that this evaluation does not systematically address three issues related to growth and adjustment in IMF-supported programs. First, it does not examine how the scale of access to IMF resources provided to program countries affected growth and adjustment outcomes. Clearly there are complicated short-term trade-offs involved. More program financing would ceteris paribus reduce the adjustment need and, hence, could in principle help to alleviate short-term growth pressures. However, such short-run benefits of larger program financing should be weighed against potential gains from stronger adjustment—such as positive confidence effects and lower indebtedness—as well as increased financial risks to the Fund itself and the need for the Fund to have access to a larger resource envelope.

Second, this evaluation does not systematically analyze the trade-offs between different types of policy adjustment (e.g., fiscal, monetary, and exchange rate policies) in program design, nor the macro-financial dimension of program design. The focus of the empirical work is on fiscal policy and exchange rate policy, although the country case studies do look at the composition of adjustment more broadly. This

choice in part reflects that fiscal and exchange rate policies are easier to compare across countries, while measurement and data issues are particularly tricky for monetary policies and macro-financial conditions. Moreover, inflation was less of an issue in most IMF-supported programs over the evaluation period due in large part to the downward global trend, while some program countries had no independent monetary policy. Indeed, the case studies prepared for the evaluation generally identified fiscal policy as the central adjustment tool, supported in some cases by exchange rate adjustment, with monetary policy playing a supporting role.

Third, the evaluation does not systematically address the impact of Fund-supported programs on the quality dimensions of growth (e.g., impact on low-income groups, on employment creation, and on the environment). Such outcomes are certainly relevant in affecting growth sustainability. Indeed, as documented in the case studies, issues related to ensuring inclusive growth and protecting the vulnerable received considerable attention in the design of virtually all programs being evaluated, and the evaluation discusses how far such policies were implemented in individual cases. However, even for countries individually there is very limited data available (and presented in IMF country reports) on actual outcomes related to the distribution of income, consumption, or employment

and cross-country comparisons are even more difficult. Thus, cross-country evidence is largely drawn from the existing literature focusing on the impact of fiscal adjustment policies.

The remainder of the report is organized as follows. Chapter 2 briefly reviews IMF policies related to program design and how they have evolved over time to give increasing attention to growth. Chapter 3 provides an overview of growth and adjustment outcomes of IMF-supported programs relative to initial conditions, program projections and growth benchmarks. Chapter 4 assesses the growth impact of IMF-supported programs empirically and discusses the role of IMF-supported programs in initiating sustained growth surges from a longer-term historical perspective. Chapter 5 assesses the growth and sustainability considerations incorporated in the macroeconomic frameworks through the lens of fiscal adjustment. Chapters 6 through 9 explore in greater depth a range of policy instruments for supporting growth in the program context, including growth-friendly fiscal policies, structural conditionality, exchange rate flexibility and debt operations. Chapter 10 concludes by summarizing the main findings of the report and provides recommendations aimed at strengthening growth outcomes in IMF-supported programs.

2 ATTENTION TO GROWTH IN IMF POLICIES ON LENDING AND PROGRAM DESIGN

The IMF's attention to growth in the program context has increased over time with different emphasis under different facilities. The EFF was created in 1974 to provide assistance to countries experiencing serious BOP imbalances because of structural impediments or slow growth and an inherently weak BOP position and provides financial support for comprehensive programs, including reforms to correct structural imbalances over an extended period. The ESAF was introduced in 1987 to provide concessional financing to support structural adjustment in LICs. The September 1999 Annual Meetings resulted in a clear mandate to more fully integrate the objectives of poverty reduction and growth into the Fund's operations for the poorest countries. This led to the creation of the PRGT in 2000.

Reflecting in part the increasing attention to growth, the Fund adopted a new set of guidelines on program conditionality in 2002, which replaced the 1979 Guidelines. The 1979 Guidelines focused on stabilization objectives while calling for the Fund to pay due regard to the domestic social and political objectives and economic priorities of the country. The 2002 Guidelines on Conditionality specified that Fund-supported programs should be primarily directed at solving the member's BOP problem without recourse to measures destructive of national or international prosperity and to achieve medium-term external viability while fostering sustainable economic growth.

Notwithstanding the increased attention to growth in the 2002 Guidelines on Conditionality, program design—particularly in non-concessional programs supported by the General Resources Account (GRA)—remained largely focused on achieving programs' primary external objectives. The 2005 ROC recognized that in the context of streamlining conditionality in GRA-supported programs (hereafter, GRA programs), growth can of course be a key aid to sustainability, but measures that would be aimed solely at increasing growth but would have no impact on external sustainability, while laudable, should not be made conditions of GRA programs. At the same time, it acknowledged a risk that streamlining efforts would result in insufficient attention to growth- and efficiency-related reforms in IMF-supported programs (IMF, 2005). Relatedly, the 2007 IEO evaluation on *Structural Conditionality in IMF-Supported Programs* found that there was extensive use of structural conditionality during the period of 1995–2004, but most structural conditions (SCs) had little structural depth, with only a weak link between compliance and subsequent reforms (IEO, 2007). The 2008 revision of the Operational Guidance Note on Conditionality (OGNC) reflected the Board's guidance in response to this evaluation to be more parsimonious in the use of structural conditionality by emphasizing criticality as well as requiring rigorous justification.

The growth impact of IMF-supported programs has received significantly more attention since the GFC. The 2009 Review of Recent Crisis Programs indicated that post-GFC programs accommodated larger deficits in order to cushion the short-run impact on

growth (IMF, 2009b). Similarly, the 2011 ROC found that fiscal adjustment was generally restrained in post-GFC programs largely out of concern for contractionary effects and that promoting growth and poverty reduction was a goal in an increasing number of GRA programs (IMF, 2012a; 2012c). It also found that key macroeconomic projections, such as growth, did not display an optimism bias in the aggregate. The 2011 ROC discussed growth in the broad macro-social context, encompassing the quality dimension of growth such as inclusiveness and income distribution. The 2015 Review of Crisis Programs noted that often tepid growth performance during 2008–13 reflected in part factors such as weak global conditions and balance sheet stress (IMF, 2015b).

Increased attention to growth was reflected in the 2013 Jobs and Growth Board paper and the 2014 Revisions to the OGNC. The Jobs and Growth paper indicated that "while ensuring that members achieve their primary goals of correcting their BOP problems and achieving external sustainability, Fund programs should help maintain and strengthen growth as much as possible." The revised OGNC sought to incorporate guidance on conditionality in relation to jobs and growth issues and specifically directed staff to accommodate to the extent possible the preferences and policy choices of country authorities, including on growth, labor market and distributional targets, subject to consistency with resolving BOP problems, macroeconomic stability and all other program goals (IMF, 2014b). It also stressed that staff should ensure that conditionality is well matched to tightly specified program goals, with due regard to the likely program effects on growth, employment and (at least where relevant for growth and stability) income distribution.

At the same time, other frameworks affecting program design were modified to take more account of the role of growth. For instance, the Fund's debt sustainability analysis (DSA) framework has been refined over time, with greater built-in scrutiny of the realism of growth projections. The 2009 and 2014 reforms of the debt limits policy (DLP) sought to ensure that IMF-supported programs strike a balance between debt sustainability and growth consider-ations—especially for LICs—by allowing greater flexibility in borrowing to create space for productive investment.

Reflecting the greater attention to growth outcomes, after a period in which structural conditionality was deliberately reduced, structural conditionality has gained greater promi-nence in recent years as prolonged slow growth has become an increasingly serious concern in many countries as the global macroeconomic environment remained persistently weak (IEO, 2018b). The IMF's increased attention to growth has also been reflected in the composition of program objectives. When program objectives recorded across the 17 categories in the MONA database are grouped into two broad categories of growth and adjustment, the share of growth objectives has increased after 2010 in GRA programs, while remaining relatively stable (at a higher level) in PRGT programs (Figure 1).[3] Still, the average share of growth objectives over the period 2008–19 was about 12 percentage points higher in PRGT programs (40 percent) than in GRA programs (28 percent).

Despite this heightened attention to growth, the 2018 ROC found evidence that growth outcomes tended to fall short of growth projections even as programs were generally quite successful in solving members' BOP problems in the period covered (2011–17). Its analysis suggested that growth optimism during this period was systematically related to an underestimation of the impact of adjustment on growth. It raised concern that, in the program context, growth optimism could trigger adjustment fatigue and undermine debt sustainability and ultimately program success. The 2018 ROC recommended increased scrutiny of the realism of program baselines and strengthened analysis of the growth impact of program policies. To follow up, the staff is now revising the 2014 OGNC, although this work has been delayed by the heavy work demands of responding to the COVID-19 pandemic.

---

[3] See Kim and others (2021) for technical details about the classification of program objectives.

FIGURE 1. COMPOSITION OF PROGRAM OBJECTIVES: 2008–19
(Percent share of total)
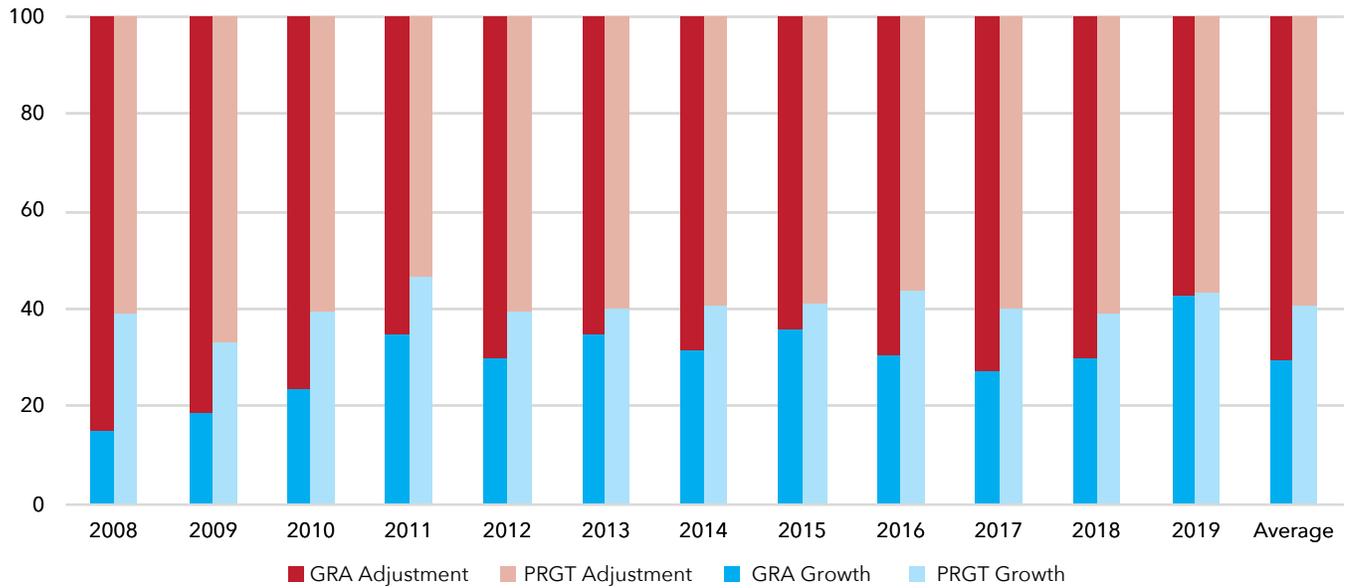
Sources: MONA database; IEO staff calculations.
Note: For each year, the left (right) bar is for GRA (PRGT) programs approved in that year.

GROWTH AND ADJUSTMENT OUTCOMES OF IMF-SUPPORTED PROGRAMS

This chapter provides an overview of growth and adjustment outcomes of IMF-supported programs over the evaluation period, looking at outcomes both during the program and afterwards, based on a range of empirical metrics.[4]

TIME PATTERN OF GROWTH AND ADJUSTMENT OUTCOMES

In GRA programs, growth outcomes typically exhibited a U-shaped trajectory with the trough in the first year of the program (T) followed by a rapid recovery in growth in the next year and more modest acceleration afterwards (Figure 2). Notable is the wide range of growth outcomes for year T as indicated by the interquartile range in shade. About 41 percent of GRA programs in the sample experienced real GDP contraction (i.e., negative growth) in the first year (T) of the program, two-thirds of which are accounted for by programs for countries in the context of the GFC and the euro area crisis ("crisis programs") for which the U-shaped pattern in growth trajectories is particularly pronounced (Figure 3).[5] Growth outcomes of other programs (which include some programs in response to home-grown BOP crises) were much steadier and show relatively small cross-country variation as indicated by the relatively narrow interquartile range.

Consistent with the 2018 ROC, growth outcomes consistently underperformed growth projections, indicating optimism bias embedded in initial program projections. This bias was particularly pronounced in the first year of GRA programs (median bias of 1.5 percentage points) but is also visible in later years (median bias averaging 1.1 percentage points in years T+3 to T+5).[6] The first-year optimism bias is particularly related to the experience of crisis programs.

In PRGT programs, there was a less marked pattern in the trajectory of growth outcomes. In the median program, an initial modest recovery at T was followed by a steady decline in growth until T+3 before leveling off (see Figure 3), in contrast to the steady recovery until T+3 shown in growth projections.

As in GRA programs, growth outcomes under PRGT programs generally fell short of projections but with a different pattern. The median outturn was close to projection in the first

---

[4] This chapter draws on Kim and others (2021) and country case studies prepared for the evaluation.

[5] Crisis programs refer to GRA programs arranged in response to a global or major regional crisis. Specifically, crisis programs include GRA programs approved during 2008–09 in response to the GFC (18 programs in total) and five Eurozone programs arranged in response to the euro area crisis (see Table A1 in Annex 1 for further details). Some other programs have also taken place in the context of BOP crises (e.g., Ukraine 2014 and 2015) but where the source has been internal imbalance rather than an exogenous shock. Their experience has typically followed a similar pattern of adverse growth outcomes.

[6] This metric does not distinguish between the later years of a multi-year program and the years after the program has been concluded. Thus, the empirical results are not always fully consistent with those discussed in the section "Growth and Adjustment Outcomes Relative to Projections," which are based on the data for program periods only.

FIGURE 2. GROWTH AND ADJUSTMENT TRAJECTORIES: GRA AND PRGT PROGRAMS
(Cross-country medians)
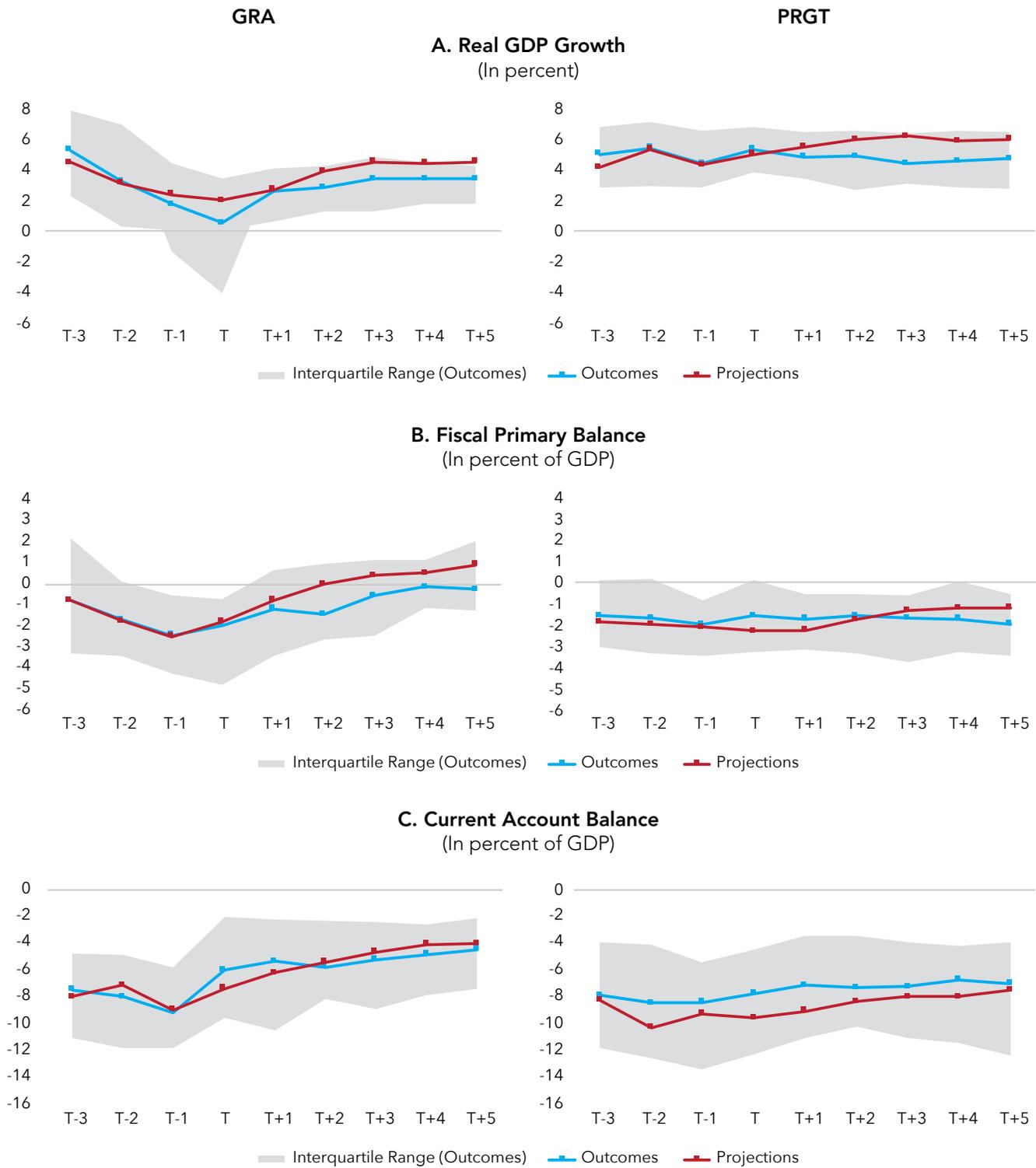
GRA

PRGT

A. Real GDP Growth
(In percent)

Interquartile Range (Outcomes) — Outcomes — Projections

B. Fiscal Primary Balance
(In percent of GDP)

Interquartile Range (Outcomes) — Outcomes — Projections

C. Current Account Balance
(In percent of GDP)

Interquartile Range (Outcomes) — Outcomes — Projections

Sources: WEO database; IEO staff calculations.
Note: All projections refer to initial program projections made at program approval (T). Outcomes and projections represent cross-country medians. Data availability is not uniform across periods mainly because post-program outcome data are not yet available for recently completed programs. Due to the presence of successor programs for some countries in the sample, there is overlap in the data presented over the period and, therefore, the results are not always fully consistent with those based on program periods only.

FIGURE 3. GROWTH AND ADJUSTMENT TRAJECTORIES: GRA CRISIS AND OTHER GRA PROGRAMS
(Cross-country medians)
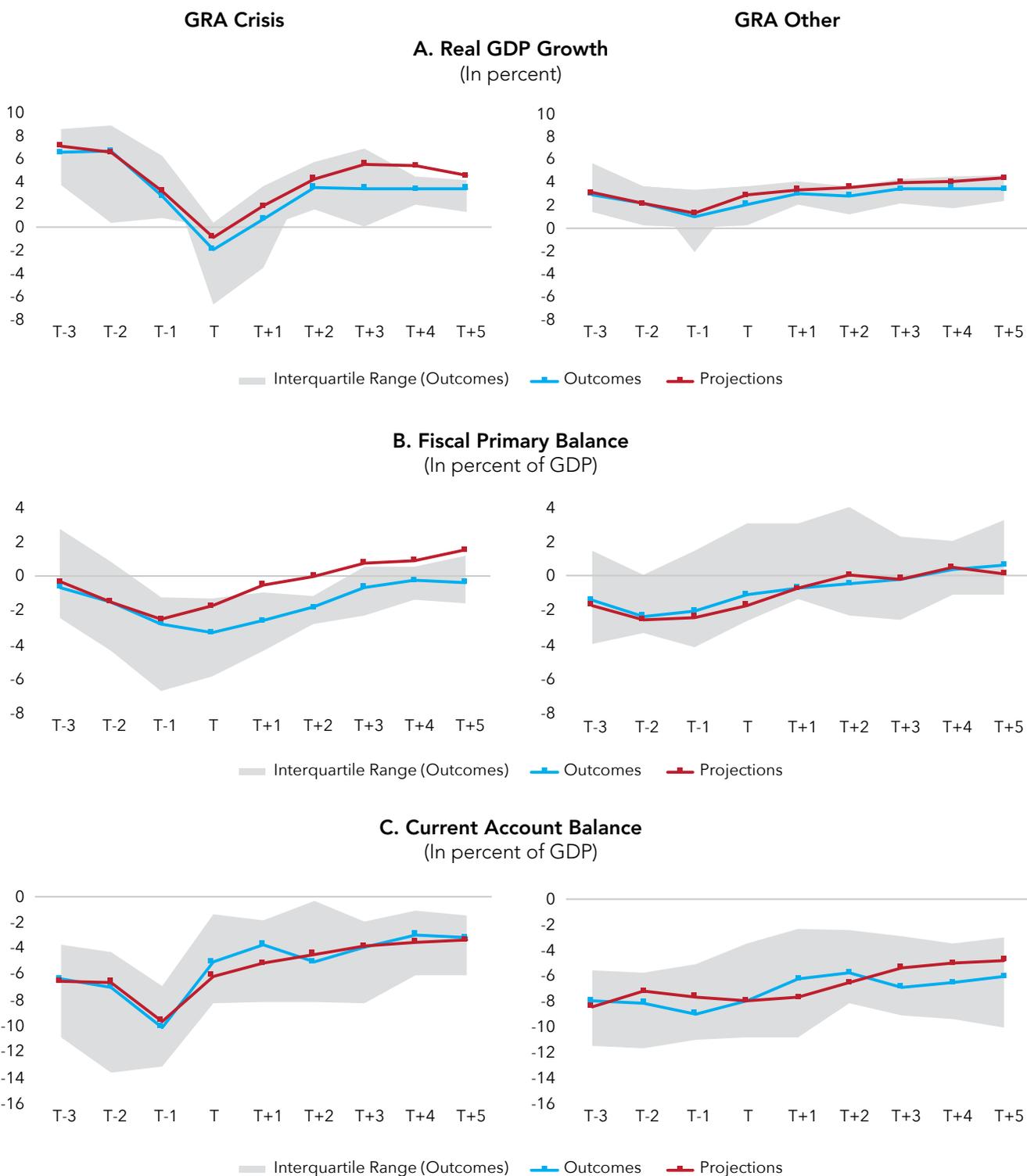
Sources: WEO database; IEO staff calculations.
Note: All projections refer to initial program projections made at program approval (T). Outcomes and projections represent cross-country medians. Data availability is not uniform across periods mainly because post-program outcome data are not yet available for recently completed programs. Due to the presence of successor programs for some countries in the sample, there is overlap in the data presented over the period and, therefore, the results are not always fully consistent with those based on program periods only.

program year, but fell increasingly short in subsequent years, with the gap reaching 1.8 percent by year T+3.

Like growth trajectories, both fiscal and current account (CA) balances exhibited a U-shaped pattern in GRA programs, but the trough was in year T–1 in the case of the CA balance (see Figure 2). Again, the U-shaped pattern observed for fiscal and CA balances was primarily driven by crisis programs. Such a pattern was far less visible in PRGT programs where the trajectories of fiscal and CA outcomes were quite stable over time. GRA programs showed on average smaller fiscal and CA deficits in outcomes and projections than PRGT programs.

In GRA programs, median fiscal outcomes were in line with the program in the first and second program years but subsequently underperformed projections by rising margins, particularly in crisis programs, while CA outcomes overperformed initially. In PRGT programs, the median gap between fiscal outcomes and projections was much narrower while CA outcomes consistently outperformed projections. It is notable, however, that CA outcomes varied widely across PRGT programs as indicated by the large interquartile range.

GROWTH AND ADJUSTMENT OUTCOMES RELATIVE TO PROJECTIONS

A closer look at growth outcomes relative to projections in programs provides further granularity on the growth optimism observed in IMF-supported programs. The analysis in this section is based on the data for program periods only. For consistent comparison between program outcomes and projections, the program sample is limited to 114 programs for which both projection and outcome data are available for one year or longer.[7]

Growth. Optimism bias in initial program projections averaged 1.3 percentage points in GRA programs—somewhat larger in crisis programs than in other GRA programs—and ½ percentage points in PRGT programs

FIGURE 4. GROWTH: PROGRAM PROJECTIONS AND OUTCOMES
(In percent; cross-country averages)
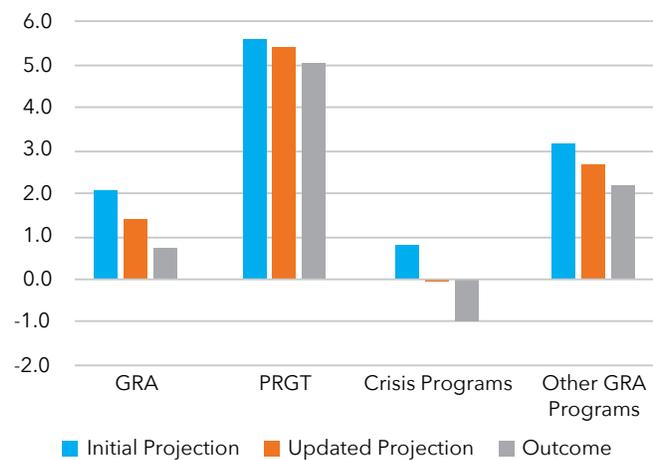
Sources: WEO database; IEO staff calculations.
Note: Updated projection refers to one-year-ahead projection (i.e., program projection made in year t-1 for growth outcome in year t).

(Figure 4).[8] In both GRA and PRGT programs, growth projections were revised downwards over the program period, generally more so in GRA programs and particularly in crisis programs than in PRGT programs. Thus, updated (one-year-ahead) program projections were typically more realistic than initial projections in both GRA and PRGT programs.

Growth shortfalls were widely dispersed across both GRA and PRGT programs (Figure 5). For GRA programs, growth fell short of projections by more than an annual average of ½ percentage points in 58 percent of cases; in 25 percent of cases the growth shortfall was greater than 2.2 percentage points. For PRGT programs, growth shortfalls were larger than ½ percentage points of GDP in 42 percent of cases and larger than 1.5 percentage points in 25 percent of cases.

External Adjustment. In GRA programs, on average actual CA adjustment exceeded modestly (by ½ percentage points) the programmed adjustment (Figure 6). Within GRA programs, both programmed and actual CA adjustments

---

[7]   Some programs in the evaluation sample went quickly off track; as a result, no observations are available for program outcomes under the conventions used to determine program duration for analytical purpose (see Annex 1). Comparison is based on annual averages over the program period. See Kim and others (2021) for further technical details.

[8]   Baseline projections at program approval assume full program implementation, implying that less than full program implementation could ex post lead to optimism bias. Similarly, ex post data revisions including GDP rebasing could be a source of optimism bias by itself and by affecting modeling errors in program design.

FIGURE 5. DISTRIBUTION OF GROWTH OUTCOMES RELATIVE TO INITIAL PROJECTIONS
(In percentage points)
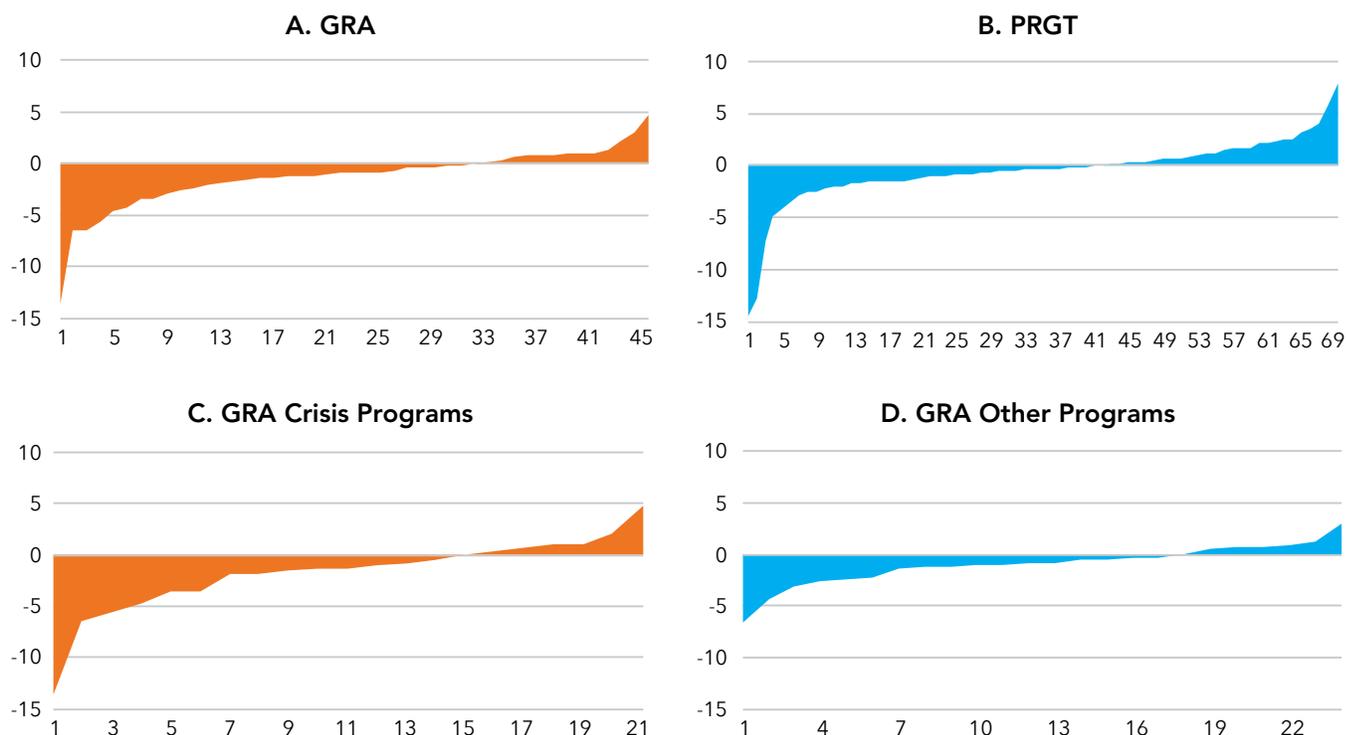
Sources: WEO database; IEO staff calculations.
Note: Data represent growth deviations (actual minus projection) in percentage points.

were stronger and relied more on import compressions in crisis programs than in other programs where improved exports played a greater role than import compression. Programmed and actual CA adjustments were both far smaller in PRGT programs but subject to large cross-program variation. Programmed CA adjustment was front loaded in GRA programs but back loaded in PRGT programs (Figure 7). In GRA programs, front loading was even more pronounced in program outcomes, largely driven by import compression. In sharp contrast to program projections, actual CA adjustment in PRGT programs was evenly phased, largely because projected increases in imports in the early phase of the program did not materialize.

Fiscal adjustment. In contrast to growth, actual improvement in the primary balance on average was slightly higher than program projections in GRA programs but fell short in PRGT programs (Figure 8). Within GRA programs, differences between fiscal outcomes and

projections were on average larger in crisis programs where fiscal outturns were significantly stronger than projected, particularly in expenditure adjustment. Programmed fiscal adjustment was on the order of 1.2 percent of GDP in GRA programs on an annual average basis but small in PRGT programs. Adjustment was dominated by expenditure adjustment in both projections and outcomes in GRA programs, while the composition was more even in PRGT programs. In terms of phasing, fiscal adjustment was front loaded in GRA programs (more so on the expenditure side and in program outcomes) while backloaded in PRGT programs with initial fiscal easing in the first year (T) of the program followed by fiscal tightening in the rest of the program period (Figure 9).

Debt. The combination of somewhat weaker growth outcomes (particularly in GRA cases) with more modest fiscal consolidation efforts (particularly in PRGT cases) has meant that public debt-to-GDP ratios have tended to rise rather than decline as programmed in both GRA and

FIGURE 6. EXTERNAL ADJUSTMENT: PROGRAM PROJECTIONS AND OUTCOMES
(In percent of GDP; annual average)

A. Change in CA Balance

Initial Projection ■ Updated Projection ■ Outcome

B. Change in Exports

Initial Projection ■ Updated Projection ■ Outcome

C. Change in Imports

Initial Projection ■ Updated Projection ■ Outcome
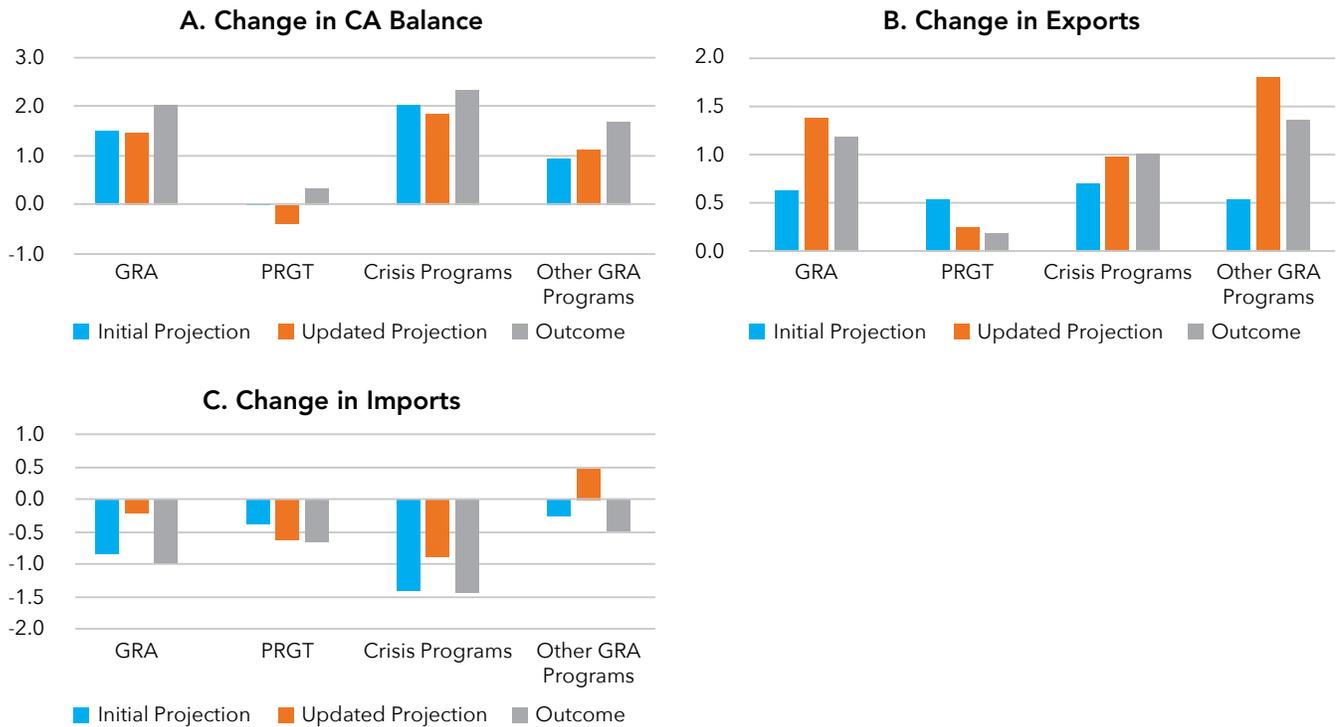
Sources: WEO database; IEO staff calculations.

FIGURE 7. PHASING OF EXTERNAL ADJUSTMENT: INITIAL PROJECTIONS AND OUTCOMES
(In percent of GDP)

A. GRA: Projections

First year (T) ■ Rest (annual average)

B. GRA: Outcomes

First year (T) ■ Rest (annual average)

C. PRGT: Projections

First year (T) ■ Rest (annual average)

D. PRGT: Outcomes

First year (T) ■ Rest (annual average)
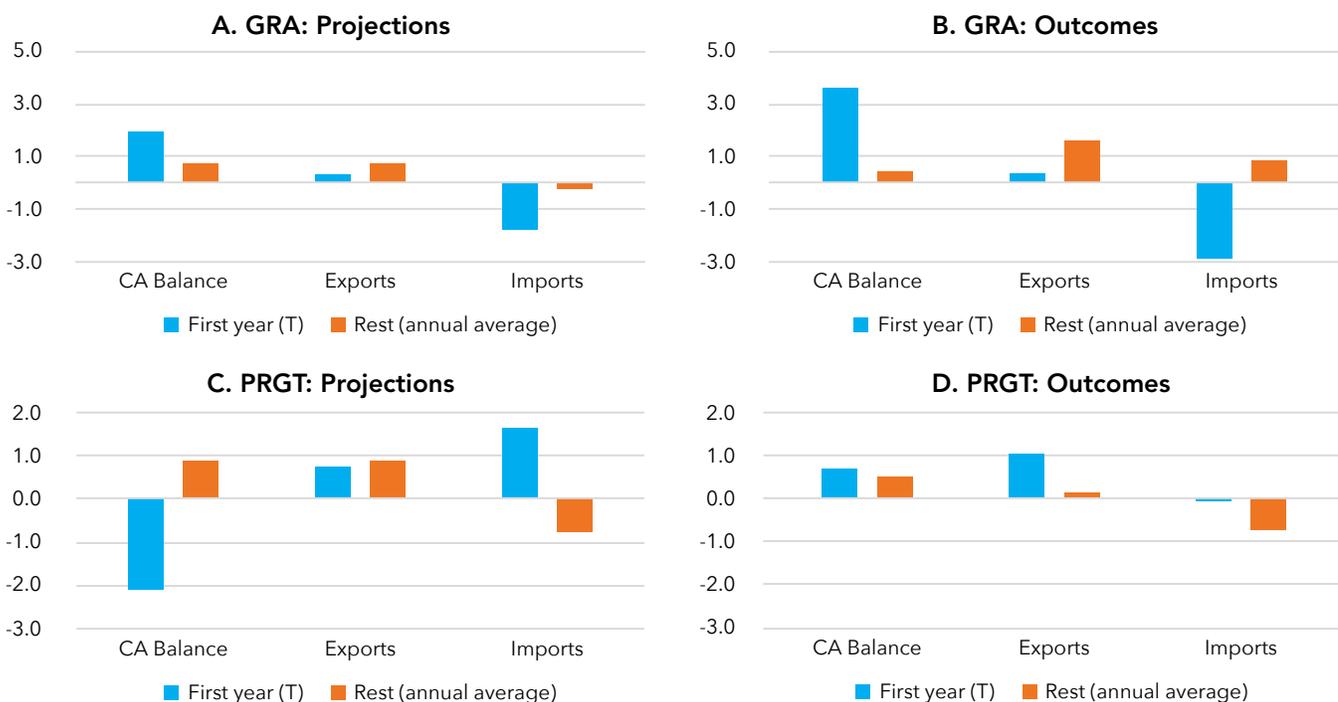
Sources: WEO database; IEO staff estimates.

FIGURE 8. FISCAL ADJUSTMENT: PROGRAM PROJECTIONS AND OUTCOMES
(In percent of GDP; annual average)

A. Change in Primary Balance

Initial Projection   Updated Projection   Outcome

B. Change in Revenue

Initial Projection   Updated Projection   Outcome

C. Change in Expenditure

Initial Projection   Updated Projection   Outcome
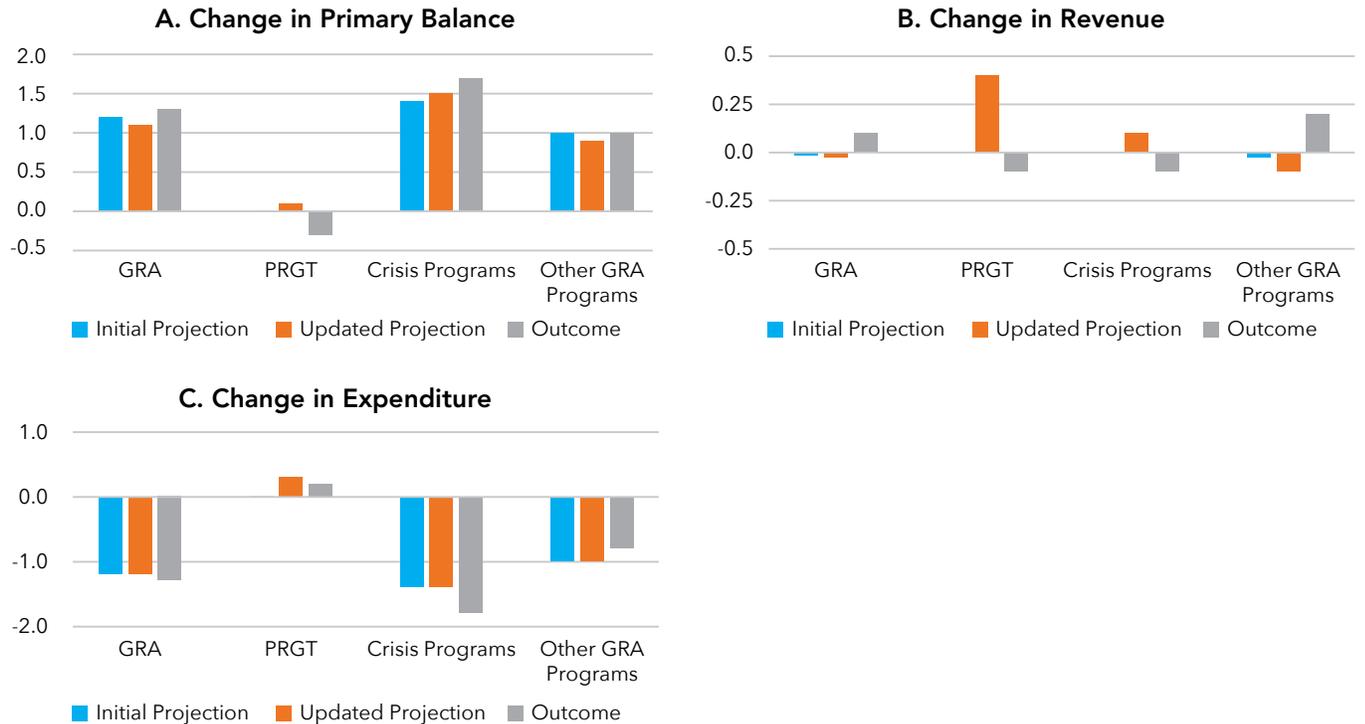
Sources: WEO database; IEO staff calculations.

FIGURE 9. PHASING OF FISCAL ADJUSTMENT: INITIAL PROJECTIONS AND OUTCOMES
(In percent of GDP)

A. GRA: Projections

First year (T)   Rest (annual average)

B. GRA: Outcomes

First year (T)   Rest (annual average)

C. PRGT: Projections

First year (T)   Rest (annual average)

D. PRGT: Outcomes

First year (T)   Rest (annual average)
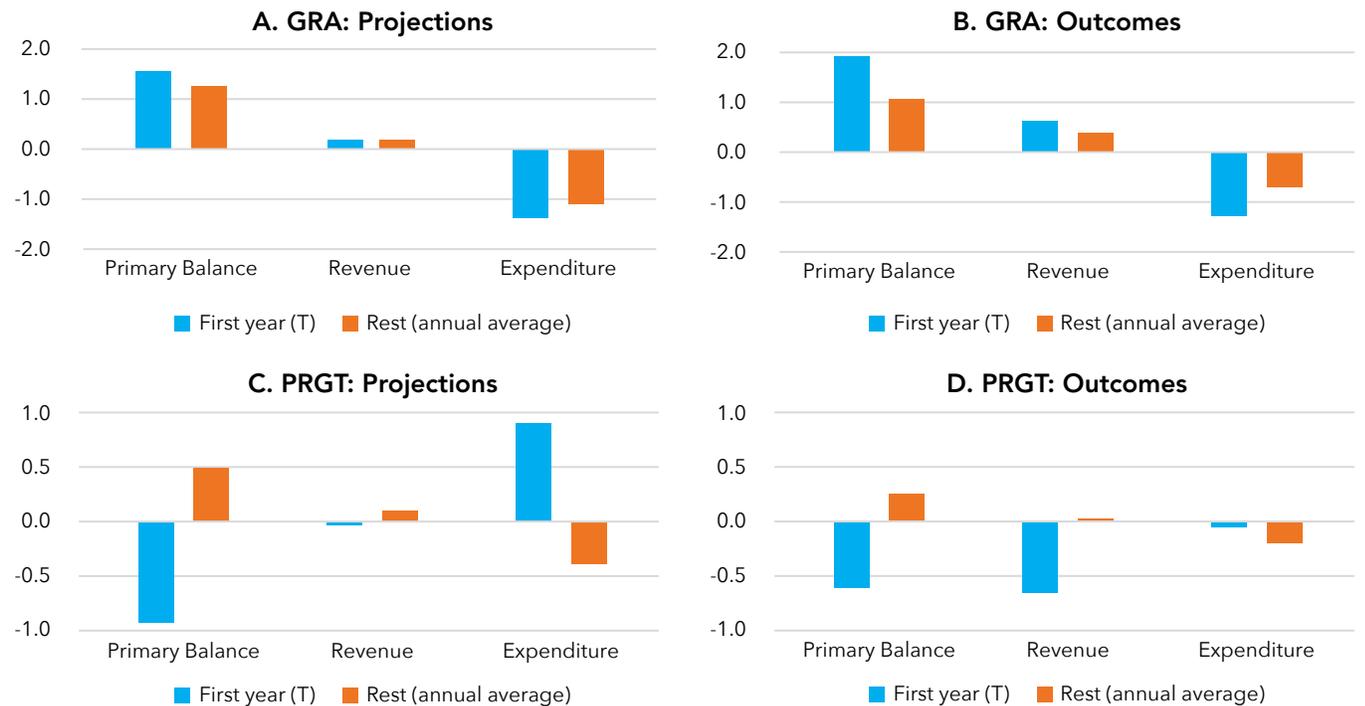
Sources: WEO database; IEO staff calculations.

PRGT programs. Moreover, there is large cross-country variation in debt projections and outcomes especially in GRA programs as indicated by large interquartile ranges (Figure 10).

FIGURE 10. DISTRIBUTION OF CHANGE IN DEBT
(In percent of GDP)
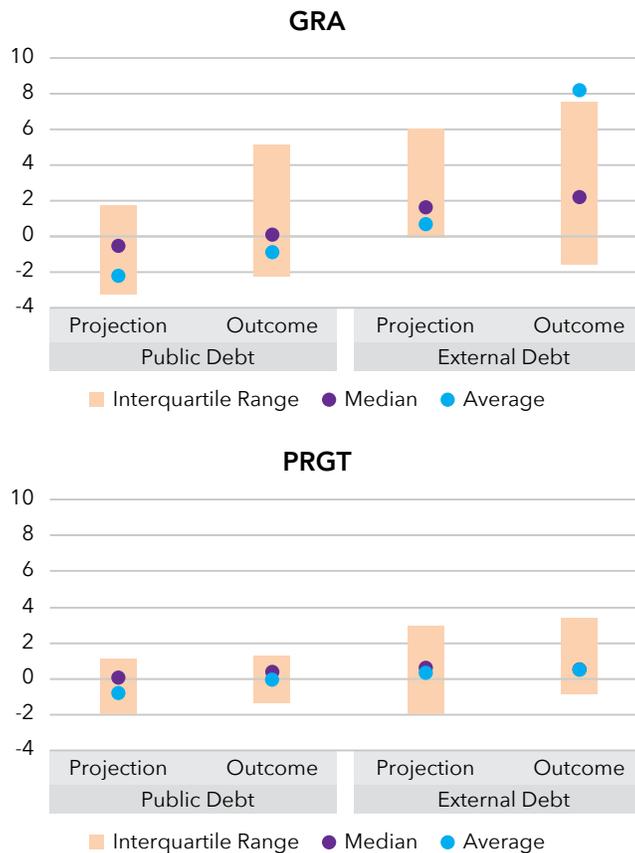
Sources: WEO database; IEO staff calculations.

GROWTH OUTCOMES RELATIVE TO BENCHMARK

This section seeks to compare growth outcomes during program periods to a historical growth benchmark for each country that seeks to reflect the impact of exogenous changes in the country's growth environment unrelated to the program but does not take into account a country's adjustment needs. This exercise recognizes that IMF-supported programs in the evaluation sample were

approved and completed at different times against different cyclical situations for the global economy; countries also experienced different terms of trade and external demand shocks depending on their economic structure and regional context. Moreover, program countries in the sample differ widely in historical growth trends. These differences pose an empirical challenge in making consistent cross-country comparison of growth outcomes over programs which span a few years at most.

The growth benchmark used in the evaluation is constructed to capture the variation in actual growth explained by external factors as well as country-specific historical trend growth. The benchmark is based on panel regressions linking growth to exogenous factors estimated for 174 countries over the period 1990–2019, including non-program as well as program periods. Growth deviations from the estimated benchmark are then calculated for program periods, and by construction should reflect primarily the influence of domestic factors such as domestic policy adjustments and supply shocks.[9]

Applying this approach, average growth deviations from the estimated benchmark ranged widely from –11.2 percent for Ukraine (2008 SBA) to 9.5 percent for Afghanistan (2011 PRGT) (Figure 11). While growth deviations are relatively evenly split between positive and negative values, the GRA sample is populated largely by negative deviations while the opposite is the case for the PRGT sample. As a result, the sample median diverges significantly between GRA (–1.5 percent) and PRGT programs (0.9 percent). Within GRA programs, sample medians also differ significantly between crisis programs (–3.5 percent) and other programs (–0.7 percent).

This exercise suggests that there were relatively few cases (12 percent of the full sample) in which the program growth outcome fell significantly below the country's historical norm, mostly associated with crisis programs with large adjustment needs, as well as some later programs with countries like Ukraine facing acute home-grown BOP problems. Overall, positive or negative growth deviations from the benchmark were statistically significantly different from zero (at 10 percent or higher) in 24 out of the 120 programs in total. The distribution of statistically

---

9   See Kim and others (2021) for detailed discussion of the estimation of growth benchmarks and related empirical findings. The benchmark is not intended to be a counterfactual (e.g., growth outcome that would have prevailed with no Fund engagement).

FIGURE 11. DISTRIBUTION OF GROWTH OUTCOMES RELATIVE TO BENCHMARK
(In percent)
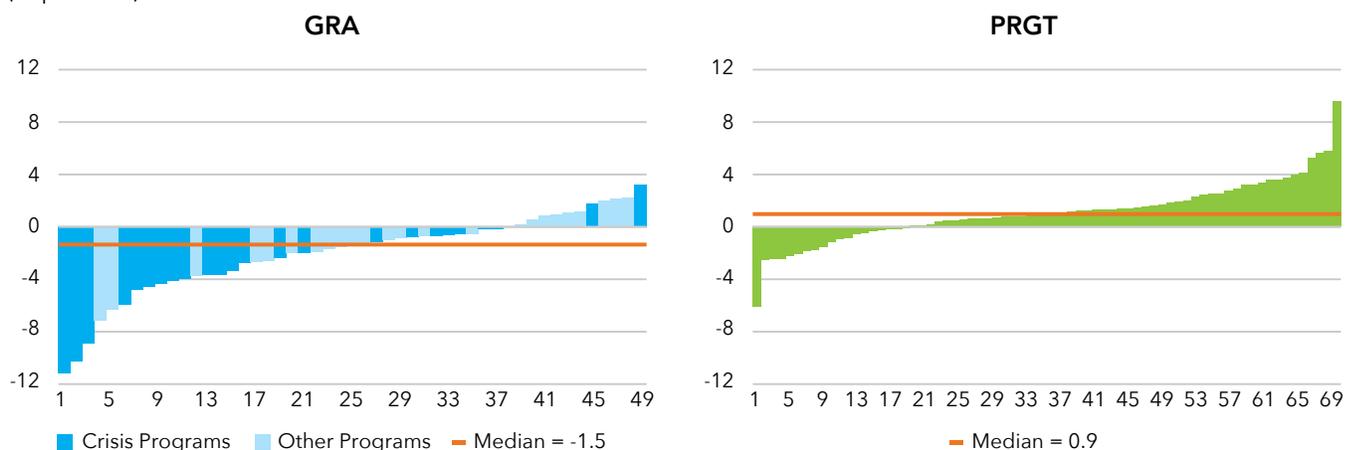
Source: Kim and others (2021).

significant deviations is quite uneven between the GRA and PRGT samples—positive deviations are entirely from the PRGT sample while almost all negative deviations are from the GRA sample (Table 1). Within the GRA sample, crisis programs dominate other programs in accounting for negative and significant growth deviations—11 out of 13 negative and significant deviations in the GRA sample are associated with crisis programs.[10]

Country case studies undertaken for the evaluation provide some further insights about the country-specific drivers of growth in cases of statistically significant growth deviations from benchmark. Positive growth deviations were often associated with favorable supply-side factors such as new mines coming on stream and good harvest in Ghana (2009

PRGF), a post-flood rebound in agriculture and buoyant tourism in Grenada (2014 ECF), and a demand stimulus from surge in public investment financed by capital inflows in Senegal (2015 PSI).

Negative and significant growth deviations found in case studies were mostly associated with crisis programs and driven by a range of negative demand and supply shocks as well as political factors. Latvia (2008 SBA) and Romania (2009 SBA) were both afflicted by an unwinding of an unsustainable economic boom and severe credit crunch in the aftermath of the GFC, as well as fiscal consolidation. Mongolia (2009 SBA) was hit hard by a slowdown in investment flows to the mineral and construction sectors and further by strong fiscal consolidation implemented in

TABLE 1. DISTRIBUTION OF GROWTH DEVIATIONS BY PROGRAM TYPE

PROGRAM TYPE	POSITIVE	NEGATIVE	TOTAL
GRA	12 (0)	37 (13)	49 (13)
Crisis	2 (0)	20 (11)	22 (11)
Other	10 (0)	17 (2)	27 (2)
PRGT	52 (10)	19 (1)	71 (11)
Total	**64 (10)**	**56 (14)**	**120 (24)**

Source: Kim and others (2021).
Note: Figures in parentheses are the number of programs for which the average growth deviation is statistically significantly different from zero at 10 percent or higher.

_____

[10] The remaining two GRA programs with negative and significant growth deviations are Ukraine (2014) and Suriname (2016), both of which went off track. Sierra Leone (2013) is the only PRGT program indicated with a negative and significant growth deviation.

the early phase of the program. For Ukraine (2008 SBA) which went quickly off track, restricted access to international capital markets after the GFC and a sharp fall in exports led to a sharp recession. Moreover, pre-existing domestic vulnerabilities, reflecting a stalled transition to a market-oriented economy and poor economic governance, weighed on growth. Ukraine (2014 SBA), which also went off track, suffered from political unrest and military conflict in the Eastern region which overwhelmed the government and depressed economic confidence.

SOURCES OF GROWTH OPTIMISM

It has long been recognized that the IMF's short-term growth forecasts are subject to optimism bias, particularly outside the advanced economies (Timmermann, 2007; IEO, 2014). The 2018 ROC confirmed optimism bias in growth projections in the program context and sought to identify its origin, following the approach used by Blanchard and Leigh (2013). More specifically, in the full ROC sample, short-run optimism bias was found to be slightly more than 1 percentage point, about one-quarter of which was accounted for by underestimation of the growth impact of fiscal and CA adjustments. Another one-quarter was explained by forecast errors of external conditions.

For this evaluation, we investigated growth optimism by undertaking a cross-country analysis of growth forecast errors drawing on the approach used in the 2018 ROC but for a slightly different purpose of assessing whether the contribution of macroeconomic modeling errors to growth optimism differed between program and non-program periods and between GRA and PRGT programs. Regression analysis was undertaken for a panel sample of 75 countries included in the evaluation sample over the period 2009–19.

Based on this approach, we found, similar to the analysis of the 2018 ROC, that growth forecast error regressions explain about one-quarter of total sample variation in growth forecast errors, leaving a large unexplained variation. The estimation results reported in Kim and

others (2021) suggest that large planned fiscal adjustments were associated on average with smaller optimism bias than average-sized fiscal adjustments. This finding, although at odds with evidence found by the 2018 ROC and other related studies,[11] may reflect that confidence effects associated with larger fiscal adjustments helped to offset income effects captured by standard multiplier analysis. Another finding is that macroeconomic modeling errors related to too low fiscal multiplier assumptions (relative to the estimated actual) were a statistically significant source of growth optimism in GRA programs other than crisis programs, although not in crisis or PRGT programs.[12]

While macroeconomic modeling errors have played a role, persistent growth optimism across programs seems to be substantially related to other factors. Drawing on case study evidence, an important role seems to have been played by political economy considerations in difficult program negotiations. Authorities have an incentive to provide the public with prospects of a robust payoff from adjustments and reforms to garner needed political support. Fund staff may also have an incentive to agree to unrealistic growth projections, which make it technically easier to close fiscal gaps and reach favorable conclusions about debt sustainability, while hoping to convince authorities to advance difficult adjustment and reforms. Several case studies illustrate how such factors played out in practice. In Latvia (2008), Fund staff anticipated a GDP decline of 6 percent to 8 percent in 2009 given the data already pointed to a sharp recession, but agreed to program a 5 percent decline as the authorities viewed such a forecast as overly pessimistic; the eventual outturn was a 14 percent contraction. In the case of Jamaica, staff noted in interviews with the IEO that medium-term growth forecasts were probably overoptimistic but cautioned that it would have been challenging to get domestic support for a program with even lower medium-term growth projections. In Jordan, Pakistan, and Tunisia, Fund staff underestimated the complexity of the political transition and the impact of intervening political, security-related and regional shocks. At the same time,

---

[11] For instance, Ismail, Perrelli, and Yang (2020) find for a large panel sample of 170 countries for the period of 2003–17 that large fiscal adjustments (one-half standard deviation or more above the sample average) are associated with higher growth optimism in surveillance and non-concessional program forecasts.

[12] The variance decomposition results reported in Kim and others (2021) show that modeling errors related to fiscal multipliers explain 30 percent of sample variation in growth forecast errors (after country and vintage fixed effects) in GRA programs other than crisis programs while little in crisis and PRGT programs.

country officials wanted to show hope to sustain political support for challenging reforms. The consequence was a disconnect between optimistic growth projections and actual outcomes.

ASSESSMENT

Overall, the evidence suggests that IMF-supported programs in the evaluation sample did not demonstrate consistently adverse growth outcomes after accounting for adjustment needs and the external environment. The most negative growth outcomes occurred in the first year of GRA-supported crisis programs. More generally, in the large majority of programs, growth outcomes did not fall significantly short of a historical growth benchmark that corrects for the influence of exogenous external factors and the difference in historical trend growth, but does not take into account adjustment needs. By this metric, PRGT programs appear more successful in achieving growth than GRA programs, which can be attributed in part to the fact that adjustment needs were generally smaller in PRGT programs than in GRA programs. By the same token, the adverse growth outcomes of crisis programs can be attributed to strong policy adjustments needed to address acute BOP pressure, financial fallout from the GFC and associated severe credit crunches and, in some cases, political unrest.

IMF-supported programs also delivered substantial adjustment in terms of the external current account balance and the fiscal primary balance. Fiscal adjustment was large (on the order of 1 percent of GDP per year) and significantly front loaded in GRA programs, especially in crisis programs where restoring investor confidence early on would likely be key to program success. In these programs, adjustment was achieved almost entirely through spending cuts. In PRGT programs, by contrast, fiscal adjustment was backloaded (i.e., initial easing followed by tightening). Such different magnitude and pattern of fiscal adjustment between GRA and PRGT programs are reflective of the differences in the nature and sources of BOP pressure, market access, adjustment need, debt sustainability concerns, and program objectives.

Program projections over the evaluation period were subject to considerable growth optimism bias, reinforcing the findings of the 2018 ROC. Over the programs covered in the evaluation, around one-half experienced an average growth shortfall (relative to initial program projections) during the program period of ½ percentage points or more, while one-fourth had a growth shortfall of over 1.5 percentage points. Optimism bias was on average larger in GRA programs (particularly in the first year of crisis programs) than in PRGT programs, although PRGT programs showed rising growth shortfalls in later years. Macroeconomic modeling errors, particularly those arising from unrealistic program assumptions on fiscal multipliers, seem to have been an important source of growth optimism in GRA programs other than crisis programs but less so in crisis or PRGT programs. In GRA crisis programs where fiscal adjustment was far stronger than in other programs (see Figure 6), the seemingly limited role of fiscal modeling errors in accounting for variation in growth forecast errors may be related to positive confidence effects that large and front-loaded fiscal adjustment can entail and help to offset in part negative income effects of fiscal adjustment.

Case study evidence suggests that while macroeconomic modeling errors played a role, political economy factors in difficult program negotiations also contributed to growth optimism in program design. Several case studies illustrate pressures on staff and the authorities to agree on excessively sanguine projections, hoping to sustain domestic support but underplaying the risks of subsequent growth disappointments and the challenges of program implementation.

Persistent growth optimism raises serious concerns because growth outcomes below program goals tend to contribute to adjustment fatigue, undercut program ownership, and fuel skepticism and rising opposition to reforms. While greater scrutiny of the realism of program projections as recommended by the 2018 ROC could help to reduce growth optimism, it seems even more relevant to consider whether IMF-supported programs can achieve growth outcomes more in line with growth projections by paying greater attention to growth-friendly policies in program design, as assessed in the later chapters of this report.

4

ASSESSING THE GROWTH IMPACT OF IMF-SUPPORTED PROGRAMS

This chapter presents more formal empirical approaches to assessing the growth impact of IMF-supported programs.[13] It builds on what is already a large academic literature examining the short-run growth impact of IMF-supported programs (i.e., during the program period) relative to a notional counterfactual of no Fund engagement. In broad terms, this literature is inconclusive, although recent studies have found more favorable evidence for the positive impact on growth of IMF-supported programs than earlier studies.[14] Many factors may explain the mixed findings of the literature, including different samples and empirical strategies adopted across studies. Moreover, there are significant econometric challenges involved, ranging from inherent difficulties to establish reliable counterfactuals to challenges in addressing sample selection bias, which could arise because programs are a deliberate choice of a country and the IMF and arranged only for countries with actual or prospective BOP needs.

Recognizing these challenges, the evaluation pursued a number of empirical approaches to assess the growth impact of IMF-supported programs during the program, in the post-program period, and in the longer run.

GROWTH IMPACT DURING THE PROGRAM

The growth impact of an IMF-supported program during the program period is estimated by using a recently developed statistical approach based on the propensity scoring method to correct for sample selection bias.[15] In this approach, the growth impact during the program period relative to a counterfactual of no Fund program engagement is identified as the average treatment effect (ATE) of IMF-supported programs. Thus, a positive ATE suggests that growth during the program period is higher than it would have been for a country in similar circumstances that did not undertake a Fund-supported program.[16] Estimation is undertaken for 152 countries in total (92 GRA and 60 PRGT countries) over the period 2008–19.

The estimated ATE is positive and highly significant (Figure 12). For completed programs, engagement in an IMF-supported program is estimated to raise annual growth, relative to a counterfactual of non-participation, by about 0.7 percentage points on average. The impact is higher for PRGT countries than GRA countries (1.1 percentage points versus 0.7

---

[13] This chapter draws on Kim and others (2021) and Atsebi and Wojnilower (2021).

[14] See Appendix V in Kim and others (2021) for a select literature review of empirical studies on the growth impact of IMF-supported programs.

[15] See Kim and others (2021) for greater technical detail.

[16] Note that a country may have a positive ATE even though output may contract during the program because the counterfactual in the absence of stabilization measures and financial support would have been an even deeper economic downturn.

FIGURE 12. SHORT-RUN GROWTH IMPACT OF IMF-SUPPORTED PROGRAMS
(In percentage points)
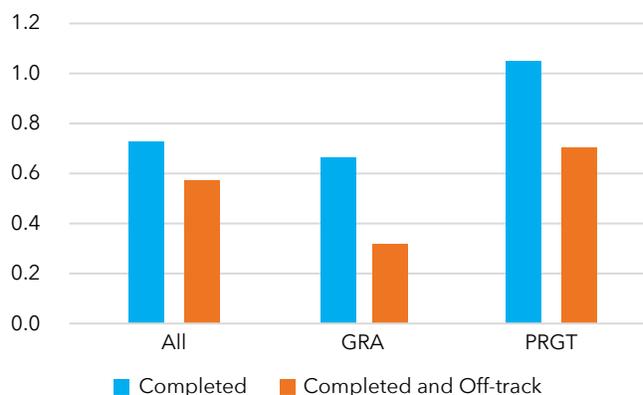
Source: Kim and others (2021).

FIGURE 13. SHORT-RUN GROWTH BENEFITS BY IMPLEMENTATION AND QUALITY OF SCs
(In percentage points)
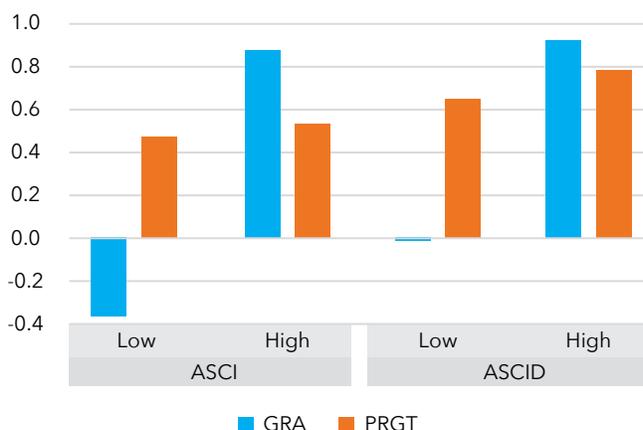
Source: Kim and others (2021).

percentage points). These results are broadly in line with the estimates reported in recent academic studies.[17] If the program sample is expanded to include both completed and off-track programs, estimated growth gains are smaller in both GRA and PRGT countries, implying that program implementation matters for short-run growth benefits of IMF-supported programs. Overall, these results provide strong support for short-run growth benefits of IMF-supported programs relative to a counterfactual of no Fund program engagement and underscore the importance of program implementation in realizing growth benefits.

Structural conditionality is an important aspect of IMF-supported programs aimed at ensuring progress with structural reforms. To assess whether structural conditionality has mattered for short-run growth gains, the analysis compares the ATE on short-run growth between programs with different records of implementation and quality of SCs.[18] The analysis uses the quantitative score index of SCs developed by the IEO (Kim and Lee, 2021), which is discussed in greater detail in Chapter 7 of this report.

The estimation results are all highly significant and provide strong support for the role played by SCs and the importance of their implementation and quality in determining the size of short-run growth gains from IMF-supported programs (Figure 13). Focusing on the average implementation score (denoted by ASCI), the difference in growth gains between high and low ASCI cases is particularly striking for GRA programs (1.2 percentage points), although negligible for PRGT programs. The results taking into account the depth of the measure as well as implementation (denoted by ASCID) paint broadly the same picture although in this case growth in PRGT countries is also seen to benefit from structural conditionality. These findings provide some support for the hypothesis that structural conditionality can generate short-term confidence effects which would be particularly important in GRA program

---

[17] Bas and Stone (2014) find for a sample of 104 countries over 1970–2008 that the average growth impact of IMF-supported programs is on the order of 1.4–3.5 percentage points and rises steadily with the cumulative number of years under programs, that is, for longer-term IMF engagement. Bal Gündüz (2016) reports an average growth impact of 0.4 percentage points for PRGT programs in 55 LICs over 1980–2010 and finds that the growth impact rises to 1.5–3.5 percentage points in LICs facing substantial imbalances or large exogenous shocks. For programs with 66 LICs over 1989–2008, Bird and Rowlands (2017) report a significant growth impact of 1.0–1.7 percentage points for concessional programs up to two years after approval but negative effects for non-concessional programs.

[18] Quality of SC refers to the depth and growth orientation of SC. Depth is the degree of structural change that a SC would bring about if implemented. Growth orientation describes whether the SC is intended primarily for enhancing growth and economic efficiency that would help the economy adapt better to changes in economic conditions. Depth and growth-orientation scores are constructed based on descriptive information on SCs in the MONA database. See IMF (2019b) and Kim and Lee (2021) for further discussion.

countries with market access. Strong implementation of high-depth SCs could signal firm commitments of program countries to a durable recovery and help boost investor confidence, ease external financing constraints and ultimately boost growth.[19]

POST-PROGRAM GROWTH IMPACT

The post-program growth impact of IMF-supported programs is assessed by examining the effect on post-program *potential* growth of stabilization and structural reforms implemented during the program. To this end, post-program growth regressions are estimated for the three-year average of post-program potential growth rates relative to the previously described IEO constructed benchmark that corrects for the influence of external factors and country fixed effects.[20] Our focus on potential growth rather than actual growth is motivated by two considerations. First, in conceptual terms, the potential growth rate is more appropriate to capture the slow-moving medium-term effects on growth of stabilization and reforms implemented during the program. Second, use of the potential growth rate, which corrects for cyclical variation associated with macroeconomic policies and shocks, should help to produce sharper estimates of the medium-run growth benefits of stability gains and reforms achieved in the program context.

Overall, the regression results provide good evidence for medium-run growth benefits of both stabilization and reform efforts during the program (Kim and others, 2021). As to the role of stabilization efforts, cumulative debt reduction—both public (DPDY<0) and external (ΔEDY<0)—during the program is found to affect post-program potential growth positively and statistically significantly (Figure 14). Specifically, a 10 percent of GDP reduction in public or external debt is found to help boost post-program potential growth by 0.2 percentage points. Growth-friendly fiscal adjustment with increased

public investment (DPUBINVY>0) and social spending (DSOCIALY>0) during the program also appears to have produced lasting growth benefits, of the order of 0.07 percentage points and 0.18 percentage points for a 1 percent of GDP increase in public investment and social spending, respectively. The growth effect of revenue mobilization (not shown) is found to be statistically insignificant, perhaps because positive benefits of higher revenue mobilization may have already been captured by public debt outcomes.

Market debt operations (measured by a simple dummy) in the program context are found to have a negative and statistically significant impact on post-program growth, particularly in the specification where public debt outcomes are also used as a control. The dummy variable is admittedly too coarse to adequately capture the diverse modalities and coverage of debt operations across countries. Moreover, a large part of growth benefits from debt operations may have already been captured by improved public debt outcomes included in the regressions. As such, the estimated negative impact of debt operations on post-program growth is likely to reflect the lingering effect on macro-financial conditions and investor attitudes associated with debt operations—such as increased borrowing cost, reduced market access, and lost investor confidence (particularly in cases where the country defaulted on sovereign debt).

As is the case for within-program growth impact, SCs implemented during the program have affected post-program potential growth positively and significantly, even though the considered post-program period of three years would likely be too short to observe the full impact of structural reforms (IMF, 2019e). The estimation results confirm that the quality of SCs, especially depth, matters significantly for post-program growth benefits. However, the mere implementation of SCs alone is not found to deliver significant growth benefits and could even harm post-program growth if too many low-quality SCs are

---

[19] In PRGT countries, Fund engagement itself is viewed as the most critical factor that helps boost donor confidence. This may explain generally smaller differences in growth gains between low and high SC scores (ASCI and ASCID) in PRGT programs than in GRA programs in Figure 13.

[20] The results do not distinguish between GRA and PRGT programs since the regression sample is relatively small with 54 programs at most due to the limited availability of post-program data since many programs were completed only recently or followed by successor programs in less than a year or two. The regression sample is further limited by programs that quickly went off track for which no valid annual data are available for program outcomes. See Kim and others (2021) for further detail.

implemented as indicated by the negative coefficient of SCI (see Figure 14).[21]

FIGURE 14. POST-PROGRAM POTENTIAL GROWTH REGRESSION COEFFICIENTS
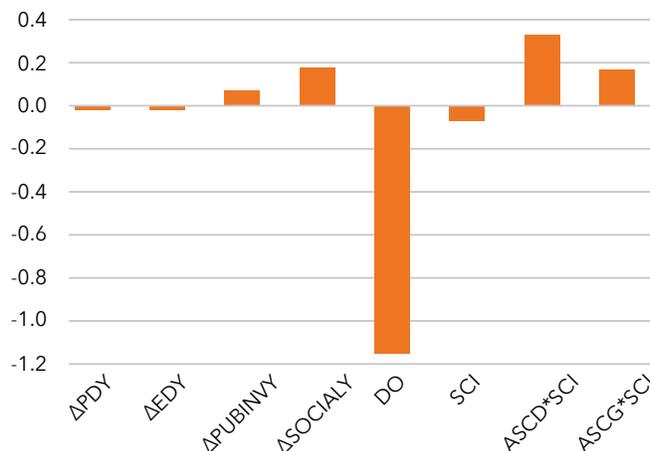
Source: Kim and others (2021).
Note: Based on the average across various specifications in Tables AVI.4 and AVI.5 in Kim and others (2021). DO denotes market debt operation. SCI, ASCD, and ASCG stand for the aggregate implementation score, the average depth score, and the average growth-orientation score of SCs, respectively.

SUSTAINED GROWTH SURGES AND IMF-SUPPORTED PROGRAMS

Taking a longer-term perspective than the rest of the evaluation, a background paper prepared for the evaluation (Atsebi and Wojnilower, 2021) assesses the role of IMF-supported programs in initiating sustained growth surges. In a seminal paper on growth surges, Hausmann, Pritchett, and Rodrik (2005) suggested that "accelerating the process of economic growth in a sustained manner is just about the most important policy issue in economics." Their study launched an expanding economic literature that aims to identify determinants of sustained growth accelerations, that is, growth surges. However, this literature has not considered the potential role of IMF-supported programs in helping countries to embark on sustained growth surges.

Anecdotal evidence suggests that IMF-supported programs, by supporting macroeconomic stabilization and structural

reforms, have played a role in initiating growth surges in a number of countries where deep-seated distortions and macroeconomic instability long hampered growth. Examples of countries that experienced sustained growth accelerations following IMF-supported programs include Thailand (1985), transition economies (e.g., Romania, 1997; Ukraine, 2003), the Baltic countries (Estonia and Lithuania, 1999), and Côte d'Ivoire (2010), among others (Figure 15).

For a more rigorous cross-country analysis and following the approach used by Hausmann, Pritchett, and Rodrik (2005), episodes of a sustained growth surge were identified as meeting two main criteria: (i) sustained rapid growth over at least eight years, and (ii) average growth significantly higher than in the preceding eight years. Applying these criteria and correcting for cyclical effects, 132 growth surges were identified in total in 117 countries during 1980–2017. More than half (56 percent) of these were associated with an IMF-supported program (Figure 16). It should be noted at the outset that growth surges are rare events—the unconditional probability of a growth surge starting in a given year is only 3.6 percent in the full sample. Interestingly, growth surges were significantly more frequent in the 2000s and in emerging market economies relative to previous decades and other income groups. Also noteworthy is that the share of growth surges associated with a program is significantly higher in the 2000s than in previous decades and in LICs compared to other income groups.

The empirical results in Atsebi and Wojnilower (2021) suggest that IMF-supported programs improve the chances of achieving a growth surge when policies are improved. Cross-country regressions based on a signal-extraction model show that significant improvements across the broad spectrum of relevant growth determinants (macroeconomic stability, structural reforms, investment and productivity) preceded nearly all growth surges. Moreover, such improvements were 10 percent to 20 percent more likely to trigger growth surges when associated with IMF-supported programs, an influence that was broadly similar between GRA and PRGT programs. While good luck (e.g., favorable external conditions) seems to have mattered significantly

---

[21] Since all three SC scores shown in Figure 14—i.e., SCI, ASCD*SCI, and ASCG*SCI where ASCD and ASCG stand for the average depth and growth-orientation scores respectively—are considered simultaneously in the post-program potential growth regressions, the coefficient of SCI captures the effect of SC implementation when all SCs are of the lowest depth and growth orientation. Thus, the negative coefficient of SCI suggests that implementing too many low-quality SCs could harm growth. See Kim and others (2021) for further discussion.

FIGURE 15. SELECTED COUNTRIES WITH GROWTH SURGES AND IMF-SUPPORTED PROGRAMS

Real GDP per Capita (LHS)   Average Annual Growth Rate (past 8 years; RHS)   Growth Surge   IMF-Supported Program (including 2 years after)
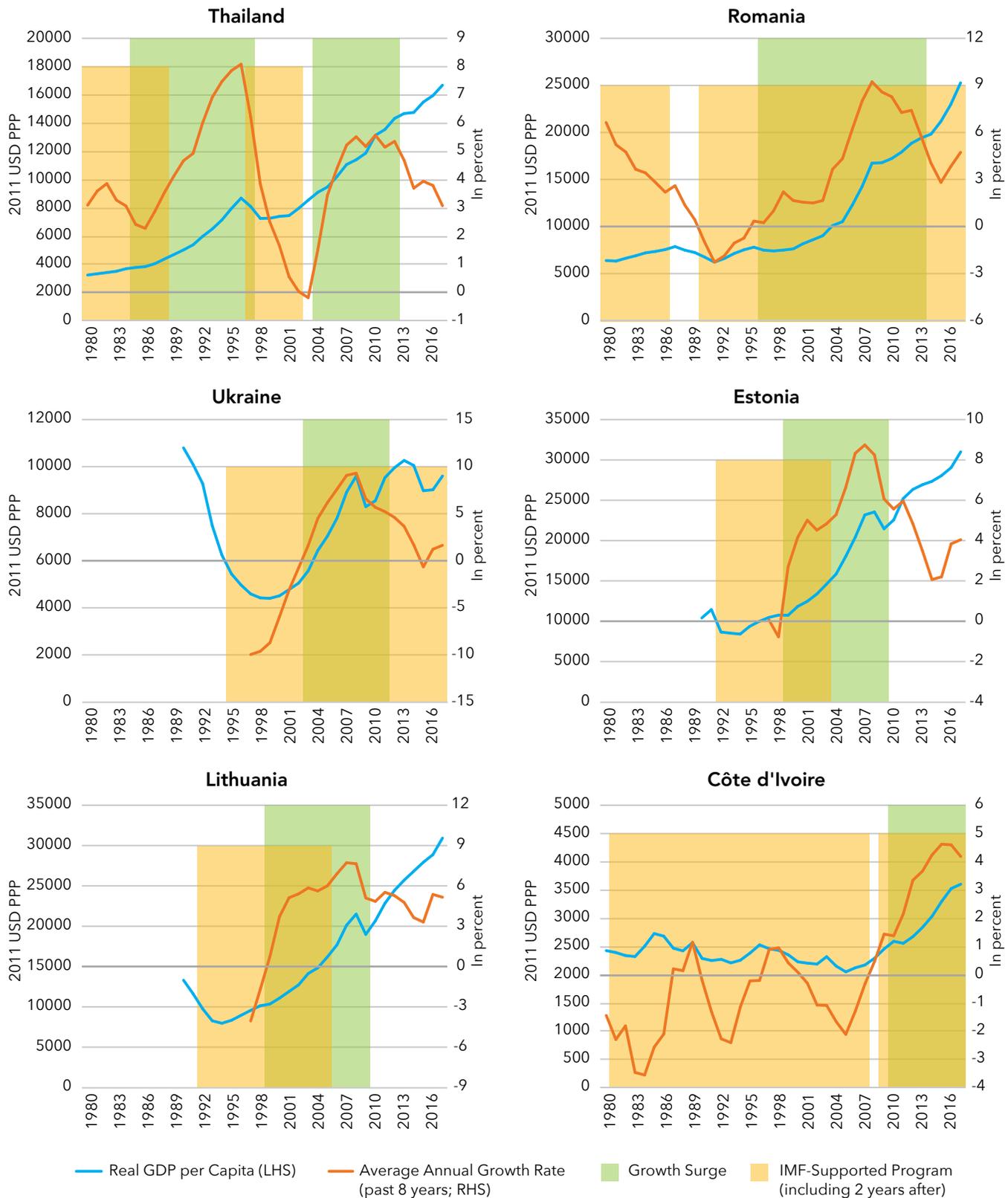
Source: Atsebi and Wojnilower (2021).

FIGURE 16. GROWTH SURGES AND THEIR ASSOCIATION WITH IMF-SUPPORTED PROGRAMS
(Number of growth surges)

A. By Decade

B. By Region

C. By Levels of Development

Legend: ■ Total ■ IMF-supported Program ■ GRA-supported Program ■ PRGT-supported Program
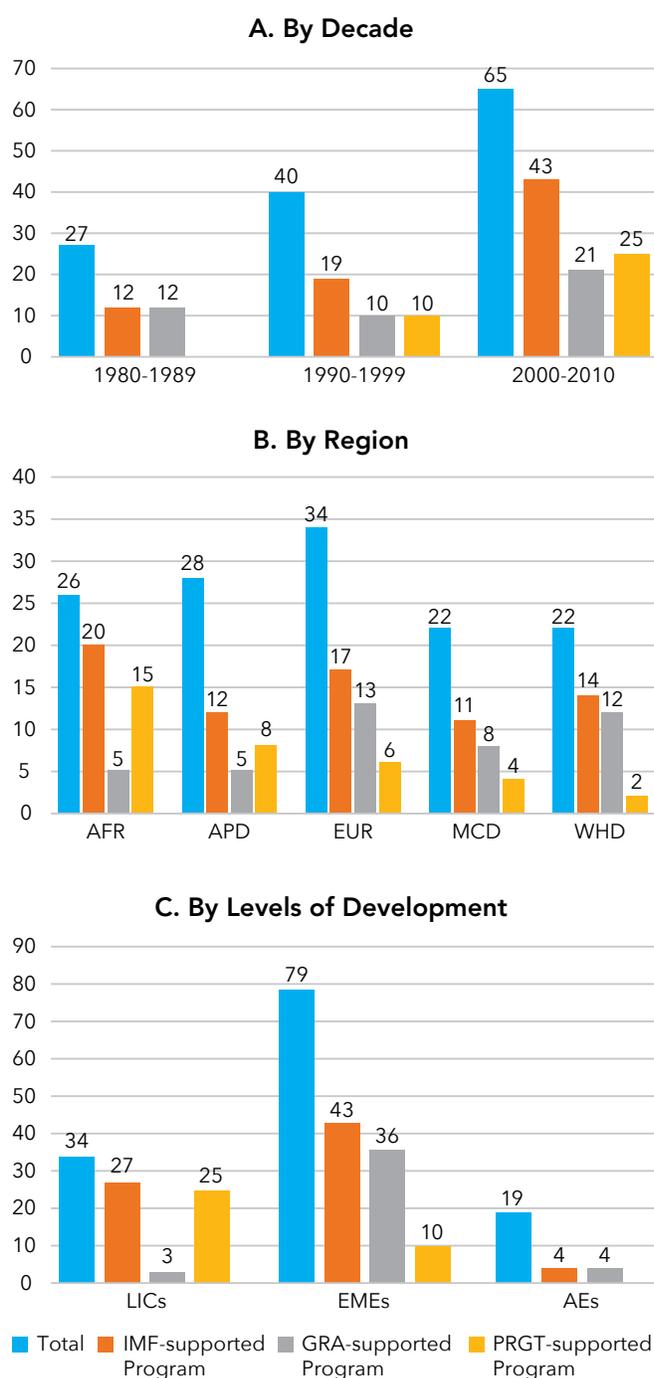
Source: Atsebi and Wojnilower (2021).
Note: AFR, APD, EUR, MCD, and WHD are abbreviations for IMF departments covering Africa, Asia and the Pacific, Europe, the Middle East and Central Asia, and the Western Hemisphere (i.e., the Americas and Caribbean), respectively. LICs, EMEs, and AEs stand for low-income countries, emerging market economies, and advanced economies, respectively.

for countries to embark on a sustained growth path, even in such cases the presence of an IMF-supported program seems to have enabled the country to take better advantage of favorable external conditions for sustained growth. Furthermore, IMF-supported programs have become increasingly effective at initiating growth surges over time, which suggests that the IMF's increasing attention to growth seems to have borne some fruit.

ASSESSMENT

Empirical evidence presented in this chapter suggests that IMF-supported programs have played a positive role in promoting growth in program countries in the short, medium, and longer runs.

The estimated growth impact during the program period, relative to a notional counterfactual of no Fund engagement, is significant, and found to be somewhat higher for PRGT programs than GRA programs. The estimated impact is found to be lower than the ranges reported in some recent academic studies but nonetheless seem quite impressive in light of the fact that such growth benefits were achieved against the background of a global growth and productivity slowdown in the aftermath of the GFC (IMF, 2015a; Adler and others, 2017). Another relevant finding is that SCs in the program (more precisely the implementation status and quality of SCs) have mattered significantly for the growth impact of IMF-supported programs, even in the short run. Given that structural reforms tend to affect growth with considerable time lag of 5–7 years (IMF, 2019e), the short-run growth impact of SCs may be attributed to positive confidence effects in the program context that implementation of high quality SCs can generate.

Evidence is also found for the medium-run growth impact of IMF-supported programs. Both stabilization policies and structural conditionality implemented during the program are estimated to have affected post-program potential growth positively and significantly. Debt reductions and increases in public investment and social spending during the program have generated lasting growth benefits beyond the program horizon. SCs also have mattered significantly for post-program potential growth, with stronger growth impact for higher quality SCs.

As to the longer-run growth impact, there is evidence that IMF-supported programs increase the likelihood of a sustained growth surge, by helping countries to implement policies conducive to macroeconomic stability and growth-enhancing structural reforms, as well as by strengthening the effectiveness of such policies in generating a growth surge. The results also suggest that IMF-supported programs have become increasingly effective at initiating growth surges over time.

5

GROWTH ASPECTS OF MACROECONOMIC PROGRAM DESIGN

The IMF's attention to growth has been increasingly reflected in program objectives (see Figure 1). This chapter provides cross-country evidence on how attention to growth was incorporated in program design and examines country experience in this regard.[22] Given that the primary objective of programs is to correct BOP problems and restore external viability, attention to growth needs to be assessed in conjunction with sustainability considerations.

ATTENTION TO GROWTH AND SUSTAINABILITY IN PROGRAM DESIGN

Growth and sustainability considerations in program design are assessed through the lens of fiscal policy because it is at the center of macroeconomic stabilization in programs and relates to policy instruments under control of country authorities. To be more specific, our assessment focuses on how fiscal policy was calibrated to address growth and debt sustainability concerns and how it reacted to interim macroeconomic developments. This assessment is undertaken both for initial program design and for program adaptation. Initial program design at program approval provides the most comprehensive snapshot of the macroeconomic framework given the financing envelope of the program. However, focusing only on initial program design would miss a crucial aspect of program design—the flexible adaptation of programs in response to interim macroeconomic outcomes in the context of periodic reviews of program implementation (Mussa and Savastano, 1999).

Initial Program Design

For initial program design, the assessment is guided by the analytical framework developed by Bohn (1998, 2008) and used in subsequent research on debt sustainability and fiscal space (e.g., Mendoza and Ostry, 2008). The analytical framework identifies a positive response of the primary balance to lagged debt as a sufficient (but weak) condition for debt sustainability. Allowing for a nonlinear fiscal reaction function incorporates the notion of fiscal fatigue (Ostry and others, 2010).

This analytical framework is used to estimate a fiscal reaction function which relates programmed fiscal policy to the lagged debt ratio and the output gap.[23] The estimated reaction coefficients provide evidence to assess how growth and sustainability considerations were reflected in the design of fiscal policy.[24] Growth considerations would suggest on average a positive response of the primary balance to the output gap so that programmed fiscal policy would be counter-cyclical in nature.

---

[22] This chapter draws on Kim and others (2021) and country case studies prepared for the evaluation.

[23] In the empirical analysis, the output gap is constructed as a percentage deviation of (projected) real GDP from the log-linear trend in real GDP over the 10-year period prior to program approval.

[24] See Kim and others (2021) for the estimation results and related discussions.

Scatter plots based on data for initial projections in the programs covered in the evaluation seem to support the hypothesis that fiscal policy has been set to reflect both stabilization and growth objectives by responding to lagged public debt and the output gap, more so in GRA programs than in PRGT programs (Figure 17). The nonlinear trend lines in the scatter plots (left panels) suggest a positive response of the primary balance to lagged debt over the interval of debt ratio between 50 percent and 140 percent of GDP, after which positive fiscal reaction is weakened. This feature in fiscal outcomes is attributed to fiscal fatigue in Ghosh and others (2013). In the context of program design, it could reflect some feasibility constraints in fiscal adjustment or growth considerations beyond what is captured by the reaction to the output gap.

Formal multivariate regression results broadly confirm the bivariate relationships in the data. The estimated linear reaction coefficients are of the expected sign in most cases but statistically significant only in GRA programs, suggesting that both growth and sustainability considerations were well reflected in programmed fiscal policy in GRA programs but less clearly so in PRGT programs (Figure 18). The results for nonlinear fiscal reactions (not reported here) provide stronger support for growth and sustainability consideration embedded in initial program design in both GRA and PRGT programs, although the results continue to be weaker in the latter, as the fiscal reaction to the output gap continues to be small and not significant in PRGT programs.[25]

FIGURE 17. FISCAL REACTIONS IN INITIAL PROGRAM DESIGN
(In percent of GDP)
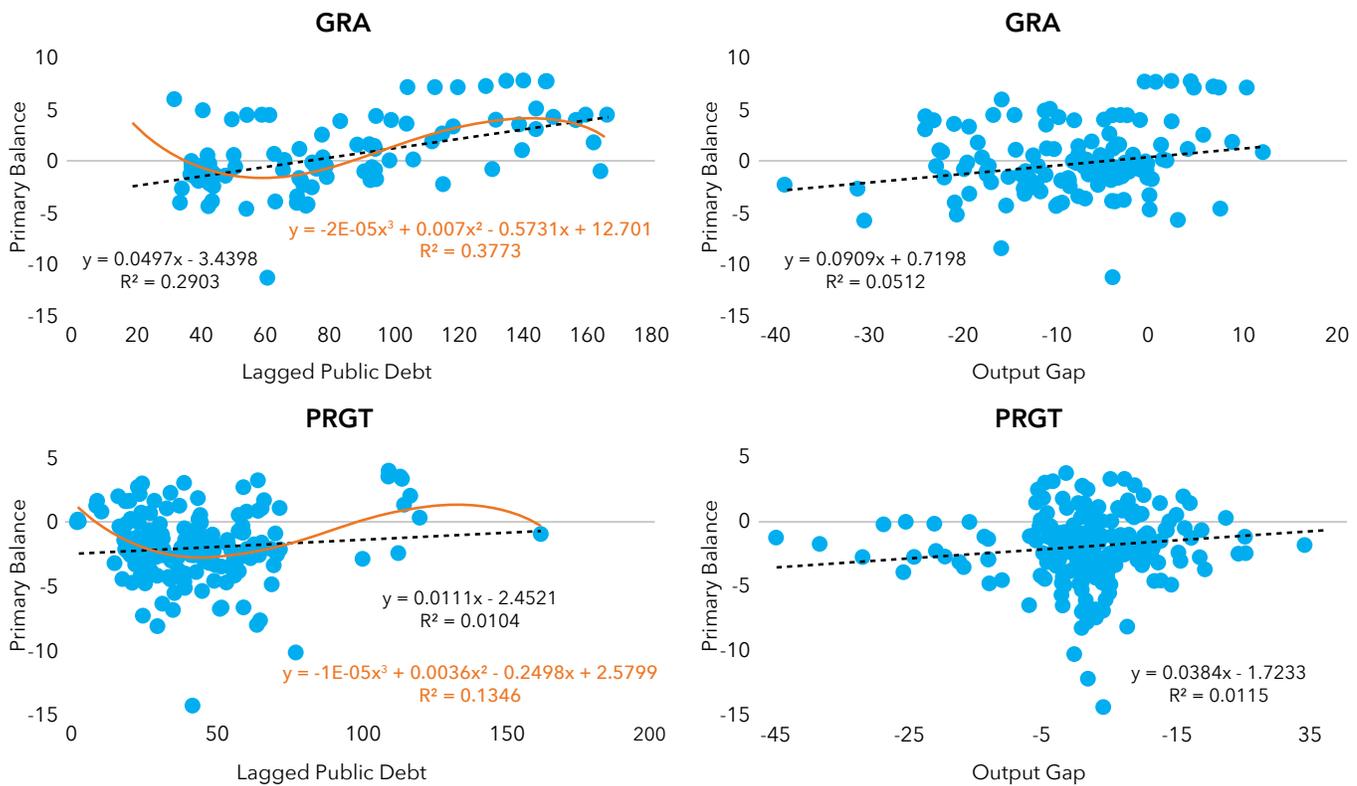
Source: Kim and others (2021).

---

[25] See Kim and others (2021) for a fuller discussion on nonlinear fiscal reactions and their implications.

FIGURE 18. LINEAR FISCAL REACTION COEFFICIENTS: INITIAL PROGRAM DESIGN

Source: Kim and others (2021).
Note: Asterisks denote statistical significance. *p<0.1, ** p<0.05, *** p<0.01.

FIGURE 19. FISCAL REACTION COEFFICIENTS: PROGRAM ADAPTATION
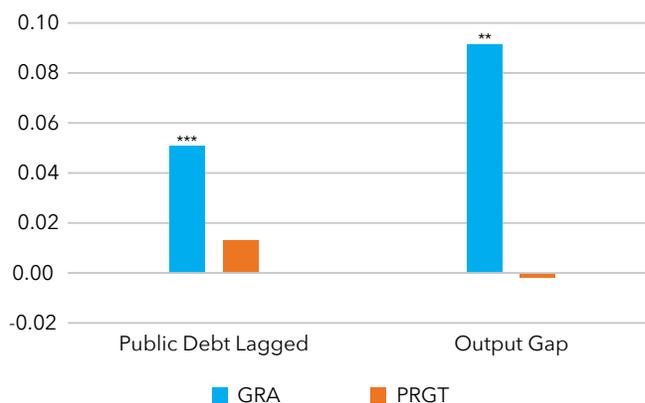
Source: Kim and others (2021).
Note: Asterisks denote statistical significance. * p<0.1, ** p<0.05, *** p<0.01.

Program Adaptation

Drawing on the analytical framework used by IEO (2014), the evaluation examined how growth and sustainability considerations were reflected in modifications to the policy framework in program reviews. In particular, we examined how updated (one-period-ahead) projections for fiscal adjustment in the next period were modified from previous (two-period-ahead) projections in response to interim growth and fiscal adjustment forecast errors (defined as actual minus projection) observed in the current period. Growth considerations generally call for positive fiscal reaction to growth forecast errors (GFEs), so that growth shortfalls in the current year lead to less ambitious fiscal adjustment and ceteris paribus higher growth than initially programmed in the next year. Sustainability considerations would require negative reaction to fiscal adjustment forecast errors (FAFEs) so that adjustment shortfalls in the current year lead to stronger fiscal adjustment than initially programmed in the next year.

Regression results again show that both growth and sustainability considerations were at play in calibrating fiscal reactions in program adaptation. In the estimation, the dependent variable is the modification in fiscal projection, measured as the difference between one- and two-year-ahead

fiscal projections.[26] The estimated fiscal reaction coefficients are of the expected sign and statistically significant in most cases (Figure 19, bars on the left side). The reaction coefficients of FAFEs are more negative for GRA programs than for PRGT programs, suggesting that sustainability considerations have generally been stronger in GRA programs than in PRGT programs. By contrast, growth considerations seem to have played a relatively stronger role in adapting PRGT programs than in GRA programs when assessed by the reaction coefficients of GFEs.

The disaggregated results between positive and negative forecast errors provide a more detailed account of how growth and sustainability considerations were addressed in program adaptation (Figure 19, bars on the right side). For instance, about 90 percent of adjustment shortfalls (FAFE < 0) were programmed to be recovered in the next period in GRA programs but less than half in PRGT programs. While fiscal reaction to growth shortfalls (GFE < 0) was stronger in GRA programs than in PRGT programs, reaction to upside adjustment surprises (FAFE > 0) gave more weight to growth considerations in PRGT programs than in GRA programs. Specifically, about 87 percent of upside adjustment surprises (FAFE > 0) were to be reversed in the next period in PRGT programs while only about 40 percent were to be reversed in GRA programs.

---

[26]   To approximate the information available to country authorities and Fund staff at the time of projection, actual data used to construct growth and fiscal adjustment forecast errors are real time data as recorded in the WEO vintages matched with program years, rather than the latest data from the most recent WEO vintage. See Kim and others (2021) for further technical details.

LESSONS FROM COUNTRY EXPERIENCE

The country case studies in this evaluation broadly support the cross-country empirical evidence of efforts to reflect both growth and sustainability considerations in program design and adaptation. In virtually all cases, staff and country officials discussed the appropriate degree of upfront adjustment in the specific country circumstances, typically seeking to moderate the pace of fiscal adjustment to avoid too adverse an impact on activity while still providing a credible path to achieving stabilization goals. In some cases, particularly earlier in the evaluation period (e.g., Grenada 2010, Jamaica 2010, and Pakistan 2008), authorities felt that staff were too inflexible in insisting on front-loading of fiscal adjustment that was hard to sustain. In other cases, officials and staff felt that front-loaded adjustment was essential to restore confidence in the face of wide imbalances, and some officials (e.g., in Honduras 2014, Latvia 2008, and Romania 2009) in fact preferred to follow a tougher adjustment path than proposed by staff feeling this would pave the way for more vigorous recoveries by supporting recoveries of business sentiment and regaining market access.

It is striking, however, that program documents presented to the IMF Board seldom provided much analysis of the potential short-term trade-offs between adjustment and growth or how it could be affected by different policy mixes. Moreover, program documents discussed the specific actions for authorities to take in response to the materialization of growth-related risks in less than 20 percent of the programs studied, including Egypt, Ghana, Grenada, Jordan, Mongolia, Tunisia, and Ukraine. Even in the few programs with explicit program contingencies, measures considered in the risk assessment matrix were focused largely on policy implementation risk while indicated responses to adverse shocks to growth were often limited to avoiding pro-cyclical fiscal tightening rather than easing the adjustment effort in the face of an adverse growth shock.

The relative paucity of discussion of program contingencies in initial program design notwithstanding, program reviews generally adapted policy settings to respond to adverse growth outcomes where applicable in most case studies, consistent with the broader empirical evidence. Specifically, program reviews adapted fiscal targets due

to weaker growth outcomes or fiscal overruns in many programs, including Bangladesh, Cameroon, Grenada, Latvia (where Fund staff sought less fiscal consolidation than the authorities), Mongolia, Pakistan (not in the first review but in subsequent reviews), Romania, Senegal, and Ukraine. Program reviews were also combined and/or extended to provide the authorities more time to take corrective policy actions after policy slippages in a range of programs, including Bangladesh, Ghana, Honduras, Jordan, Malawi, Mongolia, Pakistan, Tunisia, and Ukraine.

Most authorities and staff viewed flexibility in program adaptation as contributing to program success. In Tunisia, staff viewed adaptations to the program during quarterly reviews as the key instrument for adjusting the program framework and taking remedial actions. In Honduras, authorities felt that the Fund's more flexible attitude during the 2014 program contributed to its success, while the lack of flexibility during the 2010 program contributed to its going off track irretrievably. In some other programs, authorities were less supportive of the way program adaptation was handled. In Ghana, authorities thought the Fund should have been more flexible in completing reviews. In Ukraine, staff and authorities, with the benefit of hindsight, agreed that greater emphasis should have been placed on contingency planning.

However, the case studies also illustrate a clear risk related to more extended adjustment paths and program adaptation in response to weaker than expected growth outcomes, particularly when the envisaged stabilization of public debt is not achieved. Cameroon and Senegal provide examples of countries in which fiscal adjustment (over a sequence of programs in case of Senegal) fell short as deviations from adjustment targets have been accommodated in the presence of weak growth and external borrowing has been used to support public investment, while the private sector response has remained lackluster. As a result, these countries have faced increasing medium-term debt sustainability risks.

ASSESSMENT

When assessed through the lens of fiscal policy, both growth and sustainability considerations were well incorporated in initial program design in GRA programs but less clearly so in PRGT programs. In GRA programs, fiscal primary balance targets reacted positively to the lagged debt

ratio (satisfying a weak debt sustainability condition), as well as to the output gap (implying counter-cyclical fiscal policy). In addition, fiscal reaction to the lagged debt ratio appears to be nonlinear, providing some further support for growth considerations embodied in initial program design. In contrast, such systematic fiscal reaction was less clear if not absent in PRGT programs.

In adapting programs for interim outcomes, updated fiscal projections balanced growth and sustainability considerations not only in GRA programs but also in PRGT programs. Fiscal adjustment targets tended to be revised downwards in response to interim growth short-falls and upwards in response to adjustment shortfalls.

Sustainability considerations were generally stronger in GRA programs than in PRGT programs. Country case studies broadly confirm these cross-country findings on program adaptation. Explicit discussions on program contingencies were relatively infrequent among programs studied, but nonetheless program reviews generally eased fiscal targets due to weaker growth outcomes or fiscal overruns in many programs. Program reviews were also combined and/or extended to provide time for corrective policy action and compliance with program conditions. Moreover, most authorities and staff viewed flexibility in program adaptation as contributing to program success.

6

FISCAL ADJUSTMENT, FISCAL REFORMS, AND GROWTH

This chapter examines growth-related aspects of fiscal policy in IMF-supported programs, looking first at how the growth impact of fiscal adjustment has been calibrated and then at how different growth-related aspects of fiscal policy have been reflected in program design and implementation.[27] It provides evidence on the extent to which IMF-supported programs have succeeded in mobilizing revenue, which can then be used to support growth-friendly public spending while delivering fiscal adjustment, and in modifying the composition of public spending in favor of public investment and social support. It then reviews the use of structural conditionality to support growth-friendly fiscal reforms.

FISCAL MULTIPLIERS IN PROGRAM DESIGN AND OUTCOMES

In practice, analysis of the growth implications of fiscal adjustment in the program context has generally been ad hoc and quite limited. There is no official IMF-wide guidance, although country desks have typically used a so-called bucket approach suggested by Batini and others (2014), which identified the key structural characteristics of an economy that influence the size of fiscal multipliers and proposed a plausible range of fiscal multipliers for different advanced, emerging, and LIC groups. However, while the paper suggested that staff fine-tune fiscal multiplier assumptions for the economy's cyclical position and monetary policy stance, this additional step has seldom been taken. Moreover, despite increased awareness among Fund staff of the relevance of fiscal multipliers in program design, they are not discussed widely in program documents (IMF, 2019c). Insufficient attention to fiscal multipliers implies that staff risks underestimating the adverse growth impact of fiscal adjustment in circumstances that could increase the size of the multiplier (for example, when the economy is already weak and room for offsetting monetary or exchange rate policy is limited) or overestimate the growth impact where there are significant offsetting benefits from higher confidence.

This section presents evidence on how program design incorporated fiscal multipliers reflecting the short-run relationship between fiscal policy and output, how multiplier assumptions were modified after program approval, and whether they differed from actual multipliers. For this purpose, we used formal regression analysis to estimate the short-run relationship between adjustment and growth embodied in program projections (using both initial and updated projection data) and program outcomes. The estimated coefficients of fiscal adjustment in the regressions are taken as fiscal multipliers.[28] It should be noted at the outset that the purpose of the regression analysis for program projections is not to

---

[27] This chapter draws on Gupta (2021), Kim and others (2021), and country case studies prepared for the evaluation.

[28] Strictly speaking, the estimated coefficient of fiscal adjustment should be considered as a proxy for the fiscal multiplier because in the regression analysis fiscal adjustment is measured by the change in the fiscal primary balance and not the change in the *structural* primary balance. Kim and others (2021) discuss in greater detail the growth regressions.

establish a causal relationship between adjustment and growth, but to assess the underlying assumptions—particularly program assumptions on fiscal multipliers—used to formulate program projections.

The regressions found a statistically significant short-run trade-off between fiscal adjustment and growth embodied in the initial program design in both GRA and PRGT programs (Table 2, Panel A). The estimated fiscal multipliers are on the order of 0.35–0.5 in GRA programs, a range which is consistent with the buckets proposed by Batini and others (2014), the broader literature on fiscal multipliers, and program assumptions in several country case studies as discussed below. In addition, GRA programs on average set revenue multipliers to be smaller than expenditure multipliers, which is again in line with existing evidence in the literature. According to the regressions, PRGT programs on average used smaller fiscal multipliers (on the order of 0.17–0.22) than assumed in GRA programs, which is consistent with the findings in the literature that fiscal multipliers are generally lower in emerging and low-income countries than in advanced economies.[29] As in GRA programs, revenue multipliers were assumed to be lower than expenditure multipliers although the difference between them was small and less marked than in GRA programs.

The same regression analysis was undertaken for updated program projections to assess how the macroeconomic framework evolved over the program period to incorporate new information from interim macroeconomic outcomes. Updated fiscal multipliers for GRA programs were on the order of 0.24–0.35, in general lower than those assumed in initial program design (Table 2, Panel B). By contrast, updated fiscal multipliers for PRGT programs were on the order of 0.1–0.3 (albeit rarely significant) and on average modestly larger than assumed in initial program design. Unlike in initial program design, updated revenue multipliers were on average larger, and not smaller, than expenditure multipliers in both GRA and PRGT programs. Moreover, none of the updated expenditure multipliers were statistically significant.

Actual multipliers were estimated using a similar approach but based on cross-section data for program averages of growth deviations from the growth benchmark discussed in Chapter 3. The use of cross-section rather than panel data was motivated by the desire to obtain sharper estimates of fiscal multipliers given that in practice the growth impact

TABLE 2. FISCAL MULTIPLIERS IN PROGRAM DESIGN

PROGRAM	FISCAL MULTIPLIERS		
	OVERALL	REVENUE	EXPENDITURE
A. Initial Program Projections			
GRA	0.37*** -- 0.51***	0.24* -- 0.39*	0.34** -- 0.60***
PRGT	0.17** -- 0.22**	0.14** -- 0.16**	0.17* -- 0.20*
B. Updated Program Projections			
GRA	0.25** -- 0.35**	0.43** -- 0.50**	0.01 -- 0.18
PRGT	0.07 -- 0.32*	0.03 -- 0.41*	0.11 -- 0.20
C. Program Outcomes			
GRA	0.35*** -- 0.45***	1.12** -- 1.48***	0.34** -- 0.63**
PRGT	0.61***	1.02**	0.05

Source: Kim and others (2021).
Note: Fiscal multipliers are based on the estimated coefficients of fiscal adjustment, revenue adjustment, and expenditure adjustment in growth regressions. Asterisks denote statistical significance. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

---

[29] See Batini and others (2014), Gupta (2021), IMF (2017) and the references therein. IMF (2017) found that estimated fiscal multipliers in Sub-Saharan Africa tend to be lower than those typically identified in advanced or emerging market economies.

of fiscal adjustments could materialize over a period longer than a year.[30]

The results suggest that overall fiscal multipliers were in practice on the order of 0.35–0.45 in GRA programs, which lies within the range reported in the literature for advanced and emerging countries and also the multipliers effectively used in program design (Table 2, Panel C). For PRGT programs, the overall fiscal multiplier is estimated to be 0.6, which is substantially higher than suggested in the literature and the multipliers used in program design. When fiscal adjustment is disaggregated between revenue and expenditure adjustments, revenue multipliers exceed unity and are significantly larger than expenditure multipliers in both GRA and PRGT programs. The estimated expenditure multiplier is particularly low in PRGT programs (where it is not statistically different from zero). These findings are at odds with existing evidence that revenue-based adjustment is generally less contractionary than expenditure-based adjustment (Batini and others, 2014; Gupta, 2021).

Lessons from Country Experience

Based on the case studies, fiscal multiplier discussions between staff and officials were typically quite limited, with few examples in which staff offered more than ad hoc adjustments to standard multipliers. Staff reports supporting program requests rarely provided explicit analysis of multiplier assumptions.[31] In some cases, staff sought to modify standard assumptions, for example lowering the assumed multiplier in Grenada (2014) as a small very open economy, in Honduras (2014) to reflect beneficial effects through improved confidence, in Latvia (2008) because the tightening followed a boom, and in Mongolia (2009) because large cutbacks in public investment would mainly affect imports. In contrast, it appears that a mis-specified multiplier assumption contributed to a growth shortfall in Jordan and Ukraine. In the case of Jordan, while multiplier assumptions were not discussed explicitly in program documents, the case study

cautiously pointed to a possibility that too low a multiplier assumption may have contributed to the program's growth shortfalls. In the case of Ukraine, staff suggested in ex post analysis that the deeper than expected contraction in output reflected excessive reliance on expenditure restraint rather than revenue measures to achieve fiscal objectives given their estimation of a higher expenditure multiplier (Mitra and Poghosyan, 2015).

In practice, in programs with a moderate degree of fiscal adjustment (typically PRGT programs and GRA programs out of a crisis), fiscal drag seems to have been quite modest, and countries were able to grow at or even better than their IEO-calculated benchmark, especially when benefiting from favorable supply conditions. By contrast, in a number of GRA crisis cases (e.g., Latvia 2008, Mongolia 2009, and Romania 2009), deep fiscal adjustment was associated with even deeper output contractions than projected. However, it is hard to assess the counterfactual given that weaker policy commitments could have undercut confidence gains and other factors such as credit constraints were also at play.

ATTENTION TO GROWTH-FRIENDLY FISCAL ADJUSTMENT

Increasing attention to growth objectives in Fund programs has led to efforts to design "growth-friendly" fiscal adjustment that can limit the adverse short-term effects from fiscal adjustment and contribute to raising medium-term growth potential. In this context, the distributional consequences of fiscal policy actions are relevant, particularly the impact on low-income and vulnerable groups.

Revenue Mobilization

Revenue mobilization is growth friendly to the extent that it provides additional resources for priority spending, occurs in a non-distortionary way, and avoids imposing a burden on low-income groups. Evidence of improvements in tax mobilization is particularly marked in PRGT

---

[30] Unlike in the regression analysis for fiscal multiplier assumptions in program projections, the focus here is to establish a causal effect of fiscal adjustment on growth. In this respect, it is important to allow a longer horizon than a year for fiscal adjustment to affect growth. Moreover, the use of cross-section data—that is, program period averages—helps to average out the cyclical component in the primary balance measured on an annual frequency and bring the resulting average primary balance closer to the structural primary balance. See Kim and others (2021) for further discussion.

[31] For example, fiscal multiplier assumptions were not mentioned in the program documents of Egypt, Honduras, Latvia, Malawi, Pakistan, Romania, Senegal, Tunisia, and Ukraine.

programs (Table 3). Relative to the pre-program period, the tax-to-GDP ratio rose on average by 1.1 percentage points in the post-program period, largely aided by increases in taxes on goods and services. Not surprisingly, tax mobilization was substantially stronger in completed programs (where the tax-to-GDP ratio increased on average by 3 percentage points between pre- and post-program periods) than in programs that went off track. Those countries with completed programs collected more revenues not only from taxes on goods and services but also taxes on income, thereby making the tax structure more progressive by bringing individuals and businesses with rising incomes into the tax net.

In GRA programs, the increase in the average tax-to-GDP ratio was much more modest, less than ½ percentage points. However, there is some evidence that the tax structure became more growth friendly in the post-program period.

The dependence on distortionary trade taxes fell by ½ percent of GDP while reliance on more efficient taxes on goods and services increased by 0.6 percent of GDP. Crisis programs were no different in terms of tax mobilization outcomes from other completed GRA programs.

Expenditure Trends

Growth-friendly expenditure adjustment prioritizes resources for capital spending that can bring long-term benefits for the economy and social spending that supports low-income and vulnerable groups. Overall, total general government outlays increased in PRGT programs, buttressed by higher tax revenues (see Table 3, which is based on reported budgetary data). Moreover, there was a clear shift in the composition of spending in favor of capital projects. Capital spending rose on average by 1.1 percent of GDP between pre- and post-program

TABLE 3. TAX AND EXPENDITURE TRENDS ASSOCIATED WITH IMF-SUPPORTED PROGRAMS
(In percent of GDP)

	PRGT PROGRAMS			GRA PROGRAMS		
	Pre-program	Program	Post-program	Pre-program	Program	Post-program
Taxes	13.6	14.4	14.7	20.1	20.5	20.4
On income, profits, and capital gains	5.5	6.3	5.6	9.8	11.3	9.7
On goods and services	3.7	4.5	4.3	7.6	6.9	8.2
On international trade and transactions	2.7	2.6	2.7	2.3	2.1	1.8
Social contributions	0.9	0.9	0.8	6.9	6.6	6.3
General government total expenditure	25.1	24.8	26.2	35.7	33.7	33.0
General government expense	17.4	16.8	17.2	32.2	29.9	30.1
Compensation of employees	6.8	6.4	7.3	9.5	9.3	9.2
Purchases/use of goods and services	4.3	4.1	3.7	5.6	5.6	4.9
Interest	1.2	1.5	1.6	2.9	2.9	2.8
Social benefits	2.6	3.0	2.2	9.4	9.2	9.3
Capital expenditure	7.3	7.4	8.4	5.0	4.5	3.7
Net acquisition of financial assets	0.1	0.4	0.1	0.6	−0.1	−1.4

Source: Gupta (2021).
Note: "Pre-program" captures the three years prior to a program's starting year; "Program" captures program years; "Post-program" captures the three years following a program's ending year.

periods, a desirable outcome for addressing the country's infrastructure needs. However, little progress was made in raising social benefits—these rose during the program itself but subsided thereafter.[32] Outlays for compensation of government employees increased during the post-program period, but this data set does not distinguish compensation to health and education workers, which could be associated with providing social support. As with revenues, performance was significantly stronger in on-track programs than in off-track programs, suggesting that program completion matters not just in the short run but also in the medium run.

In contrast, government spending declined in GRA programs by about 3 percent of GDP between pre- and post-program periods, reflecting in part greater focus on bringing down the fiscal deficit and providing room for the private sector to grow, as well as the more modest progress in raising revenues. Capital spending declined by 1.3 percent of GDP between pre- and post-program periods, while social benefits remained flat.

Data from the World Development Indicators allow for a more focused look at trends in education and health spending related to IMF-supported programs, and suggest that overall progress in raising such social spending in the program context was limited.[33] In GRA programs, health and education spending showed little movement relative to GDP, although health spending increased in the post-program period as a share of the total budget, suggesting some increased prioritization of the health sector in government budgets (Table 4). In PRGT programs where such spending was relatively low, health spending remained unchanged as a share of GDP while spending on education declined somewhat after the program ended. These results occurred even though PRGT program conditionality actively sought to protect or increase social spending, suggesting that program conditionality to protect or raise social spending in the short term was not enough to achieve sustained increases in such spending.[34]

TABLE 4. TRENDS IN HEALTH AND EDUCATION SPENDING IN IMF-SUPPORTED PROGRAMS

GRA PROGRAMS						
	In Percent of GDP			In Percent of Government Expenditure		
Expenditure	Pre-program	Program	Post-program	Pre-program	Program	Post-program
Health	3.5	3.5	3.6	10.2	10.3	11.3
Education	4.7	4.2	4.6	13.8	12.9	13.8
PRGT PROGRAMS						
	In Percent of GDP			In Percent of Government Expenditure		
Expenditure	Pre-program	Program	Post-program	Pre-program	Program	Post-program
Health	1.9	2.0	1.9	7.1	7.7	6.3
Education	4.7	4.5	4.2	17.6	16.9	16.3

Source: Gupta (2021).
Note: "Pre-program" captures the three years prior to a program's starting year; "Program" captures program years; "Post-program" captures the three years following a program's ending year.

---

[32] Social benefits are current transfers to households, which may be paid in cash or in kind, to provide for needs arising from events such as sickness, unemployment, retirement, housing, or family circumstances.

[33] The effects of IMF-supported programs on education and health spending have been widely debated in the literature. Some studies argue that austerity measures and particularly conditionality on the wage bill have lowered such spending (Ooms and Hammonds, 2009; Rowden, 2009). In contrast, Clements, Gupta, and Nozaki (2013) show that spending in the education and health sectors increased at a faster pace in countries supported by IMF programs than in other developing countries. Similar results are found by IMF (2017).

[34] The 2017 IEO evaluation on *The IMF and Social Protection* found that social and other priority spending floors in IMF-supported LIC programs were generally not very useful for safeguarding social protection expenditures in part because of the difficulties of establishing a useful measure.

ATTENTION TO GROWTH IN FISCAL REFORMS

IMF-supported programs have typically incorporated reforms in revenue and expenditure structure aimed at strengthening countries' long-term growth prospects. To foster such reforms, programs have included structural conditionality and been supported by IMF capacity development work.[35]

Fiscal SCs have played a particularly important role in PRGT programs, given the criticality of building and strengthening fiscal institutions in low-income economies. Nearly four-fifths of fiscal SCs were classified as intended to support fiscal adjustment. Fiscal SCs (covering actions related to revenues, expenditures, debt, civil service reform, and fiscal transparency) constituted two-thirds of all SCs in PRGT programs while half in GRA programs (Figure 20, Panel A). Most fiscal SCs were of low or medium depth, and only around 10 percent were of high depth in both GRA and PRGT programs (Figure 20, Panel B). A high proportion of fiscal SCs was implemented in both GRA and PRGT programs, but only a small fraction of fiscal SCs required a permanent institutional change, reflecting in part the low average depth of fiscal SCs.

On the revenue side, the bulk of the structural conditionality in IMF-supported programs focused on taxes on goods and services, followed by conditionality on taxes on international transactions and on income (Figure 21). It is notable that in PRGT programs the share of SCs on taxes on international transactions (mostly related to customs management and control) was as large as that on taxes on goods and services and significantly larger than the corresponding share in GRA programs, while the share for income tax was much smaller than that of GRA programs. In general, revenue measures focused on broadening the tax base including by curtailing exemptions and enhancing tax compliance through strengthened revenue administration (IMF, 2019d). In addition, a larger number of administrative measures were included as fiscal SCs in PRGT

FIGURE 20. SHARE AND DEPTH OF FISCAL SC
(In percent)

A. Share of Fiscal SC in Total SC

B. Fiscal SC by Depth
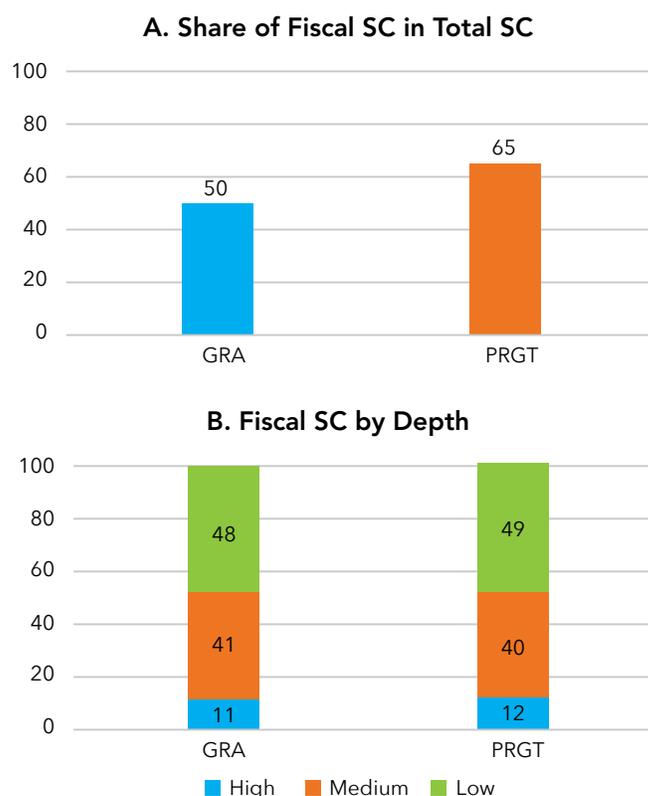
Source: Gupta (2021).

programs given the emphasis on mobilizing more domestic resources in PRGT countries to finance the Sustainable Development Goals by building new or strengthening existing revenue institutions.

On the spending side, SCs covered strengthening of budget preparation, debt management, and fiscal transparency. Conditionality on fiscal transparency was more prevalent in PRGT programs given the weaknesses in public financial management systems of LICs.[36] Weak public financial systems have been linked to widespread leakages of public resources and associated corruption (IMF, 2016). The IMF's new framework on governance stresses improvements in public financial management and enhancing transparency

---

[35] Chapter 7 of this report provides a fuller assessment of SCs and reforms in IMF-supported programs, including the impact of IMF capacity development support. The assessment of fiscal SCs provided here is based on the classifications and score indexes developed by the IEO (Kim and Lee, 2021).

[36] Fiscal transparency includes publication of financial statements of public institutions including state-owned enterprises. It also includes publication of details of infrastructural project costs/bids, publication of arrears and budget execution reports, passing and presentation of fiscal responsibility laws, and asset disclosures of cabinet members.

FIGURE 21. DISTRIBUTION OF TAX-RELATED SCs IN IMF-SUPPORTED PROGRAMS

Source: Gupta (2021).

of government operations as widespread and pervasive corruption can undermine program goals and thereby growth (IMF, 2018b).

Compliance with fiscal SCs was high overall but generally weaker in PRGT programs than GRA programs (Figure 22). In particular, compliance with fiscal transparency conditionality was significantly weaker in PRGT programs (by 19 percent) than in GRA programs.[37] Compliance in revenue and expenditure measures was also weaker in PRGT programs (by 12 percent), followed by those related to civil service and pension reforms (11 percent). While fiscal SCs were drawn increasingly more from IMF CD work, it is not clear from the data that growing provision of fiscal TA has helped to improve implementation of fiscal SCs. Specifically, no evidence is found for a positive and statistically significant relationship between fiscal TA and fiscal SC implementation in both bivariate and multivariate settings (Gupta, 2021).[38]

Stronger compliance and higher quality of fiscal SCs was associated with more growth-friendly fiscal outcomes.[39]

Dividing the sample of programs with overall fiscal consolidation into two subgroups depending on whether fiscal SC scores are above (first group) or below (second group) the cross-country median, the share of programs where fiscal adjustment relied more on revenue increases than expenditure cuts is on average 24–32 percentage points and 19–24 percentage points higher in the first group than in the second in GRA and PRGT programs, respectively (Figure 23).

Higher fiscal SC scores also have on average been positively and statistically significantly associated with higher social (health and education) spending, while a positive but insignificant association is found between fiscal SC scores and public investment in both GRA and PRGT programs (Figure 24). Moreover, the impact of fiscal SCs on social spending depends not only on the implementation (Figure 24, top right panel) but also on the depth and growth orientation of fiscal SCs (Figure 24, middle and bottom right panels).

---

[37] In 2018, the IMF adopted a new framework on governance (IMF, 2018a) that calls for greater attention to be paid to strengthening public financial systems and enhancing fiscal transparency in surveillance, program, and CD work. Staff are now producing detailed governance diagnostic reports for an increasing number of (mostly program) countries.

[38] See Chapter 7 for a fuller assessment of the relationship between IMF CD support and overall SC implementation in the program context.

[39] Based on the literature, a growth-friendly fiscal outcome is defined as a program where fiscal adjustment relies more on revenue increases than expenditure cuts.

FIGURE 22. IMPLEMENTATION OF FISCAL SC BY AREA

(Percentage of SCs met or met with delay)

Sources: MONA database; IEO staff calculations.
Note: "Combined" and "Civil/Pension" refer to fiscal SCs related to both revenue and expenditure and fiscal SCs related to civil service and pension reforms, respectively.

FIGURE 23. SHARE OF PROGRAMS WITH GROWTH-FRIENDLY FISCAL OUTCOMES

(In percent)
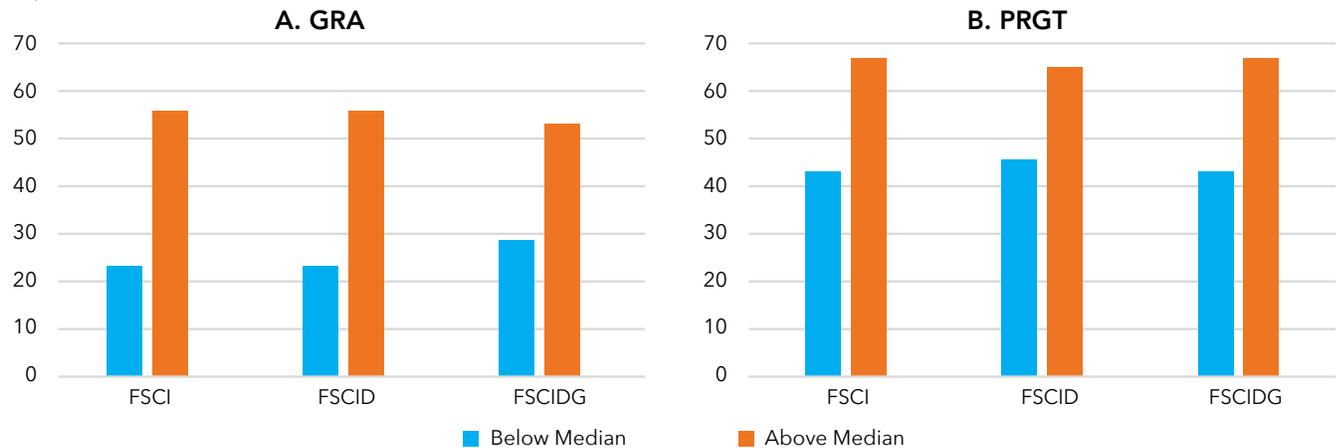
Source: Gupta (2021).
Note: Growth-friendly fiscal outcome is defined as a program where fiscal adjustment relied more on revenue increases than expenditure cuts; FSCI denotes the aggregate index for fiscal SC implementation score; FSCID and FSCIDG denote the aggregate composite indices for implementation and depth scores and for implementation, depth, and growth-orientation scores, respectively.

FIGURE 24. FISCAL SCs, PUBLIC INVESTMENT, AND SOCIAL SPENDING

FSCI vs. ΔPUBINVY

y = 0.156x - 2.7786
R² = 0.1073

y = 0.0043x - 0.9937
R² = 0.0008

FSCI vs. ΔSOCIALY

y = 0.0437x - 0.6799
R² = 0.292

y = 0.0428x - 1.5544
R² = 0.4231

FSCID vs. ΔPUBINVY

y = 0.2379x - 2.4725
R² = 0.0845

y = 0.0097x - 1.0044
R² = 0.001

FSCID vs. ΔSOCIALY

y = 0.0548x - 0.5061
R² = 0.1862

y = 0.0849x - 1.5339
R² = 0.4213

FSCIDG vs. ΔPUBINVY

y = 0.6103x - 2.5278
R² = 0.0895

y = 0.0594x - 1.1656
R² = 0.0095

FSCIDG vs. ΔSOCIALY

y = 0.132x - 0.5292
R² = 0.1789

y = 0.2033x - 1.5711
R² = 0.3976

■ GRA   ■ PRGT
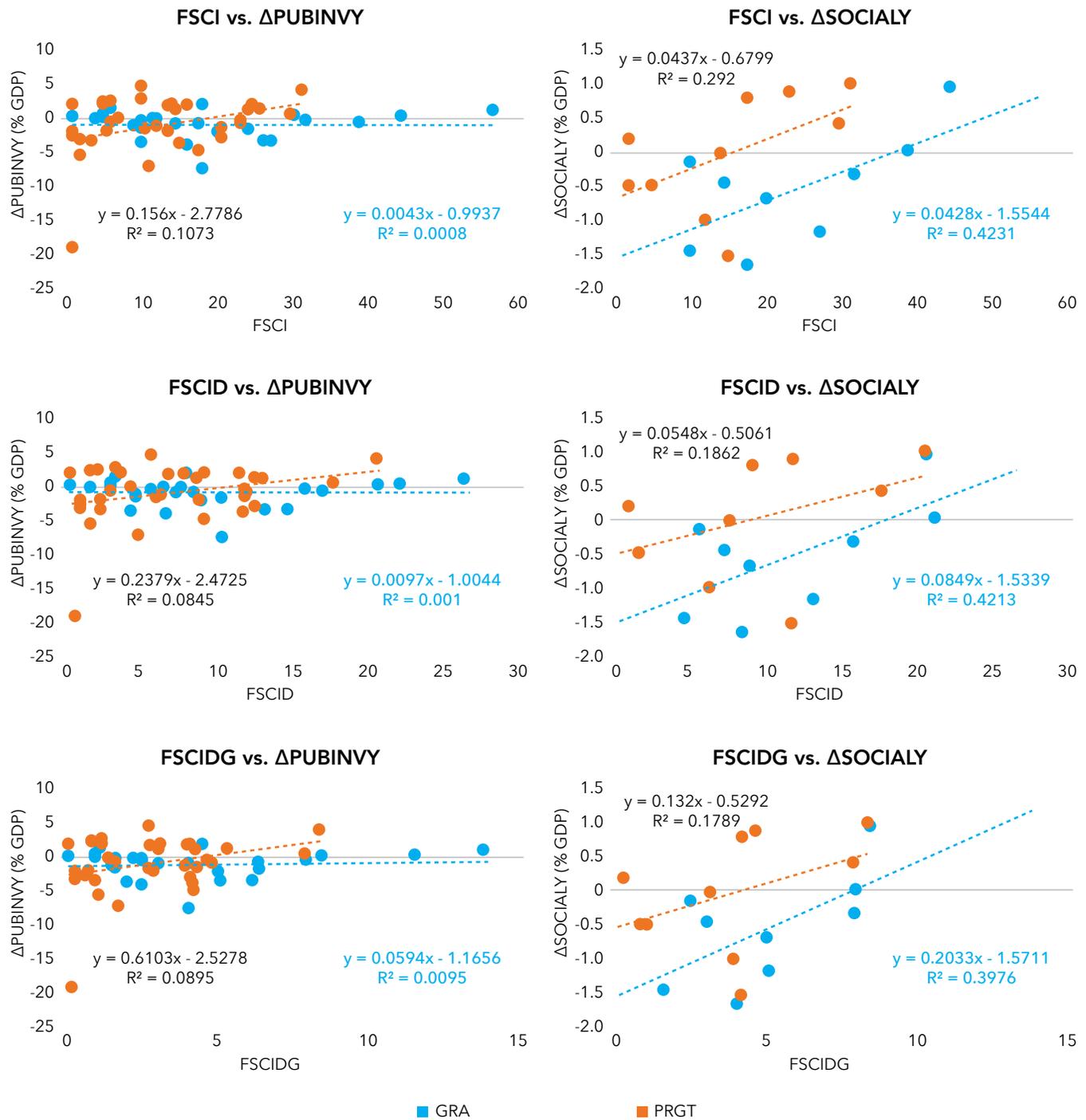
Source: Gupta (2021).
Note: See Figure 23 for the definition of FSCI, FSCID, and FSCIDG. ΔPUBINVY and ΔSOCIALY denote the cumulative change in public investment and social spending as a share of GDP during the program period.

LESSONS FROM COUNTRY EXPERIENCE

Composition of Fiscal Adjustment

The case studies reveal a wide variety of country experience within the aggregate data presented earlier. While most PRGT programs in the case studies envisaged growth-friendly revenue-based fiscal consolidation, in some programs such as Ghana (2009, 2015), Honduras (2014), and Malawi (2010, 2018), significant upfront reductions in expenditures were incorporated, given that the reforms needed to raise revenue would take time to accomplish.

Initial commitment to growth-friendly fiscal adjustment often needed to be adapted in the face of implementation slippages. Programs had to pull back from growth-friendly fiscal strategies during implementation owing to weak revenue results, shortfalls in grants, and/or pressures on current spending, including strong resistance to public wage restraint. In a number of cases, these pressures led to cuts, as opposed to programmed increases in public investment (Benin, Jordan, Malawi, and Tunisia). In some cases, weak expenditure controls contributed to large accumulation of public domestic arrears with a negative impact on growth (e.g., Benin, Cameroon, Grenada, Malawi, and Senegal); subsequent clearance or reduction of arrears supported growth by freeing up resources for the private sector (Grenada, Romania, and Senegal).

In many case studies, revenue mobilization efforts had modest success during the program period. In a number of cases, tax reforms faced substantial internal opposition and complicated implementation challenges and were delayed or watered down. In some PRGT cases, an increase in tax revenue was achieved after the program or during a successor program (Benin), while in others it remained elusive even after reform was rolled out (Bangladesh). Revenue measures proved challenging in some GRA countries too, particularly in the face of political resistance (Jordan 2016).

Public Investment for Growth

In a number of the case studies, programs were able to accommodate significant increases in public investment. For example, in the case of Honduras, increased revenues and efforts to contain non-essential spending were instrumental to increase spending on infrastructure. In Senegal, increased access to external finance helped by use of IMF signaling instruments and flexibility in adapting the fiscal program to shortfalls in adjustment allowed a build-up in such spending. In Egypt, increases in capital spending by state-owned enterprises were not restrained by program conditionality. Targeted efforts at protecting public investment were generally successful. Bangladesh (2012), for example, protected 10 mega public investment projects (in areas such as transport, ports, and power generation) that were already underway at the start of the program, despite the associated strain on public finances and the current account. In other cases, public investment spending was ring-fenced from budget cuts and supported by external financing. In Latvia and Romania, pro-growth capital spending was supported by EU financing.

At the same time, case studies also raised questions about the medium-term growth impact of higher public investment, highlighting the need for strong management and governance, including greater attention to the assessment of infrastructure gaps, transparent selection of projects based on cost-benefit analysis, effective monitoring and execution, and improved debt management. While scaling up infrastructure spending to support growth was a primary fiscal objective of many PRGT programs, some countries accumulated debt rapidly and did not always see the expected productivity gains from increased public capital, raising concerns about the quality and growth impact of public investment as well as the risks created for debt sustainability (Benin and Senegal). In Senegal, for instance, debt tripled between 2008 and 2019. A Public Investment Management Assessment (PIMA) highlighted the low quality of public investment spending in Benin, Cameroon, and Senegal. Another factor leading to rapid increases in debt was off-balance sheet operations, including from state-owned enterprises, and hidden domestic arrears, pointing to the need for monitoring comprehensive fiscal targets and debt (Cameroon).

In response to governance concerns surrounding public investment, the IMF has provided TA to enhance the efficiency of public investment (Cameroon and Senegal) and to establish a legal framework for public private partnerships to encourage private sector participation in public investment (Benin). Partner institutions have also contributed to this effort. The World Bank provided

extensive TA on public investment management for Cameroon, Jordan, Malawi, and Mongolia. The USAID also joined in this effort for Jordan, and the EU for Malawi. However, the results of such efforts have typically taken considerable time to materialize.

Subsidy Reform and Social Spending for Inclusive Growth

Case studies show widespread attention in program discussions to supporting social spending and inclusive growth. Approaches followed were adapted to account for institutional capacity and national preferences, and typically involved considerable support from the World Bank and regional development banks with deeper experience and expertise in this policy topic. Particular areas of focus included reforms of costly and distortionary energy and food subsidies and provision of social transfers to protect vulnerable groups.

Varying success was achieved in reforming energy subsidies. In Egypt (2016), energy subsidy reforms were extensively prepared, had strong political support and were accompanied by social assistance measures to protect the most vulnerable, such as additional food subsidies, cash transfers to the elderly and poor families, and other targeted social programs. Notwithstanding progress, however, concerns remained about the adequacy of targeting. In Jordan, energy subsidy reforms gained less traction because the cash transfer intended to protect the vulnerable from the impact of the automatic fuel price adjustment was poorly targeted, resulting in 70 percent of the population receiving this cash transfer. In some cases, political factors were behind delays or failures in energy sector reforms. In Ukraine, meaningful progress on curtailing gas losses was eventually achieved in the 2015 EFF after serious shortfalls in the 2008 and 2014 SBAs. Similarly, in Pakistan, progress was made in power sector reforms in the 2013 EFF after failure in this area in the 2008 SBA.

Two lessons emerge from this varied country experience. First, strong domestic political commitment and ownership are crucial to overcome resistance from vested interests. Second, particularly where progress is key to program success, the Fund may need to pay more attention to mobilizing technical support for reforms with development partners, including applying leverage related to the IMF's

access to senior decision makers, rather than taking a "hands off" attitude and effectively delegating responsibility to other agencies.

Turning to social safety nets, the case studies confirm that effective steps to safeguard vulnerable groups from the adverse impact of fiscal adjustment and energy subsidy reform can play an important part to maintain domestic support (Grenada, Honduras, and Malawi). However, in some cases (e.g., Romania 2009), reforms to the social safety net were too slow to provide protection to the most vulnerable. Nearly all programs aimed at improving the efficiency of social spending through better targeting. In some cases, the Fund adapted their recommended approach in view of national preferences (e.g., Latvia 2008) or faced a lack of political buy-in (e.g., Mongolia 2017). Given limited expertise of Fund staff in this area, social programs were often implemented with TA from the World Bank and other multilateral institutions (Bangladesh, Benin, Grenada, Honduras, Latvia, Malawi, Mongolia, Romania, Tunisia, and Ukraine).

More broadly, all PRGT programs have been required to include indicative targets establishing floors on social spending. Some GRA programs also established indicative floors on social spending (Jordan and Tunisia). Design of these floors has had to be adapted to data availability and are often broad rather than well targeted. As a result, even though the floors have typically been observed, in some cases, they have not been effective for safeguarding social expenditures for low-income and vulnerable groups (IEO, 2017). Malawi provides a case in point as its indicative floor on social spending was progressively lowered and the coverage was broadened over time.

The IMF's increased attention to social spending was well-received by many, though not all, country authorities. Ghanaian authorities, for example, "appreciated that they had flexibility on the choice of fiscal measures" that allowed them to implement free high school education. Latvian authorities, however, expressed "concerns about program costs, possible adverse impacts on work incentives, and cultural preferences for family-based rather than government-based safety nets." Malawi authorities, meanwhile, thought growth-promoting development programs should be prioritized ahead of social spending, which includes items that do not directly benefit the most vulnerable such

as the wages and salaries of government employees. In some cases, authorities preferred to follow universal rights rather than targeted approaches usually advocated by the Fund, arguing that a targeting approach can be more expensive than universal programs and/or exclude large segments of vulnerable populations.[40] In Mongolia, for instance, the 2009 SBA included conditionality to shift from universal to targeted social transfers. But progress was difficult without a clear social and political consensus.

One general issue raised in many of the case studies relates to the lack of monitoring and reporting on the social impact of the overall program and specific policies to protect vulnerable groups. Very few staff reports provided much information in this area, which made it hard for this evaluation to reach conclusions on the extent to which the undoubted attention to protecting the vulnerable was actually successful in achieving its objectives. Even more important, the lack of a capacity to track effectiveness made it hard to identify emerging risks and assess the need for further reinforcing actions.

Fiscal Reforms and Growth

A clear lesson across the case studies is that meaningful fiscal reforms often take a long time to put in place, implying that the growth benefits are slow to materialize, often well beyond the program period. This experience implies a need for realism in setting program timelines for reform progress and in anticipating growth benefits. It also underlines the importance of building political support for reforms and sustaining reform efforts well after programs are completed.

A case in point is Bangladesh (2012) where the introduction of the VAT was the centerpiece of fiscal reform of the program. Notwithstanding the enactment of the VAT law in 2012 and a short extension of the program to give time for reform implementation, the tax reform was not implemented because of strong opposition from vested interests. Eventually, Bangladesh implemented a VAT four years after program completion, but a positive revenue impact has been slow to materialize.

The case of Mongolia also highlights the importance of continued efforts to preserve program achievements after the program; otherwise the hard-earned reform benefits may be lost. The major achievement under the 2009 SBA was the establishment of the fiscal stability law (FSL) which was set up to avoid boom-bust fiscal cycles related to world mineral prices and to ensure saving of a greater proportion of windfall revenues during the good times. However, the fiscal rules under the FSL were circumvented after the SBA was completed and, as a result, the boom intensified during 2011–14 as mineral production expanded, and commodity prices rose. When the boom turned to bust, authorities sought a new three-year program in 2017.

Country case studies on Egypt, Jordan, and Tunisia illustrate the importance of sequencing fiscal adjustments, garnering political support and avoiding reform fatigue. All three countries were affected by economic and social disruptions in the context of the Arab Spring uprisings and protracted regional conflicts. They all asked for Fund financial support to attain two key objectives—regaining macroeconomic stability and reinvigorating growth to address longstanding social problems at the root of the uprisings. Among the three, only Egypt was able to attain both the adjustment and growth goals although it took a number of years for Egypt to lay the groundwork before a program could be put in place. At the start of the program, the Egyptian authorities took decisive policy measures, centered on fiscal consolidation and the liberalization of the foreign exchange market, and used the savings from the elimination of fuel subsidies to strengthen social programs for the most vulnerable. Such early decisive actions helped the authorities to avoid reform and donor fatigue which emerged in the other two countries over time and fueled skepticism and opposition to reform.

The case studies illustrate the critical role of program ownership in implementing fiscal adjustment and reaping the benefit of fiscal reforms. In Latvia (2008), which involved bold, front-loaded fiscal adjustments, authorities retained strong political support despite a sharp initial growth downturn by clearly communicating to the public the national strategy for sustaining progress towards closer EU integration. The program achieved its

---

[40] While IMF policy advice on social spending often centers on targeting mechanisms based on means testing, IMF (2019d) noted that the appropriate use of targeted and universal transfers depends on country preferences and circumstances and should be consistent with fiscal and administrative constraints.

goals and resolved external and fiscal imbalances relatively quickly. After three years of recession, growth resumed at a respectable pace and Latvia joined the Eurozone in January 2014. By contrast, Pakistan (2008 SBA) is a case in which limited political support undermined critical tax reform implementation. Despite massive IMF TA and program extensions to allow more to time to unlock the VAT reform, the tax reform failed to be implemented before the program expired in 2011 due to lack of political support. As the VAT reform remained politically infeasible, another program in 2013 sought revenue mobilization via a combination of several incremental reforms (e.g., scaling back tax exemptions, broadening the tax base, increasing goods and services taxes, and improving tax administration) and achieved partial success.

Domestic Arrears

A number of case studies highlight that large accumulation of public domestic arrears has a negative impact on growth and business climate (e.g., Benin, Cameroon, Grenada, Malawi, Romania, and Senegal). These strains typically received considerable attention in program design although implementation of needed reforms was often challenging. In Senegal (2015), a large increase in domestic arrears was not reported on a timely basis and required introduction of appropriate program conditionality. In Malawi (2012), program efforts to control government spending were partly circumvented by the accumulation of arrears to domestic suppliers, which led to increased nonperforming loans (NPLs) in the banking sector. Creditors were issued promissory notes amounting to a cumulative 9 percent of GDP between 2013 and 2018. The 2012 program also included conditionality on public financial management designed to better limit and track new arrears. Cameroon (2017) aimed at a gradual elimination of domestic arrears to contractors to reduce banks' NPLs and unlock bank credit to the private sector, but efforts have not been effective, resulting in persistent domestic arrears. In Romania (2009, 2011), both authorities and staff viewed efforts to strengthen public sector arrears management as having particularly benefited growth and business climate as payment arrears by state-owned enterprises, local governments, and health sector bodies were an important burden on the business community. Pre-existing arrears were regularized, and disciplinary rules were established to deter new arrears. In

Grenada (2014), the private sector benefited from improved fiscal management as the government had a history of accumulating significant arrears on domestic payables, which were eliminated under the program. In Benin (2017), business climate benefited from the completion of an audit of government arrears to domestic suppliers.

ASSESSMENT

The experience covered in this evaluation clearly shows the Fund paying considerable attention to encouraging growth-friendly fiscal adjustment and reforms. Empirical analysis suggests that program design generally incorporated multiplier relationships broadly in line with the professional literature. Moreover, efforts were made to raise revenue mobilization, increase public investment, and at least protect social spending, by building these goals into program objectives and structural conditionality, often with capacity development support. These efforts were usually tailored to country circumstances, including through program adaptation that helped to adjust for slippages, alleviate the growth impact, and foster country ownership.

Overall, PRGT programs did manage to raise revenues and public spending on average; efforts at raising social support were less successful, although assessment is complicated by lack of data. GRA programs by contrast were less growth-friendly, relying more heavily on spending cutbacks, including to public investment and social spending. The cross-country analysis and country case studies suggest a number of lessons.

First, more systematic attention should be paid to calibrating the growth consequences of fiscal adjustment and reforms in program design. Our cross-country evidence suggests that actual short-term fiscal multipliers could differ substantially from assumed multipliers. In particular, assumed fiscal multipliers seem to have been typically too low in PRGT cases (where results suggest that the adverse growth impact of higher revenue is under-estimated and growth benefit of higher public investment is over-estimated). In GRA cases, multipliers seem to have been under-estimated in some circumstances when the economy was weak and there was little scope for offsetting monetary support—although multipliers seem to have been less under-estimated in crisis cases where confidence effects helped in part alleviate the income

effect of adjustment. In this respect, explicit analysis of short-term fiscal multipliers in staff reports would enable a better understanding of the underlying assumptions and assessment regarding the short-term growth consequences of fiscal adjustment and could help reduce growth optimism bias. In addition, explicit discussion in program documents of the longer-term growth impact of fiscal reforms would help to incentivize authorities to undertake growth-enhancing reforms as well as to underpin realistic assessment of medium-term debt sustainability.

Second, fiscal structural conditionality should give greater emphasis to fostering deep reforms, including through increasing the proportion of high-depth SCs and cutting back on the proliferation of low-depth ones. This would help focus authorities' attention on measures crucial for improving fiscal performance supportive of growth (such as enhancing the efficiency and compliance of the VAT and improving the governance of capital spending).

Third, programs should take a longer-term perspective on seeking to raise public spending on public infrastructure, education and health by focusing more on strengthening public financial practices instead of merely protecting or raising such spending through short-term conditionality.

The IMF has intensified efforts to help strengthen governance through technical assistance (TA) to member countries since the adoption of a new framework in 2018. The recommendations of those missions on fiscal transparency and public financial management should be incorporated in IMF-supported programs.

Fourth, programs should seek ways to incentivize the authorities to report domestic arrears with candor and on a timely basis and to prevent new arrears by, for example, limiting the scope for using off-balance sheet operations or using domestic arrears to meet program targets. In addition, ex post accommodation of unmet conditionality on domestic arrears should be granted as an exception based on clear justification. Fund TA related to PFM should pay particular attention to how to improve monitoring and control of domestic arrears.

Fifth, there is a need to strengthen the monitoring and reporting of the social and distributional impact of the overall program and of the specific policies to protect vulnerable groups. The lack of a capacity to track effectiveness made it hard to track progress made in achieving inclusive growth, to identify emerging risks and to assess the need for further reinforcing actions.

STRUCTURAL CONDITIONS, STRUCTURAL REFORMS, AND GROWTH

This chapter assesses the effectiveness of structural conditionality in promoting growth-enhancing structural reforms.[41] It also examines the role of IMF capacity development work and collaboration with partner institutions in supporting the design and implementation of SCs.

IMF-supported programs have used SCs to encourage needed adjustments, support structural reforms and ultimately promote growth. In the programs under evaluation, SCs accounted for more than 40 percent of total program conditions. The volume of SCs per program has increased significantly since the end of the 2000s, particularly in GRA programs, reflecting in part that more programs in the 2010s were dealing with protracted structural challenges in a weak global environment (Figure 25). The average number of SCs for GRA programs peaked in 2013 and has since been on a broadly declining trend except for 2017.[42] PRGT programs have exhibited similar time pattern but with on average fewer SCs and less time variation than GRA programs.

FIGURE 25. VOLUME OF SCs PER PROGRAM: 2009–19
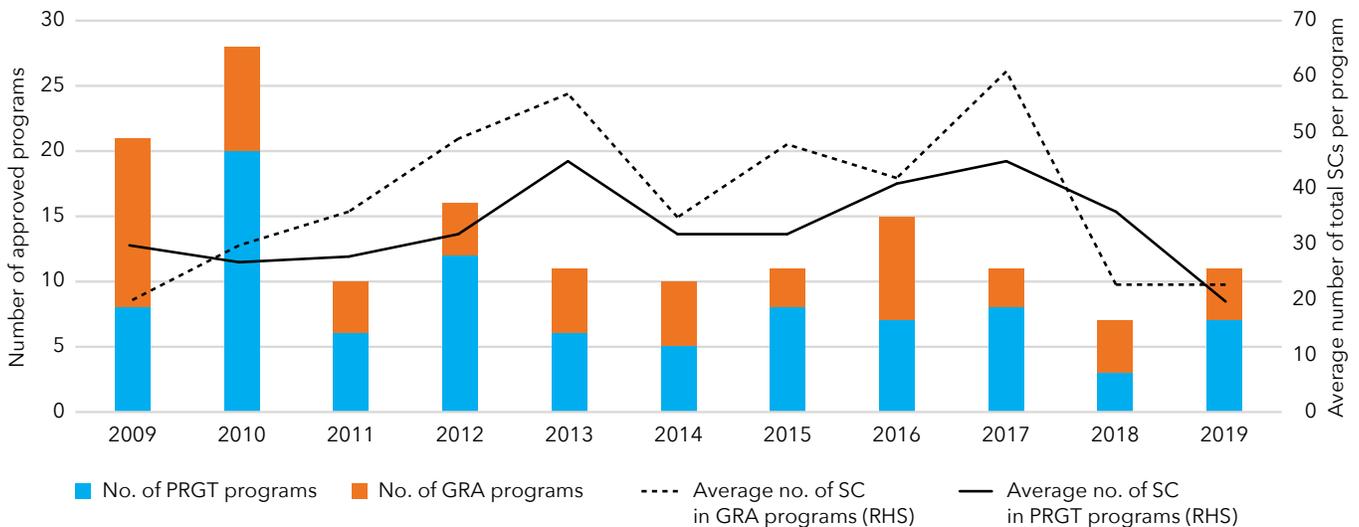
Source: Kim and Lee (2021).
Note: The average numbers of SCs for 2017–19 are preliminary estimates based on the updated data from the MONA database because some programs are still ongoing.

---

[41] This chapter draws on Kim and Lee (2021) and country case studies prepared for the evaluation.

[42] The spike in 2017 in the average number of SCs in GRA programs is explained mainly by the fact that one of the three GRA programs approved in 2017 had an exceptionally large number of SCs (80 in total).

COMPOSITION, IMPLEMENTATION, AND DEPTH OF STRUCTURAL CONDITIONS

SCs in initial program design (i.e., at approval of the program) were mostly of low to medium depth, largely related to demand management, and in the fiscal area (Figure 26).[43] Only about 10 percent of the SCs were aimed at growth and efficiency objectives. About 70 percent of SCs were in the area of the IMF's core expertise and the remaining 30 percent were in non-core areas or areas of

shared expertise with other international development institutions (IDIs). Interestingly, the shares of high depth SCs and SCs directly related to growth and efficiency were both higher in GRA programs than PRGT programs (Table 5). By contrast, PRGT programs had a significantly higher share of SCs in the fiscal sector and for demand control.

Focusing on observed SCs for which implementation status was determined in a completed program review, SC implementation was on average stronger in GRA programs and in countries in Europe and Latin America than in PRGT

FIGURE 26. COMPOSITION OF STRUCTURAL CONDITIONS BY DEPTH AND SECTOR

A. IMF-Supported Program Conditions

Structural Conditions 42%
Quantitative Conditions 54%
Indicative Targets 4%

B. SC by Sectoral Composition

Other 15%
Monetary/ Financial/ Exchange Rate 28%
Fiscal 57%

C. SC by Depth

High Depth 12%
Medium Depth 37%
Low Depth 51%

D. SC by Content

Vulnerability Management 23%
Growth/ Efficiency 10%
Demand Management 67%
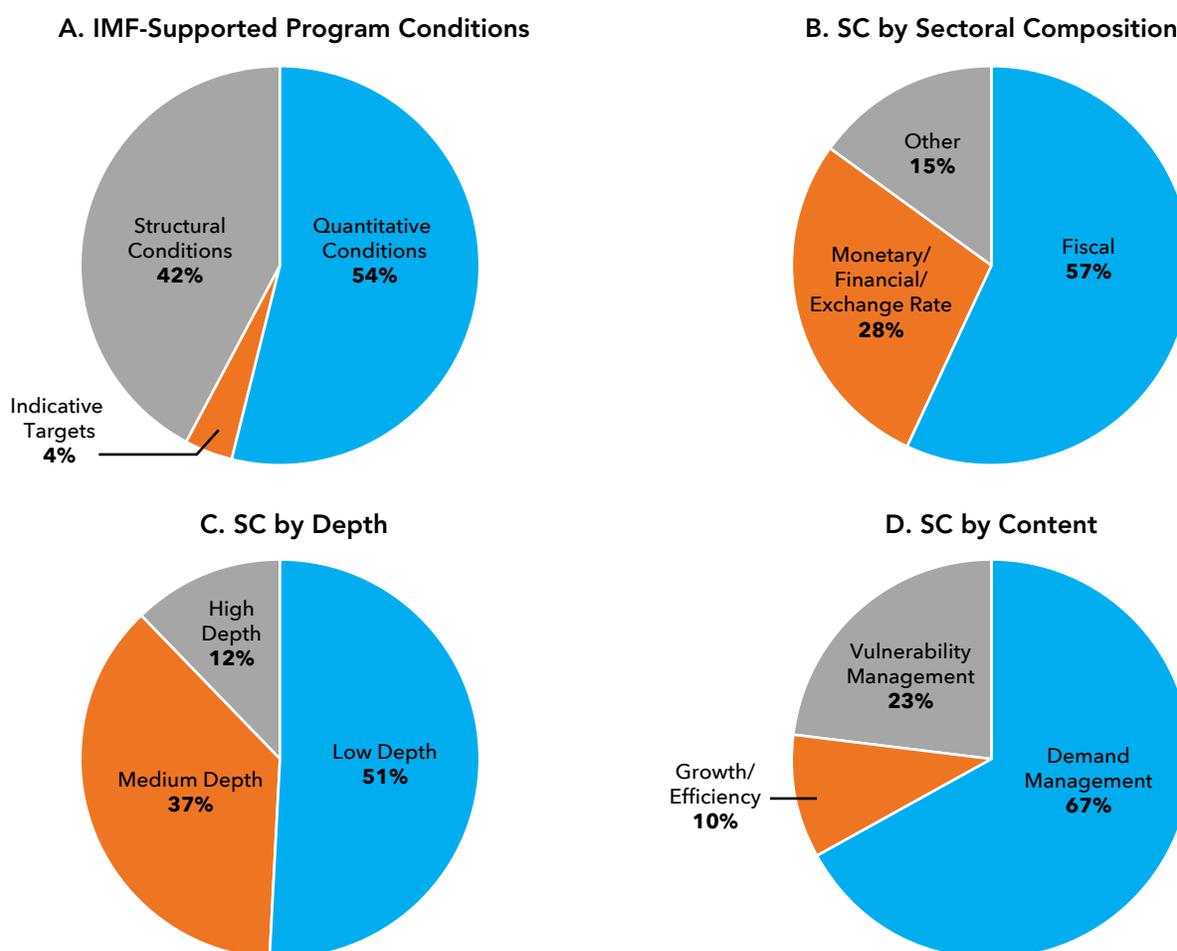
Source: Kim and Lee (2021).

---

[43] Depth of an SC is assessed based on the methodology developed by the IEO's evaluation of structural conditionality (IEO, 2007) and data put together in the 2018 ROC (also see footnote 18 for the definition of depth). An example of a high depth SC would be "Parliamentary approval of the revised PFM legislation" (Grenada 2014 ECF). An example of a medium depth SC would be "Install the new IT software at the central server site (NAIS) and commence testing" (Albania 2014 EFF). An example of a low depth SC would be "Start posting on the central bank website the national accounts and CPI data, as well as detailed methodological information, and a calendar of upcoming data releases" (The Gambia 2012 ECF). See Kim and Lee (2021) for further details.

TABLE 5. STRUCTURAL CONDITIONS BY DEPTH, CONTENT AND SECTOR
(In percent of total)

	DEPTH			CONTENT			SECTOR		
	High	Medium	Low	Demand Control	Growth/ Efficiency	Vulnerability Management	Fiscal	Monetary/Financial/ Exchange Rate	Other Structural
Total SCs	12.4	36.9	50.6	67.1	10.3	22.7	57.0	28.1	15.2
GRA	15.9	36.3	47.8	60.6	12.0	27.5	48.3	33.4	18.2
PRGT	9.3	37.5	53.1	72.8	8.8	18.4	63.9	23.5	12.5
Unobserved SCs[1]	12.9	36.2	50.9	62.5	10.0	27.5	52.9	30.8	16.3
GRA	16.1	34.2	49.8	57.6	10.7	31.7	45.0	37.1	17.9
PRGT	9.4	38.4	52.2	68.0	9.2	22.8	61.2	24.2	14.6

Source: Kim and Lee (2021).
[1] Unobserved because associated reviews were not completed.

TABLE 6. IMPLEMENTATION STATUS, DEPTH AND CONTENT OF SCs

	IMPLEMENTATION	DEPTH	CONTENT
GRA (52)	0.86	0.55	0.47
PRGT (73)	0.77	0.54	0.45
AFR (54)	0.74	0.54	0.45
APD (6)	0.78	0.48	0.43
EUR (26)	0.87	0.57	0.50
MCD (22)	0.86	0.53	0.47
WHD (17)	0.88	0.55	0.42
Mean	0.81	0.54	0.46
Median	0.83	0.53	0.46

Source: Kim and Lee (2021).
Note: Numbers in parentheses represent the total number of programs in the relevant category; implementation, depth, and content figures are average scores per SC.

programs and countries in Sub-Saharan Africa (Table 6).[44] The average implementation score was relatively stable over time in GRA programs but was on average on a downward trend in PRGT programs (Figure 27). Both depth and growth-orientation scores were relatively stable over time in both GRA and PRGT programs.

SC implementation was on average somewhat weaker for programs with a higher volume of SCs (Figure 28). SCs with higher depth would normally be considered more challenging to implement than lower depth SCs because the former require more technical input and stronger political commitment. However, no statistically significant relationship was detected between implementation and depth nor between implementation and growth orientation.

Similarly, no significant relationship was found between implementation and the country's institutional capacity measured by the Government Effectiveness Index (GEI) published by the World Bank.

STRUCTURAL CONDITIONS AND IMF CAPACITY DEVELOPMENT

The 2018 ROC concluded that the focus of SC and IMF capacity development work have been broadly well aligned, especially in the areas of the IMF's core expertise. Consistently, data on programs included in this evaluation show that more technical assistance resources (measured in full-time equivalent (FTE) units) were allocated to

---

[44] Implementation status is not provided in the MONA database for SCs in program reviews that were never completed.

FIGURE 27. IMPLEMENTATION, DEPTH, AND CONTENT OF SCs: 2009–16

GRA

PRGT

Average Completion   Average Depth   Average Growth Orientation
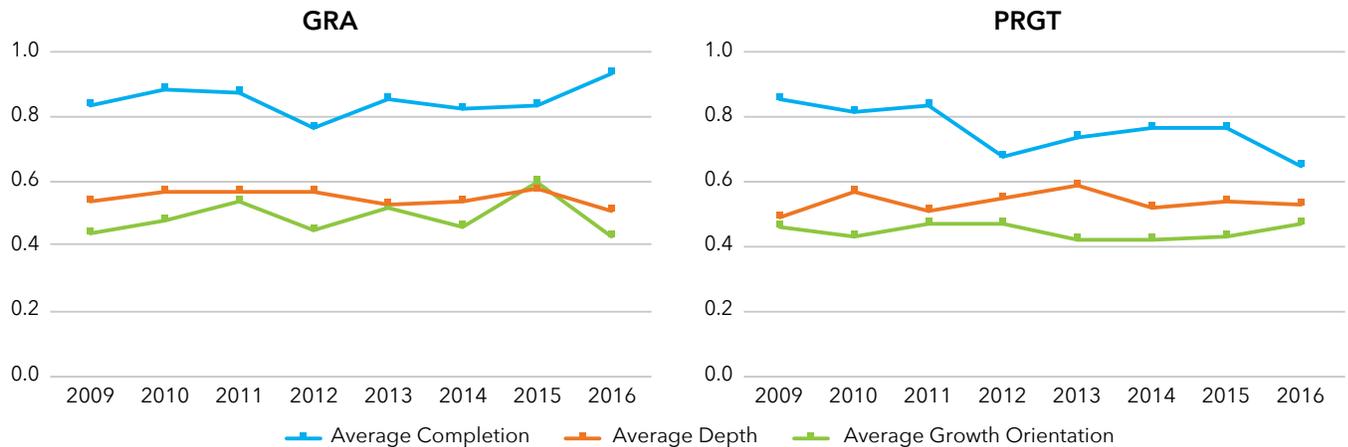
Sources: MONA database; IEO staff calculations.
Note: The year on the horizontal axis represents the year of program approval.

FIGURE 28. VOLUME OF SCs AND SC IMPLEMENTATION

$y = -0.0017x + 0.8462$
$R^2 = 0.0273$
P-value=0.0642

Source: Kim and Lee (2021).

FIGURE 29. VOLUME OF SCs AND TA DELIVERY

$y = 0.1331x + 2.9538$
$R^2 = 0.1845$
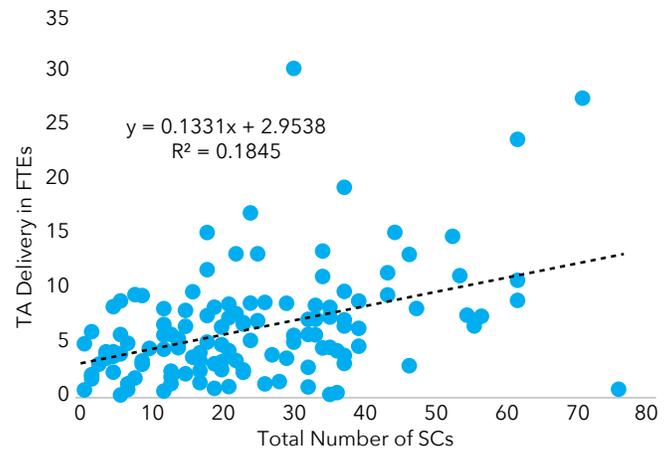
Sources: MONA database; IEO staff estimates.

programs where SCs were most actively used (Figure 29). The 2018 Review of the Fund's Capacity Development Strategy also suggested that integration between the IMF's CD and lending has strengthened as it has often been integral to a program's design and implementation framework.

Two related questions are how well CD provision has been aligned with country need or capacity and how CD has affected the implementation of SCs. Cross-country evidence is less encouraging on these questions. Data on

programs during the evaluation period suggest that more TA resources seem to have been allocated to program countries with higher, and not lower, capacity. Specifically, the bivariate relationship was positive, albeit not statistically significant, in both full and PRGT samples between country capacity (measured by the GEI) and TA delivery during programs (Figure 30, Panels A and B). Indeed, more than half of the top recipients of TA were higher-capacity LICs comprised of frontier LICs and LICs that had issued Eurobonds at least once.[45]

---

[45] Frontier LICs include Bangladesh, Bolivia, Côte d'Ivoire, Ghana, Kenya, Mongolia, Mozambique, Nigeria, Papua New Guinea, Senegal, Tanzania, Uganda, Vietnam, and Zambia. Other LICs that have issued at least one international bond are the Republic of Congo, Ethiopia, Honduras, and Rwanda. See IMF (2015b).

Data also suggest that TA delivery has been negatively associated with the average SC implementation score and that the relationship is statistically significant both in the full and PRGT samples (Figure 30, Panels C and D). This finding—that is, the lack of a positive and significant relationship between TA delivery and SC implementation—continues to hold in a multivariate setting which controls for other factors that could affect SC implementation, such as the average depth of SC, the recipient country's implementation capacity (measured by the GEI), and the total volume of SCs (Kim and Lee, 2021).[46] The sectoral breakdown of TA delivery and SC implementation shows that the fiscal sector, which accounts for the largest shares of SCs and TA allocations, was also the dominant area of unmet SCs (Figure 31).

Given the high overlap in focus between IMF TA and SCs found in the 2018 ROC, these findings raise concerns about how effectively TA is integrated with program implementation and monitoring. The average SC implementation score (ASCI) was broadly similar across sectors except for the central bank/monetary sector, while TA provision was heavily focused on the fiscal sector (see Figure 31). The thematic background paper on fiscal issues (Gupta, 2021) observes that the Fiscal Affairs Department (FAD) of the IMF, which is responsible for fiscal capacity building, had little role in assessing compliance with fiscal SCs in revenue mobilization and public financial management. A review of back-to-office reports of the IMF's fiscal CD missions to 17 case study countries during 2008–19 suggests that with a few exceptions, fiscal CD missions did not discuss the status of fiscal reforms, which is striking given that structural

FIGURE 30. COUNTRY CAPACITY, IMF TA, AND SC IMPLEMENTATION

A. Country Capacity and TA Delivery: All Programs

$y = 0.4816x + 6.017$
$R^2 = 0.0002$
P-value=0.8830

B. Country Capacity and TA Delivery: PRGT Programs

$y = 0.6746x + 6.4242$
$R^2 = 0.0003$
P-value=0.8894

C. TA Delivery and ASCI: All Programs

$y = -0.008x + 0.8521$
$R^2 = 0.0405$
P-value=0.0251

D. TA Delivery and ASCI: PRGT Programs

$y = -0.0079x + 0.8226$
$R^2 = 0.038$
P-value=0.0983
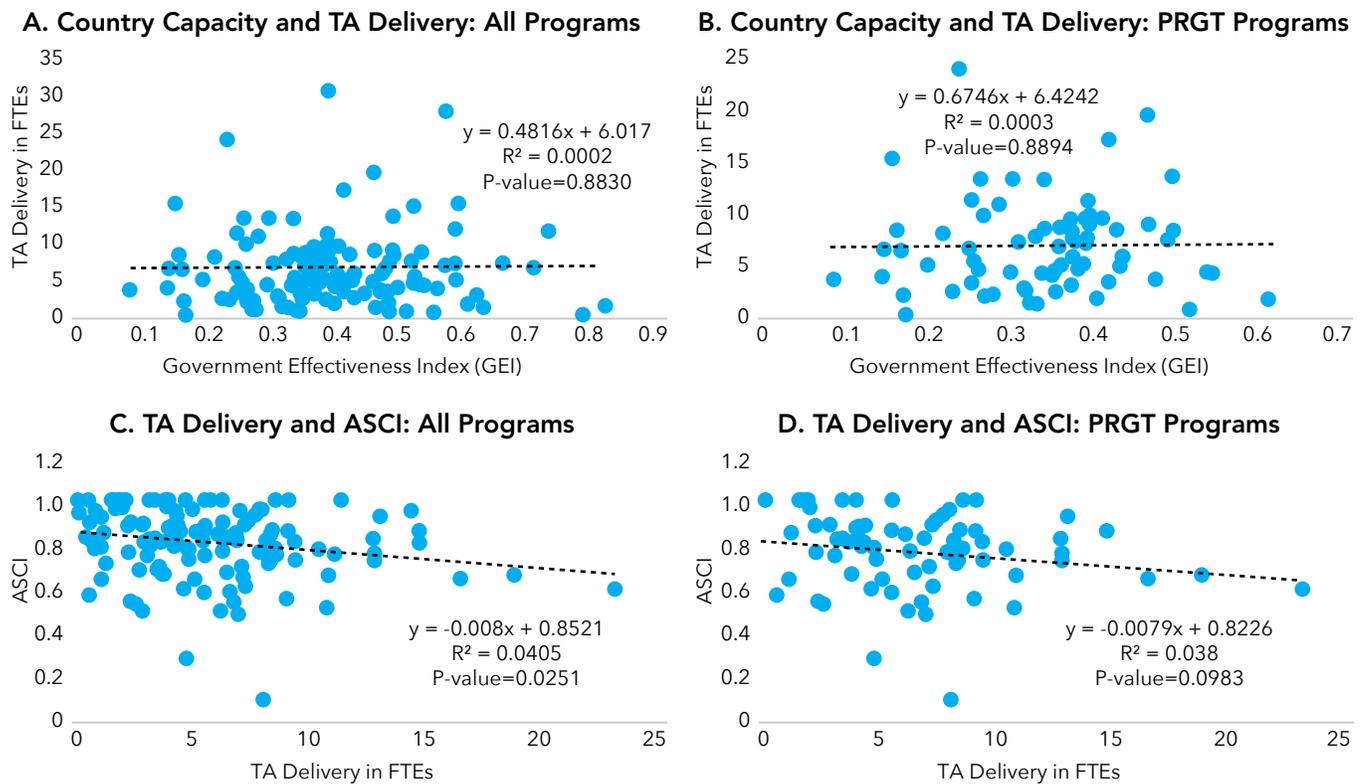
Source: Kim and Lee (2021).
Note: In Panels C and D, Greece (2012 EFF) and Ukraine (2015 EFF) are excluded where TA receipt during programs exceeded 26 in FTE (full-time equivalent) units.

---

[46] Specifically, the results of multivariate fractional logit analysis show that the relationship between IMF TA and SC implementation is statistically insignificant in both GRA and PRGT programs and continues to be negative in the latter.

FIGURE 31. IMF TA AND SC IMPLEMENTATION BY SECTOR
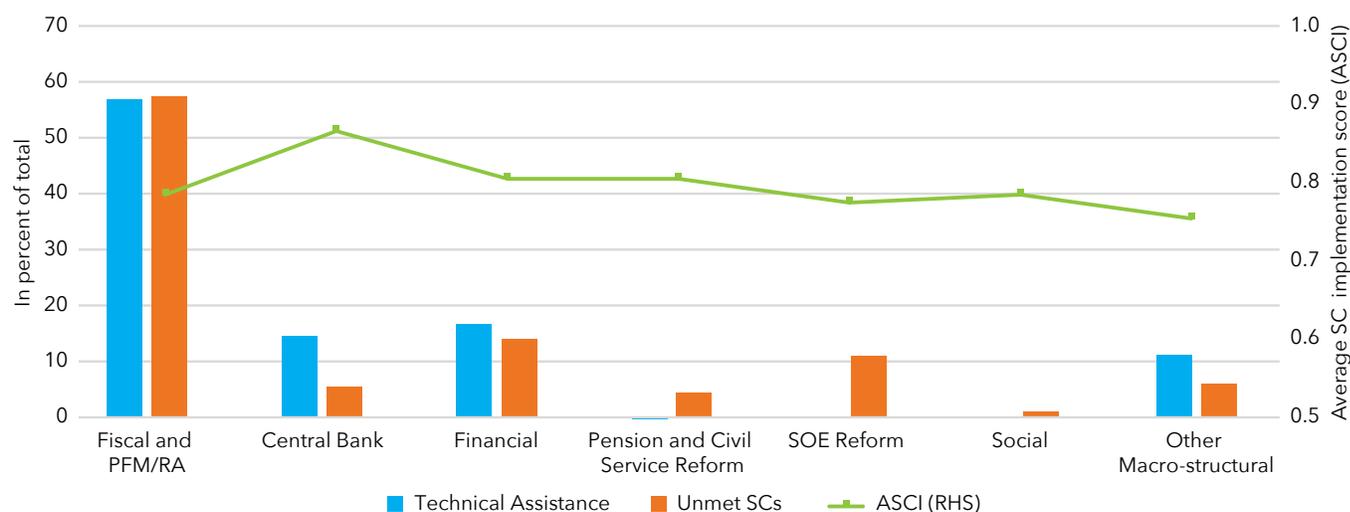
Sources: IMF, 2018 ROC; and IEO staff calculations.
Note: "Unmet SCs" refer to SCs assessed as "Not Met" in associated program reviews. PFM = public financial management; RA = revenue administration; and SOE = state-owned enterprise.

reforms covering domestic resource mobilization and public financial management are core to the department's CD work.

The question also arises as to whether TA has been delivered in the most effective way in the program context. The mode of delivery is especially important for low-income and fragile countries where institutional capacities are weak because these countries often operate on already thinly stretched human capital. See, for example, the discussions on the accompanying country case study on Malawi and the IEO evaluation of *The IMF and Fragile States* (IEO, 2018a), which concluded that TA work needs to be better tailored to be effective in difficult country circumstances. Addressing this challenge has been an important element of the Management Implementation Plan following that evaluation.

COLLABORATION WITH PARTNER INSTITUTIONS

About two-thirds of SCs in the evaluation period were in core areas of the IMF's expertise and the remaining one-third were in shared and non-core areas which

typically have a higher growth orientation and where other IDIs may lead in terms of knowledge and experience. However, fewer than 2 percent of total SCs were explicit about the collaboration with other IDIs. In this small subsample of SCs, the average score of implementation was even lower, especially among PRGT programs, suggesting that the IMF's collaboration with partner institutions with respect to the design and implementation of SCs outside the IMF's core expertise may have been less effective than desired (Table 7).

STRUCTURAL CONDITIONS AND STRUCTURAL REFORMS

Structural conditions apply to specific policy measures or actions to support structural reforms and are not a direct measure of structural reforms themselves. A recent IMF study developed a structural reform index (SRI) based on detailed information on regulatory stances and reform episodes in both real and financial sectors to quantify the degree of progress on reforms (IMF, 2019e).[47] Using the SRI, the study found positive evidence about the growth-enhancing effect of structural reforms.

---

[47] The SRI is constructed based on assessment of reforms in domestic finance (regulation and supervision); external finance (capital account openness); trade (tariffs); product market (regulation in electricity and telecommunication sectors); labor market (job protection legislation); and composite worldwide governance indicator. Each sector contains multiple sub-indicators which are scored between 0 and 1, and the aggregate reform index of each sector is obtained as an average of sub-indicator scores. See IMF (2019e) for further details.

TABLE 7. AVERAGE SC SCORES: CORE VS. SHARED/NON-CORE AREAS OF IMF EXPERTISE

		IMPLEMENTATION		DEPTH		GROWTH ORIENTATION	
Core	GRA	0.85	(0.80)	0.53	(0.54)	0.4	(0.39)
	PRGT	076		0.54		0.38	
Shared/non-core	GRA	0.82	(0.78)	0.53	(0.55)	0.65	(0.62)
	PRGT	0.74		0.57		0.6	
SCs that mention IDIs in the text	GRA	0.80	(0.73)	0.54	(0.52)	0.50	(0.56)
	PRGT	0.56		0.48		0.72	

Sources: MONA database; 2018 ROC; IEO staff calculations.
Note: Figures in parentheses are the averages of GRA and PRGT programs.

To connect the dots between structural conditionality and growth, we assessed how SC scores were related to changes in the SRI. Regression analysis for 32 programs (for which SC scores and SRI data are both available) found that the IEO's aggregate SC score indexes were positively and statistically significantly associated with the cumulative percentage changes in the SRI during the program period. Moreover, the positive relationship became stronger and more significant when the quality of SCs is accounted for. Specifically, the marginal impact of the SC score on the change in SRI was largest for SCIDG (which is an aggregate composite index of implementation, depth, and growth-orientation scores that seeks to capture the quality of SCs) and smallest for SCI (an aggregate index for implementation score only). Disaggregating between GRA and PRGT programs, the positive association between SC score indices and the SRI was statistically significant in PRGT programs but not in GRA programs. Although less reliable due to smaller sample size, this result suggests that high quality SCs may have had stronger traction in pushing for structural reforms in PRGT programs than in GRA programs.

These results provide useful empirical support for our assessment in Chapter 4 that higher quality SCs bring growth benefits in the post-program period.

LESSONS FROM COUNTRY EXPERIENCE

In general, the country case studies highlight the broad reform agendas included in IMF-supported programs supported by extensive structural conditionality. Consistent with the empirical analysis presented above, in most cases the focus of the reform efforts was in the IMF's core areas of expertise aimed at strengthening fiscal, monetary, exchange rate, and financial sector performance through developing policy making capacity and institutions. Less attention was paid to issues related to raising growth potential or improving the quality of growth by addressing market distortions and improving the business climate, although the focus increased in more recent programs, especially in repeat programs where growth performance had remained below aspirations despite progress towards macroeconomic stabilization.

Several case studies (e.g., Ghana, Grenada, Jamaica, Jordan, and Pakistan) highlight the challenges of adjusting adequately the volume and pace of structural reforms to the countries' capacity and circumstances, as well as building political and social consensus. Ambitious reform agendas often stretched the available absorption capacity, resulting in implementation delays. In this regard, country officials were generally very appreciative of the Fund's extensive technical assistance support, but commented that while helpful, the provision of IMF TA was not a full substitute for domestic implementation capacity. In some cases, officials noted that there were just too many SCs included in the program and a more focused agenda would have had more success. These examples suggest that greater selectivity in structural conditionality, better contingency planning and more cautious assumptions on feasibility of structural reforms may be called for.

Many case studies underline the limited depth and coverage of growth-relevant issues outside the Fund's core expertise and stress the need for stronger and proactive engagement with partner agencies for support in these areas. Indeed, country officials frequently mentioned a tendency for Fund staff to be more comfortable in the core areas and to be insufficiently engaged in helping countries to address growth-critical reform needs elsewhere. For example, the

Latvia, Malawi, and Pakistan (2008) programs paid little attention to structural issues outside the area of the Fund's core expertise and took a hands-off approach by relying on other agencies for SC implementation and follow-up. In Ghana, Grenada, and Jamaica, some macro-critical issues such as labor market and energy sector distortions were deep-rooted, but given that the Fund was not adequately equipped to address these issues, programs relied on interventions from partners such as the World Bank and regional development banks. In Jordan and Ukraine, officials commented that the IMF paid insufficient attention to reforms in non-core areas which were critical for faster growth, and in Romania, while the importance of reforms in non-core areas was discussed in program documents, they were not included as SCs.

Several case studies emphasize that staff had unrealistic expectations regarding the feasibility and growth payoffs of reforms. In Jordan, Pakistan, and Tunisia, case studies highlight the need for more cautious assumptions on feasibility and growth payouts of structural reforms. Fund staff underestimated the complexity of the political transition and the impact of intervening political, security-related and regional shocks. The consequence was a disconnect between optimistic growth projections and actual outcomes. This gap also reflected "the need to show hope," which was also advocated by country officials seeking to sustain political support for challenging reforms.

The discussions on growth dividends frequently emphasized the importance of the strength of program ownership and the corruption/governance problems in program countries. The case study of Latvia found that reforms were more likely to succeed if there was a strong motivating factor (e.g., EU accession), while the Romania case study showed the difficulties on making progress on state enterprise reform in the absence of consensus to support the reforms. The case studies on Grenada and Jamaica showed that program commitment can be effectively supported by energetic efforts to build broad domestic buy-in for difficult reforms, while recognizing that success of reforms hinged on many other domestic and external factors. The case studies on Honduras and Mongolia emphasized that favorable external conditions during the program period could mask insufficient reform efforts and/or diminish the incentives to implement and remain committed to reforms, with negative impact after the program ends and

when external conditions become less favorable. In Benin, Cameroon, and Senegal, the studies highlighted the need for a broader and comprehensive roadmap to improve governance, transparency, and anti-corruption efforts to benefit growth.

ASSESSMENT

Structural conditionality was extensively used to support reforms for both adjustment and growth objectives over the evaluation period. Generally, SCs were concentrated in the areas of the Fund's expertise, especially in the fiscal area, with relatively few conditions directly related to growth and efficiency. It is somewhat surprising that the share of SCs targeted directly to growth and efficiency was on average lower in PRGT programs than in GRA programs, although the share of growth objectives was on average higher in PRGT programs than in GRA programs.

Evidence presented in Chapter 4 and here suggests that well-implemented high depth, growth-oriented SCs advance reforms and bring growth benefits both during and after programs. However, the average quality of SCs was relatively low in terms of both depth and growth orientation in both GRA and PRGT programs, a situation that has been relatively stable over time with little signs of improvement.

These findings suggest that IMF-supported programs can and should do more to promote growth in program countries by strengthening the implementation, depth, and growth orientation of SCs. Greater focus on growth-oriented SCs may require the Fund to be more proactively involved in critical areas outside of its core expertise. In this respect, while recognizing that the setting, monitoring, and follow-up of SCs remain ultimately the full responsibility of the IMF, more effective collaboration with partner institutions could produce greater support for growth-enhancing reforms outside the IMF's core areas. In addition, given that higher-quality SCs take more time to implement, Fund arrangements of longer duration could allow for a more realistic time frame for reform implementation.

The high overlap in focus between IMF TA and SCs is encouraging, given that an ambitious reform agenda can stretch the available absorption capacity resulting in implementation delays. However, concerns arise about how well CD has supported program implementation

and monitoring. Cross-country data suggest that IMF TA may not have been delivered relatively more to countries with weaker capacity and that it has not been effective in strengthening SC implementation.[48] This suggests a need to consider further steps to more closely integrate programs and CD work, for example giving CD teams more of a role in designing and maintaining structural conditionality. In addition, implementation was significantly weaker for SCs outside of Fund expertise and for SCs relying explicitly on collaboration with partner institutions. Better targeting of TA resources and higher integration of TA with program implementation and monitoring could help increase traction for lasting changes in policy and institutions.

---

[48] The allocation of Fund CD resources has been guided by multiple considerations and not just country needs or capacity. The annual CD prioritization exercise reflects the membership's views on priorities for Fund work, individual members' requests for CD services, and Board decisions on the Fund's budget (IMF, 2019f). As such, there may be a trade-off between allocating CD resources to countries with the lowest capacity and allocating CD resources where it is likely to be effective. The upcoming IEO evaluation on "The IMF and Capacity Development" will take up these issues in greater detail.

EXCHANGE RATE POLICY, EXTERNAL ADJUSTMENT, AND GROWTH

This chapter examines how exchange rate policy in IMF-supported programs has helped to support growth and facilitate external adjustment.[49] It looks in turn at experience with discrete transitions in the overall exchange rate regime, changes in real exchange rates arising from both external and internal depreciations, and the links between such exchange rate outcomes and adjustment and growth outcomes. The analysis here recognizes the fundamental principles of the Fund's advice on exchange rate policy in both program and non-program contexts to respect the authorities' choice of the exchange rate regime (ERR) while seeking to ensure consistency with the broader macroeconomic policy framework and that any exchange rate depreciation should be consistent with members' obligation under Article IV to avoid manipulating exchange rates in order to prevent effective BOP adjustment or to gain an unfair competitive advantage.

In the face of an external shock, an exchange rate depreciation can help support growth and facilitate external adjustment by promoting expenditure switching and raising export profitability. Research looking into the aftermath of the GFC indicated that exchange rate adjustment may have helped to alleviate the adverse impact on growth during a crisis or when recovering from a crisis.[50] A separate strand of literature on the impact of exchange rate overvaluation or undervaluation on long-term growth finds that growth accelerations are often correlated with real exchange rate depreciations.[51] At the same time, however, exchange rate adjustment may be a double-edged sword, contributing to a "fear of floating," as it risks destabilizing inflation expectations and imposing adverse balance sheet effects on the financial sector depending on country circumstances and economic structure.

EXCHANGE RATE REGIME TRANSITION

Consistent with previous findings, transitions in the exchange rate regime (ERR) in the program context were relatively rare during the evaluation period.[52] Out of 131 programs in the evaluation sample, there were only 22 instances of regime transition (10 in GRA and 12 in PRGT programs) occurring in the program context (21 during the program period and one—Egypt (2016)—shortly before program approval as a prior action). Of these transitions, only 6 (5 GRA and 1 PRGT) were towards greater flexibility (all from intermediate to flexible

---

[49] This chapter draws on Bal Gündüz and Darius (2021) and country case studies prepared for the evaluation.

[50] Tsangarides (2012) and Terrones (2020).

[51] Hausmann and others (2005) and Rodrik (2008).

[52] This analysis uses a three-way regime classification of fixed, intermediate, and flexible regimes. This broad classification is based on the IMF's AREAER data in which regimes are classified into 10 categories (from 1 for no legal tender to 10 for free floating). The first three regimes in the AREAER (no legal tender, currency board, conventional peg) are mapped into the fixed regime, the final two categories (floating, free floating) are mapped into the flexible regime, and the rest is mapped into the intermediate regime. See Bal Gündüz and Darius (2021) for further technical details on the mapping.

regimes).[53] All transitions toward greater fixity were from flexible to intermediate regimes, except for two small states (São Tomé and Príncipe and Solomon Islands) which moved from intermediate to fixed regimes. Most program countries with no regime transition were under an intermediate regime.

The infrequency of regime transition is not unique to program periods. Transition probabilities estimated for a sample of 192 countries for 2008–19 are broadly similar between program and non-program periods and exhibit a strong tendency for status quo (Table 8, Panels A and B). A notable exception is the transition probability from flexible to intermediate regimes, which is two times higher in the program period (10.7 percent) than in the non-program period (5.3 percent). Accordingly, the probability of staying in a flexible regime is actually lower in the program period (89.3 percent) than in the non-program period (94.7 percent). Within program periods, the results show that regime transition towards greater flexibility was more likely in GRA programs than in PRGT programs, while the opposite was the case for transition towards greater fixity (Table 8, Panels C and D).

Data from a finer regime classification in the AREAER database also suggest a tendency toward greater fixity rather than greater flexibility in programs with no regime transition in broad categories. Specifically, the average regime score (scaled between 1 and 10 based on the AREAER classification) during the program period was lower (i.e., closer to fixity) than the corresponding score for the initial regime at T–1, particularly in PRGT programs that operated under intermediate regimes. This result suggests that a change in the exchange rate regime was not used actively as a tool to facilitate external adjustment and support growth in IMF-supported programs under evaluation, particularly in PRGT programs.

A simple analysis of growth outcomes for programs with different exchange rate regimes provides some evidence that more flexible regimes have been associated with somewhat stronger growth performance in PRGT programs. For such programs, growth was on average higher for countries under intermediate and flexible regimes than under fixed regimes (including currency union and currency board arrangements), while import compression was particularly pronounced in countries under the hardest peg (currency union, currency board) (Figure 32). In GRA programs, interestingly, growth during the program was on average highest under fixed regimes (Panel B) and lowest and negative under the hardest peg (Panel A).

TABLE 8. EXCHANGE RATE REGIME TRANSITION PROBABILITY
(In percent)

	A. PROGRAM PERIOD			B. NON-PROGRAM PERIOD		
	Fixed	Intermediate	Flexible	Fixed	Intermediate	Flexible
Fixed	100.0	0.0	0.0	98.2	1.1	0.7
Intermediate	2.1	93.8	4.1	1.1	94.4	4.5
Flexible	0.0	10.7	89.3	0.0	5.3	94.7
	C. GRA PROGRAMS			D. PRGT PROGRAMS		
	Fixed	Intermediate	Flexible	Fixed	Intermediate	Flexible
Fixed	100.0	0.0	0.0	100.0	0.0	0.0
Intermediate	0.0	89.7	10.3	2.9	95.6	1.5
Flexible	0.0	8.2	91.8	0.0	12.9	87.1

Source: Bal Gündüz and Darius (2021).
Note: Each entry in the matrix represents the probability of transition from the regime in the row to the regime in the corresponding column. "Non-program Period" covers the non-program periods of both program and non-program countries.

---

[53] These six instances are Armenia, Egypt, Georgia, Jamaica, Malawi, and Sri Lanka. However, Armenia, Egypt, Malawi, and Sri Lanka subsequently shifted back to intermediate regimes. Only Georgia and Jamaica still maintain a floating currency (since 2013 and 2017, respectively). Indeed, several countries had more than one regime transition during the program, leaving the number of programs with enduring regime transitions at 12 in the evaluation sample.

FIGURE 32. PROGRAM OUTCOMES BY EXCHANGE RATE REGIME
(Annual Averages)

A. Currency Union/Currency Board

B. Fixed Excluding Currency Union/Currency Board

C. Intermediate

D. Flexible

■ GRA   ■ PRGT
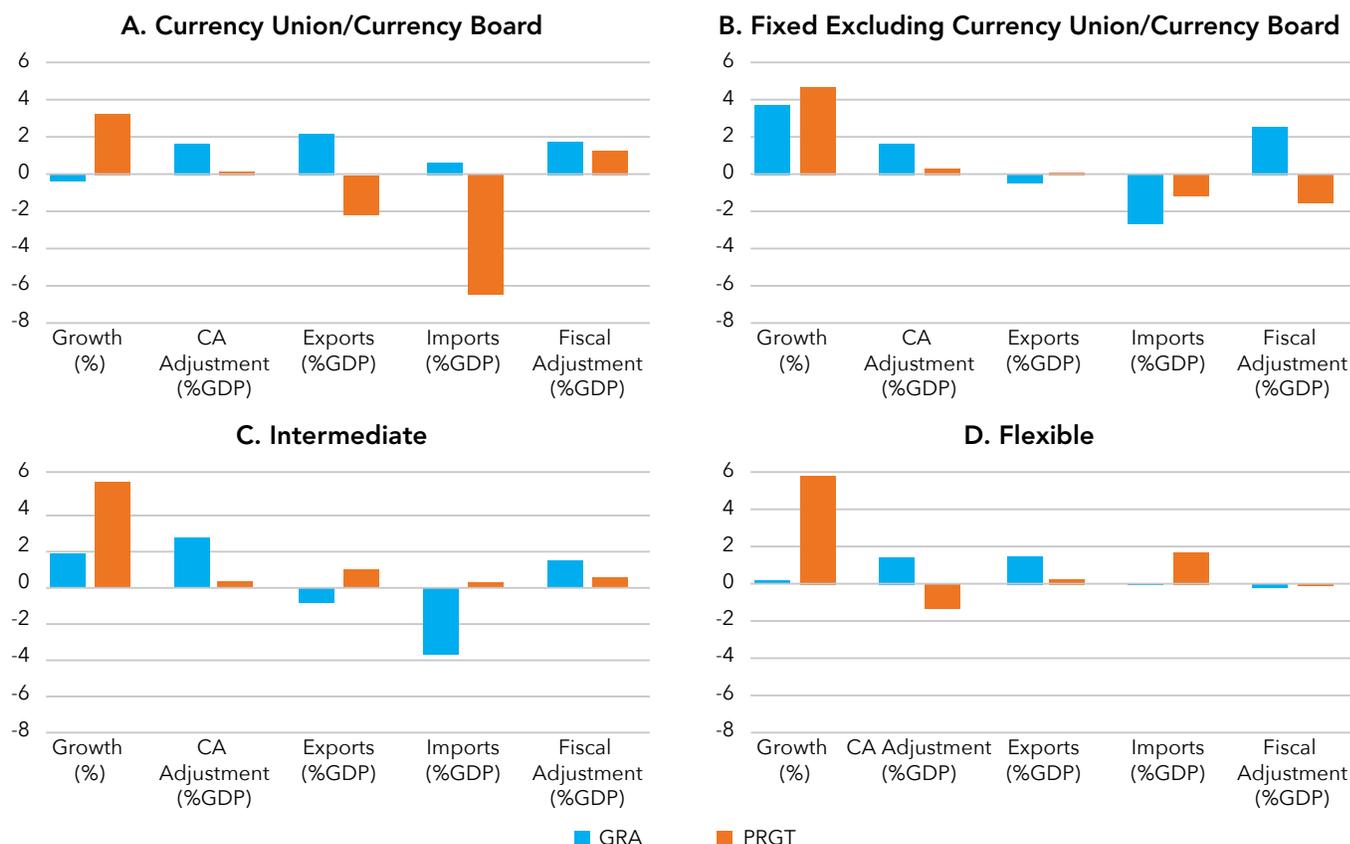
Source: Bal Gündüz and Darius (2021).
Note: Exchange rate regimes are based on the average regime score during the program period. The euro area programs are classified as programs under a currency union arrangement.

DEVELOPMENTS IN NOMINAL AND REAL EXCHANGE RATES

This section first looks at the higher-frequency developments in bilateral nominal exchange rates to capture the dynamics before and in the early phase of the program when movements are likely to be more pronounced than later in the program, particularly for countries facing sharp reversals in capital flows. It then examines how nominal exchange rate movements were translated into adjustments in nominal effective exchange rates (NEERs) and real effective exchange rates (REERs).

Developments in the Bilateral Nominal Exchange Rate

IMF-supported programs have often involved a significant nominal depreciation against the U.S. dollar. Specifically, the bilateral nominal exchange rate vis-à-vis the U.S. dollar (NER) depreciated by on average about 13 percent over the period between T–6 (6 months prior to program approval) and T+36 (36 months after program approval) in both GRA and PRGT programs (Figure 33). Cross-country variation was significantly larger in GRA programs than in PRGT programs, as indicated by the interquartile range in shade. It is notable, albeit not surprising, that NER depreciation prior to program approval was quite sharp for the bottom quartile of GRA programs, many of which were exceptional access and crisis programs. In contrast, for a quarter of GRA programs, the NER remained unchanged (pegged to

FIGURE 33. DEVELOPMENTS IN THE BILATERAL NOMINAL EXCHANGE RATE
(Increase = appreciation)

A. GRA Programs
(T-6 = 100)

B. PRGT Programs
(T-6 = 100)

Interquartile Range — Median
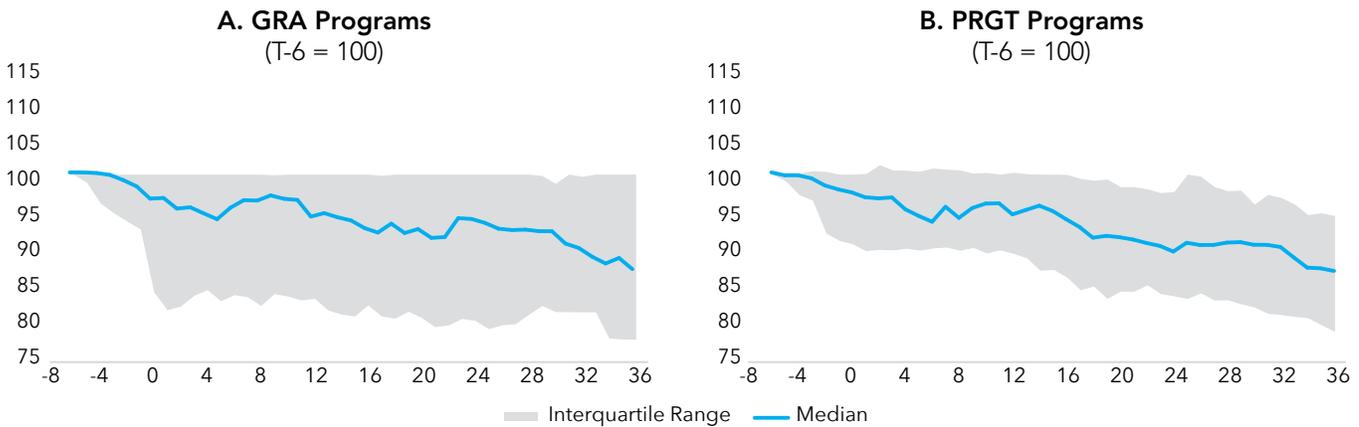
Source: Bal Gündüz and Darius (2021).
Note: Exchange rate against the U.S. dollar. The numbers on the horizontal axis represent months from program approval with negative numbers for pre-program period.

the U.S. dollar) or appreciated.[54] In PRGT programs, the NER depreciated by more than 5 percent by T+36 in more than three-quarters of programs. While individual country experiences differed widely, it is notable that the depreciation trend on average continued until 36 months after program approval in both GRA and PRGT programs.

Developments in the NEER and REER

Taking into account movements in third currencies, movements in the NEER were on average much more muted than those in the NER, although there is substantial cross-country variation, particularly in GRA programs (Figure 34). The median NEER depreciated by a mere 1 percent and 4 percent by T+36 in GRA and PRGT programs, respectively, while cross-country variation is substantially larger in the former. For the bottom quartile of GRA programs, the NEER depreciated by 18 percent or more by T+36.

Adjusting for inflation differentials, changes in the REER—a measure of exchange rate competitiveness—were typically even more modest. The median REER depreciated

by only 1.3 percent by T+36 in GRA programs with only limited cross-country variation and the median REER appreciated by 0.8 percent in PRGT programs. Notably, about a quarter of PRGT programs had a REER appreciation of 10 percent or more by T+36, attributed to higher inflation compared to trading partners. The REER appreciated by more than 5 percent in about a quarter of GRA programs over a similar period.

While movements in the NEER were only partially reflected in shifts in the REER due to inflation pass through, cross-section data do show a highly significant bivariate relationship between changes in the NEER and REER during the program period consistent with significant potential for exchange rate policy as a tool for external adjustment (Figure 35).[55] To be specific, in completed programs, a 1 percent depreciation in the NEER translated on average into about 0.6 percent depreciation in the REER not only in GRA programs but also in PRGT programs where exchange rate pass-through to inflation (ERPT) is more likely to be a policy concern than in GRA programs (Bal Gündüz and Darius, 2021). This result does not

---

[54] GRA program countries in the top quartile that had a U.S. dollar peg are Antigua and Barbuda, Djibouti, El Salvador, Grenada, Iraq, Jordan, and St. Kitts and Nevis.

[55] Exchange rate pass-through to inflation (ERPT) affects the extent to which nominal exchange rate changes translate into real exchange rate adjustments, and ERPT itself is influenced by the monetary policy regime (Taylor, 2000). Maintaining low and stable inflation reduces ERPT, which in turn helps to sustain low inflation and stabilize inflationary expectations. Evidence in the literature suggests that the ERPT in developing countries ranges between 0.1 and 0.5.

FIGURE 34. DEVELOPMENTS IN NEER AND REER

(Increase = appreciation)

A. NEER

GRA Programs
(T-6 = 100)

PRGT Programs
(T-6 = 100)

B. REER

GRA Programs
(T-6 = 100)

PRGT Programs
(T-6 = 100)

Interquartile Range — Median
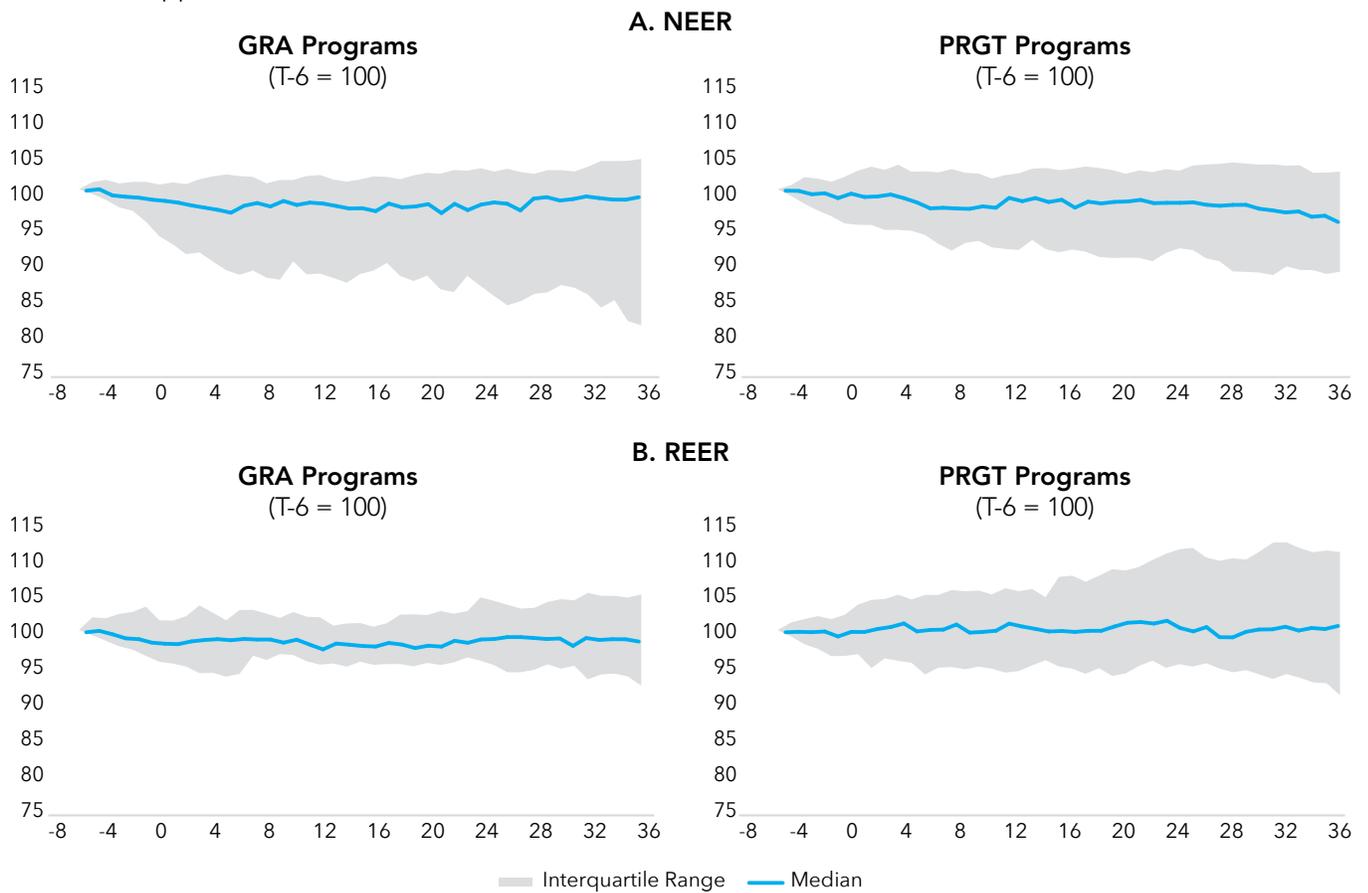
Source: Bal Gündüz and Darius (2021).
Note: The numbers on the horizontal axis represent months from program approval with negative numbers for pre-program period.

FIGURE 35. BIVARIATE RELATIONSHIP BETWEEN CHANGES IN THE NEER AND REER

All

$$\Delta REER = 0.594\Delta NEER + 1.886$$
$$R^2 = 0.62$$

GRA

$$\Delta REER = 0.562\Delta NEER + 0.587$$
$$R^2 = 0.83$$

PRGT

$$\Delta REER = 0.607\Delta NEER + 2.623$$
$$R^2 = 0.48$$
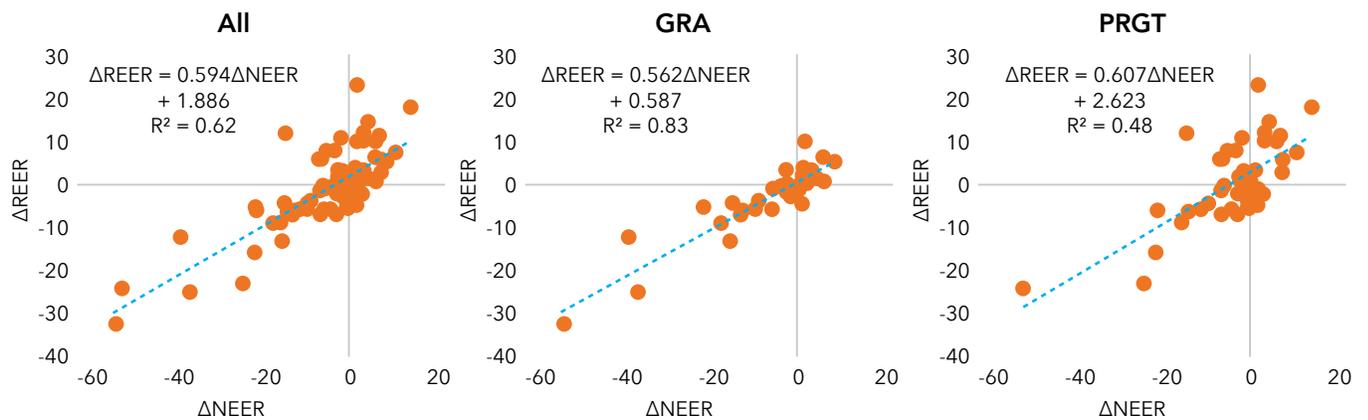
Sources: IMF, INS database; and IEO staff estimates.
Note: Changes in the NEER and REER are calculated as the percent change in the average during the program compared to the 12-month average before program approval.

hold, however, for incomplete programs, underlining the importance of ensuring confidence in the macroeconomic framework to contain ERPT.

Figure 36 provides a more detailed picture of the distribution of outcomes for the REER across GRA and PRGT programs. It confirms that REER depreciation was generally larger in GRA programs than in PRGT programs. Among PRGT programs, changes in the REER were more skewed to appreciation with a significant number of programs having REER appreciation above 5 percent (Figure 36, top panels). In contrast, a majority of GRA programs had some REER depreciation, which exceeded 4 percent for most crisis programs. Notably, REER movements towards depreciation were more limited for members of currency unions (Figure 36, bottom right panel), although three countries (Benin, Burkina Faso, and Senegal) did manage to achieve a REER correction of 5–7 percent through a combination of "internal devaluation" and NEER depreciation. A significant clustering of changes in the REER around zero suggests that in practice changes in exchange rate

competitiveness were not frequently achieved in programs to ease the adverse impact of adjustment on growth.

To put the magnitude of these exchange rate changes during programs into historical perspective, annual changes in the REER are measured relative to the country-specific standard deviation calculated over 2000–19. Focusing on the first year of the program during which exchange rate corrections were most likely, about one-fifth of GRA programs and a quarter of PRGT programs achieved a REER depreciation exceeding one standard deviation during the first program year (Figure 37, Panels A and B). As to changes in the NER, one-third of GRA programs and two-fifths of PRGT programs achieved nominal depreciation above one standard deviation (Figure 37, Panels C and D).

Notwithstanding that movements in the REER were generally quite muted during the program period, the exchange rate tool does appear to have been used to correct significant REER overvaluations prior to program approval. When assessed by the percentage deviation from the trend

FIGURE 36. CROSS-COUNTRY DISTRIBUTION OF CUMULATIVE REER CHANGES DURING PROGRAMS
(Share of programs; in percent)
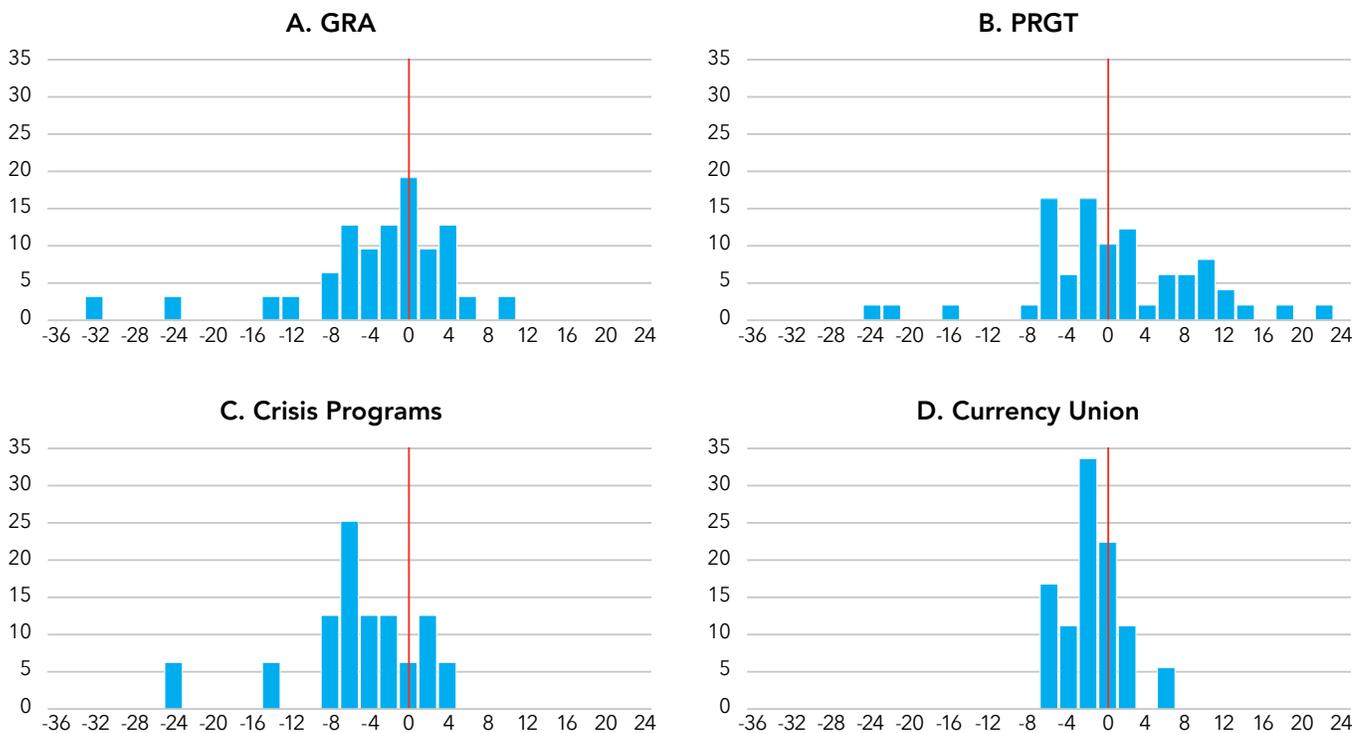
Source: Bal Gündüz and Darius (2021).
Note: Based on the percentage change in the program average (i.e., 36-month average after program approval relative to 12-month average before program approval). Positive (negative) values on the horizontal axis represent appreciation (depreciation).

REER, GRA programs had on average larger overvaluations immediately prior to program approval than PRGT programs (Figure 38). REER overvaluation exceeded 4 percent at T–1 for a quarter of GRA programs but in less than one-fifth of PRGT programs. Cross-country evidence suggests that initial REER overvaluation was corrected quite quickly and even reversed by the first year of the program as indicated by the black trend line for the full sample in Figure 39. In cases where initial overvaluation exceeded 4 percent, REER correction was on average two times larger than initial overvaluation as indicated by the red trend line.

REAL EXCHANGE RATES, EXTERNAL ADJUSTMENT, AND GROWTH

This section examines the role of REER depreciation in promoting adjustment and growth in the program context. Upfront, it should be recognized that the bivariate relationship between REER and adjustment and growth

masks the influence of a host of other factors and thus may not capture appropriately the marginal impact of REER depreciation on growth. Thus, the multivariate results presented subsequently are more reliable.

Looking first at adjustment, while bivariate results suggest only a weak impact of REER depreciation on the current account, a multivariate analysis which controls for other factors that affect CA adjustment paints a more encouraging picture, particularly for GRA countries. Specifically, bivariate data for GRA programs suggest that REER depreciation appears to have promoted not only exports but also imports, resulting in limited contribution of REER depreciation to CA adjustment (Figure 40, left panels). The bivariate relationships between REER depreciation and CA adjustment or adjustments in exports and imports in PRGT programs are broadly similar to those in GRA programs with limited contribution to CA adjustment (Figure 40, right panels). Nevertheless, the multivariate regression analysis reported in Bal Gündüz and Darius (2021), which takes account of adjustment policies, shows

FIGURE 37. CROSS-COUNTRY DISTRIBUTION OF CHANGES IN REER AND NER: FIRST YEAR OF THE PROGRAM
(Positive figures indicate appreciation)
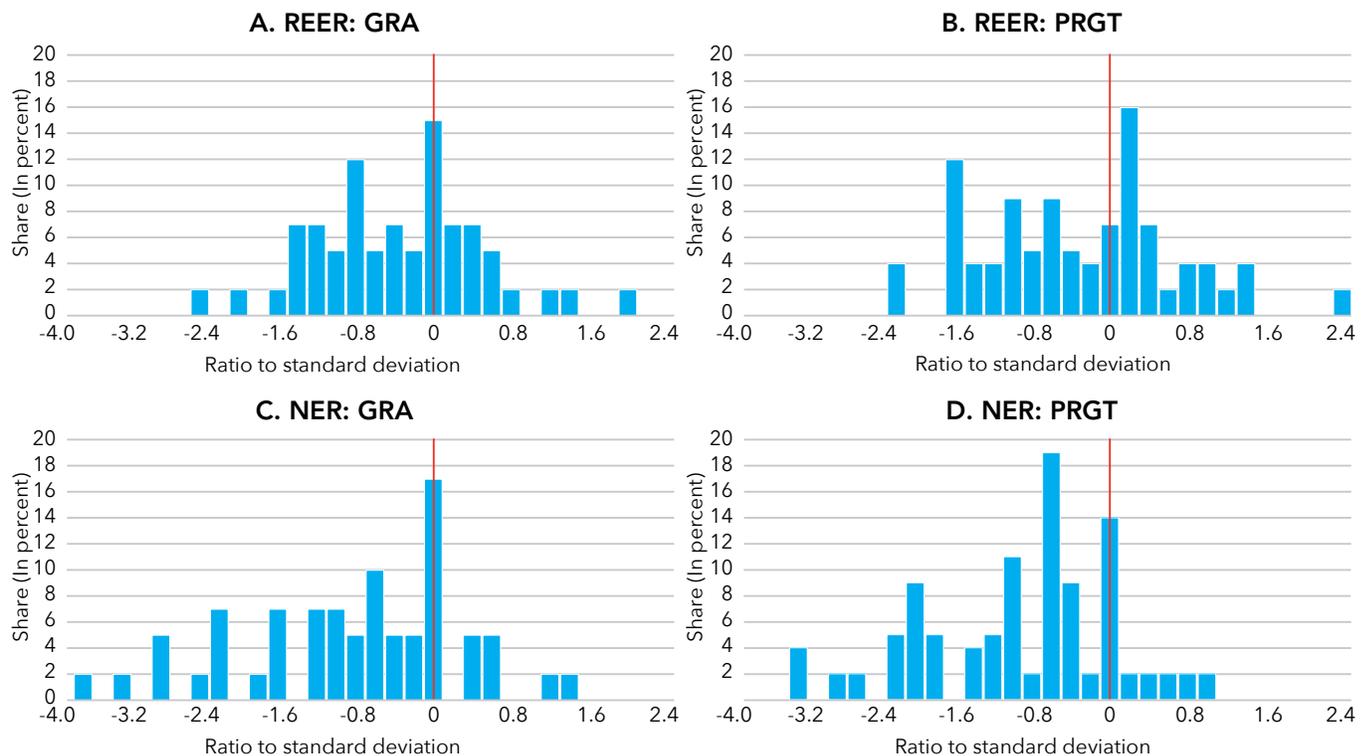
Source: Bal Gündüz and Darius (2021).
Note: Based on annual data. Standard deviations are country-specific and calculated by using the annual data over 2000–19 for each country.

FIGURE 38. PERCENTAGE DEVIATION FROM TREND REER
(Positive figures indicate appreciation)

A. GRA

B. PRGT

Interquartile Range — Median
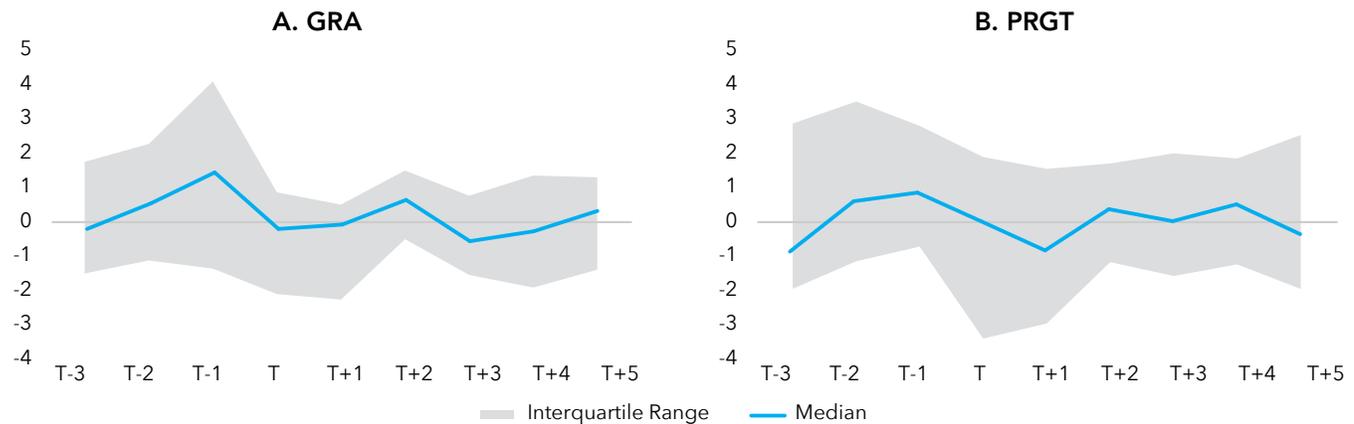
Source: Bal Gündüz and Darius (2021).
Note: The trend REER is estimated by applying the HP filter to annual REER data for each country. T refers to the first year of the program.

FIGURE 39. REER REACTION TO PRE-PROGRAM OVERVALUATION
(Positive figures indicate appreciation or overvaluation)

$y = -0.9897x - 0.8188$
$R^2 = 0.3505$

$y = -2.1621x + 7.7558$
$R^2 = 0.4352$

ΔREER at T (percent)

REER overvaluation at T-1 (percentage points)

GRA   PRGT
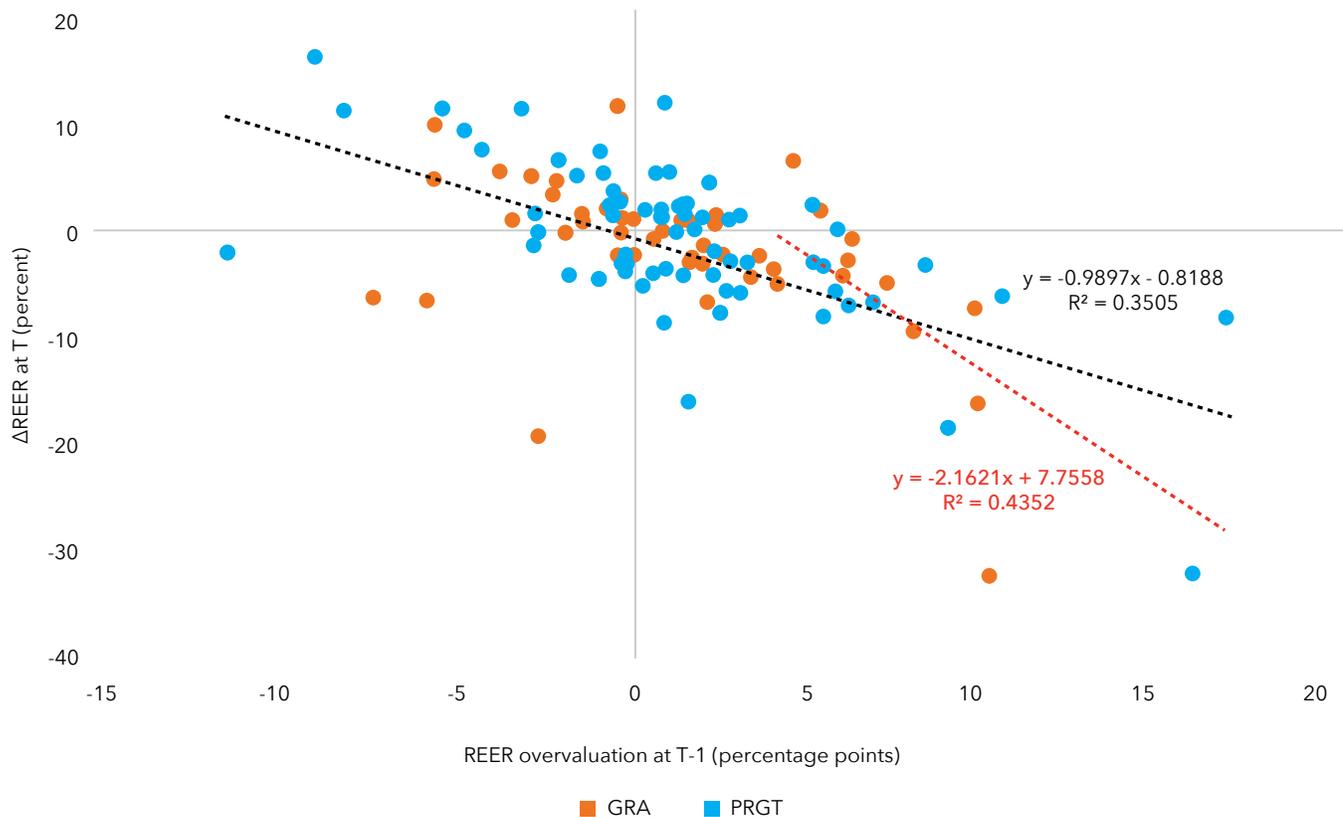
Source: Bal Gündüz and Darius (2021).
Note: REER overvaluation is measured by the percentage deviation of the actual REER from the historical trend. The linear trend line in red is for REER overvaluations exceeding 4 percentage points only.

FIGURE 40. EXTERNAL ADJUSTMENT AND CHANGE IN REER

GRA

$y = -0.0617x + 1.733$
$R^2 = 0.0053$

$y = -0.4149x + 0.53$
$R^2 = 0.1133$

$y = -0.4709x - 1.3379$
$R^2 = 0.1433$

PRGT

$y = -0.0361x + 0.4499$
$R^2 = 0.0026$

$y = -0.1328x + 0.0676$
$R^2 = 0.05$

$y = -0.0587x - 0.6727$
$R^2 = 0.0033$

■ Crisis Programs   ■ Other Programs
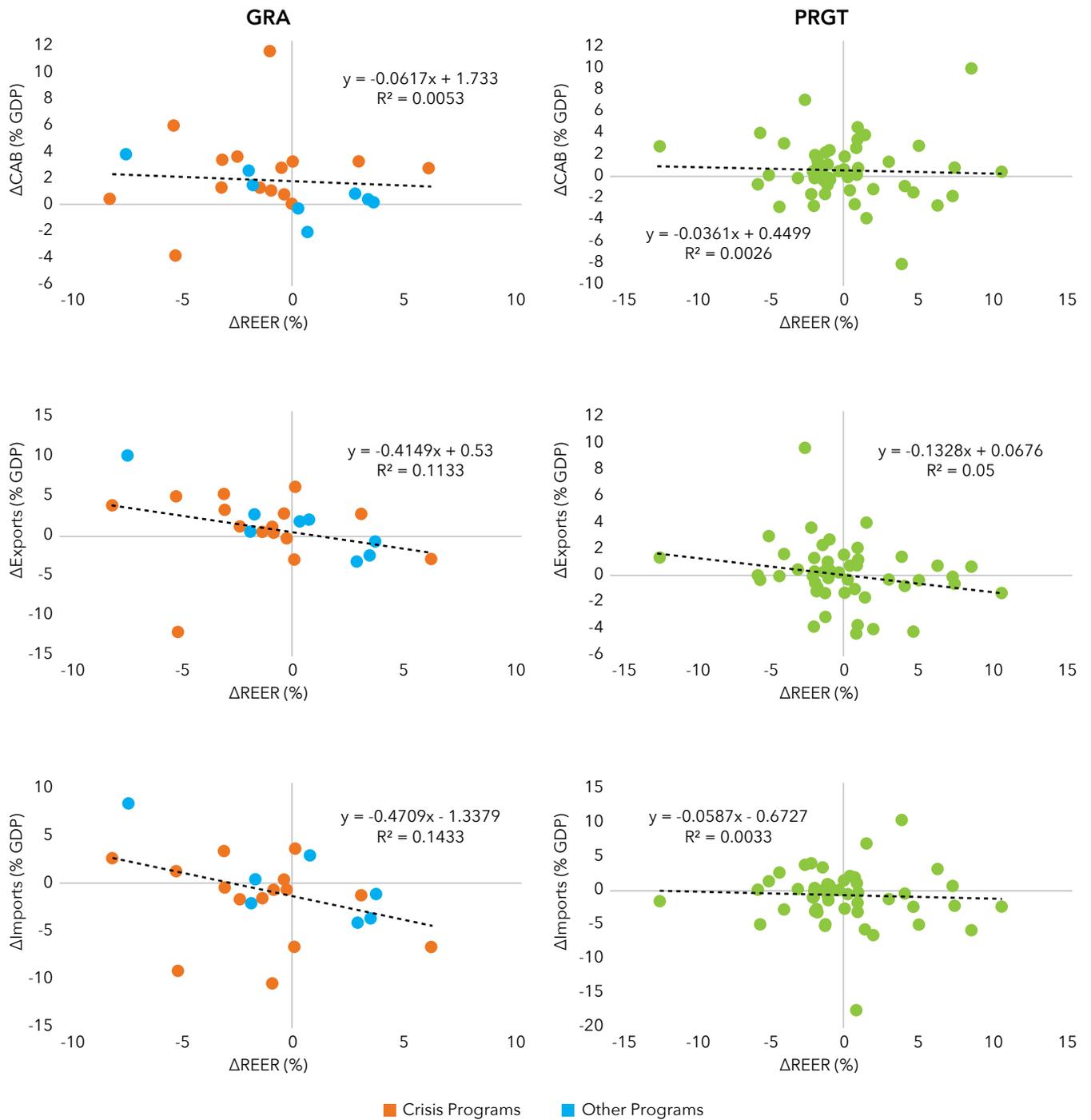
Source: Bal Gündüz and Darius (2021).
Note: CAB stands for the current account balance as a share of GDP.

that for program countries (excluding small states) over the evaluation period, the effect of REER changes on CA adjustment is highly significant with the expected sign for GRA countries, although not significant for PRGT countries. The regression results suggest that a 10 percent REER depreciation helped to increase the CA balance in the order of 1.0–2.6 percent of GDP depending on the level of trade openness (with the effect being larger for higher trade openness).

Turning to growth, the bivariate relationship between REER changes and growth is generally weak in GRA programs where more REER depreciation is associated with lower, and not higher, growth. However, this result is largely influenced by the relationship observed in crisis programs which typically experienced the largest changes in REER as well as less favorable growth outcomes in view of large adjustment needs (Figure 41, left panel).[56] In contrast, larger REER depreciation is on average associated with higher growth in PRGT programs (Figure 41, right panel).

Again, multivariate regression analysis, which takes account of adjustment policies, found clearer evidence of benefits from REER changes (Kim and others, 2021). In particular, this work found that where it has occurred, real effective depreciation (DREER <0) has helped to boost growth, particularly in PRGT programs where a 10 percent

real depreciation is estimated to boost growth by 1.1–1.7 percentage points. For GRA programs, the same 10 percent real depreciation is estimated to boost growth by 0.4–2.9 percentage points although the impact is generally not statistically significant.

LESSONS FROM COUNTRY EXPERIENCE

Consistent with the empirical analysis, only a minority of the case study countries experienced significant real depreciations during the evaluation period. Among the completed programs, six countries achieved real exchange rate depreciations exceeding 5 percent in cumulative terms during the program period (Figure 42). Egypt (2016) and Malawi (2012) had the highest real depreciations although also experiencing sharp spikes in inflation. In the former case, exchange rate management was complicated by a surge in capital inflows at a depreciated exchange rate, while in the latter fiscal and monetary policies were not tightened sufficiently. Significant nominal depreciation under Ghana's floating regime translated into relatively modest real depreciation owing to high inflation in the 2009 program. Benin (2010) and Senegal (2010), both of which are members of the West African Economic and Monetary Union (WAEMU) and share a currency pegged to the euro, achieved real depreciation through a combination of

FIGURE 41. GROWTH OUTCOMES BY CHANGE IN THE REER: COMPLETED PROGRAMS
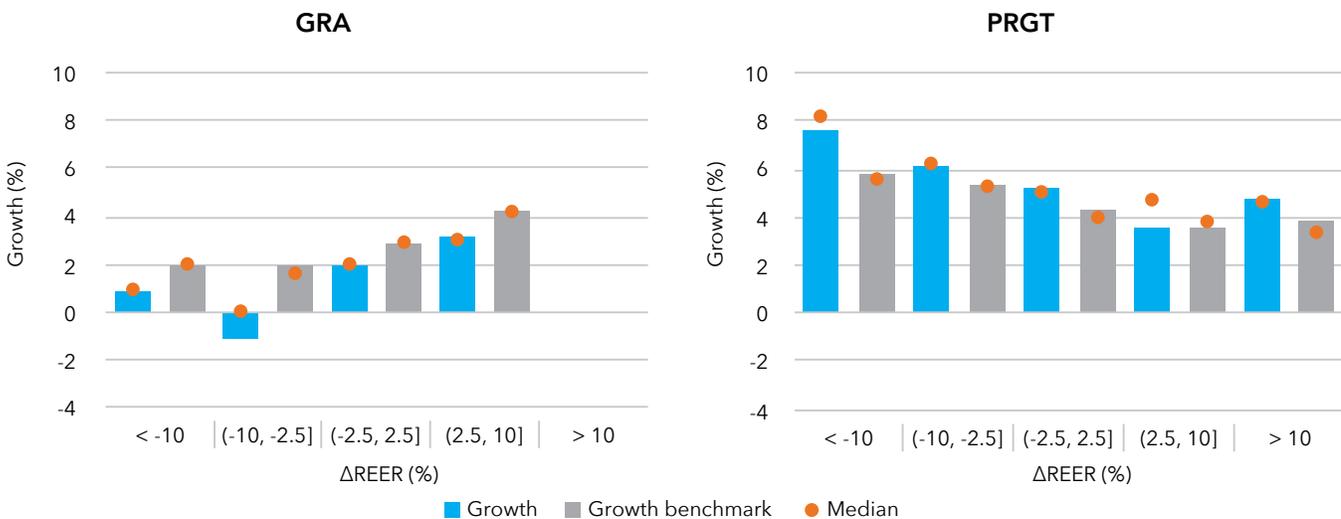
Source: Bal Gündüz and Darius (2021).

---

[56] Bal Gündüz and Darius (2021) reports that in other GRA programs, REER depreciation of more than 10 percent was associated with significantly higher growth than the benchmark.

FIGURE 42. PROGRAMS IN CASE STUDIES WITH REER DEPRECIATION OF 5 PERCENT OR MORE
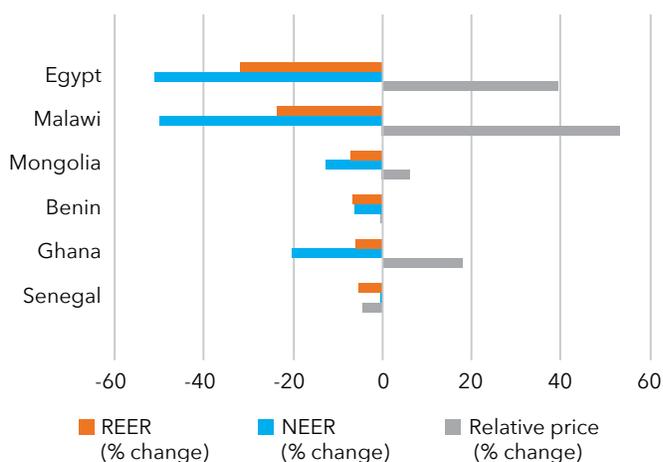
Source: Bal Gündüz and Darius (2021).

relative price deflation and depreciation of the euro against third currencies.[57]

The experience with adjustment and growth outcomes varied widely across these six cases. Egypt, Ghana, and Senegal achieved substantial current account adjustment. In Egypt, exports responded strongly to real depreciation as the program addressed upfront the significant overvaluation and the disruptive shortages of foreign exchange that had crippled the economy, which helped to support Egypt's strong growth performance during the program. In Senegal, export performance also benefitted from the real depreciation but competitiveness problems (including non-price factors) remained an issue holding back private sector performance. While Ghana's exports increased significantly following its depreciation, external imbalances reemerged at the end of the program owing to highly expansionary fiscal policy, which led to a successor arrangement. In Malawi (2012), the envisaged effects of exchange rate regime reforms on diversification did not materialize given non-price impediments. Mongolia's export sector is dominated by the mining sector, so export performance has depended more on international

commodity prices and supply factors; flexible exchange rate management in the 2009 program was intended as part of the macroeconomic policy framework rather than to gain export competitiveness.

The experience of countries with currency pegs shows that some competitiveness gains can be achieved with appropriate domestic policies while maintaining the peg. Latvia (2008) is a case in point. Latvia faced an estimated 30 percent overvaluation and in the program's first year, GDP declined by 14 percent, triple the projected contraction. While maintaining its currency peg to the euro, Latvia restored external competitiveness unexpectedly quickly, through a surge in labor productivity rather than a decline in domestic wages and prices, and exports recovered strongly (helped partly by the contraction in domestic demand). However, Latvia's experience seems to be historically unusual.[58] As mentioned above, two WAEMU members (Benin and Senegal) also experienced real effective depreciations albeit in part because of the decline of the euro against the dollar, and Benin in particular benefitted from improved export performance in the context of efforts to improve transportation infrastructure. By contrast, Cameroon, which is a member of the Central African Economic and Monetary Community (CEMAC), entered a program in 2017 alongside all other CEMAC members in the face of a large decline in oil prices and unsustainable fiscal policies, but was not able to achieve much improvement in competitiveness to support non-oil exports.

Several country case studies illustrate the practical difficulties involved in managing and sustaining a shift to a more flexible exchange rate regime. The Egypt (2016) program envisaged moving to a fully floating regime and had a prior action on an upfront devaluation but shifted back to an intermediate regime in 2017, preferring to limit wide exchange rate fluctuations in the face of volatile capital flows. Malawi (2008, 2010) aimed to introduce greater exchange rate flexibility to eliminate parallel market premia. Ahead of the 2012 program, the authorities adopted

---

[57] The depreciation of the euro against the U.S. dollar contributed to NEER depreciation in Benin but far less in Senegal due to different composition of trading partners.

[58] In many crisis programs, internal devaluation itself proved hard to achieve and the desired recovery in growth and exports did not materialize (IMF, 2015b; IEO, 2016). Difficulties in restoring competitiveness through internal devaluation was confirmed by the ex-post assessment for Greece's 2010 program (IMF, 2013c) and Portugal's 2011 program (IMF, 2016). More recent evidence also suggests that the output costs of external adjustment via internal devaluation were higher than anticipated in some euro area countries during 2010–14 (Lambertini and Proebsting, 2019).

a floating rate regime, but the float proved highly volatile in the face of swings in demand for foreign exchange, and after a high inflation episode, Malawi moved back to an intermediate regime in 2016. Mongolia (2009) had a prior action on the introduction of foreign exchange auctions. The return of commodity boom-bust cycles led to another program in 2017 amid a sharp depreciation of the exchange rate and reserve losses. In fear of worsening debt dynamics and balance sheet risks, Mongolia shifted back to an intermediate regime in 2018. Even in cases that successfully transitioned to a floating regime, concerns about the consequences of floating delayed the regime change. In Jamaica (2010), for example, the authorities initially pushed against Fund advice for greater flexibility due to concerns for pass-through to inflation, lack of national support, and delayed responsiveness of exports to the exchange rate. Jamaica (2013) included commitments to exchange rate flexibility, measures to develop fully the interbank foreign exchange market, and further steps towards a full-fledged inflation targeting regime. Helped by these measures, Jamaica eventually shifted to a floating regime in 2017 during the 2016 successor program, and has seen a pickup in export growth. Ukraine moved to a more flexible exchange rate regime in 2009 which aimed at reducing reliance on capital controls, returned to a peg in 2011, and then floated the rate in 2014 in the face of heavy market pressures in a highly uncertain policy environment. Although Ukraine needed to introduce capital controls as security-related uncertainties intensified in 2015, it has subsequently been able to maintain a flexible rate regime backed by introduction of inflation targeting.

In a number of cases with intermediate exchange rate regimes, programs tolerated gradual real effective appreciation in the context of stabilization efforts, notwithstanding the staff's assessment of overvaluation at program approval or earlier.[59] Honduras (2014) introduced a crawling peg in the 2010 SBA/SCF to provide more scope to protect competitiveness after considerable appreciation under the previous peg, but the REER continued to rise albeit at a reduced rate. In Pakistan (2013), the exchange rate was managed heavily as a contribution to inflation control in the 2008 SBA and 2013 EFF, allowing a gradual appreciation of the real effective exchange rate; the loss of competitiveness was

eventually reversed by a sharp market driven depreciation ahead of the 2019 EFF.

The case study experience clearly indicates that a shift towards active use of a flexible exchange rate as a policy tool needs careful technical preparation and policy support—in particular to establish a liquid foreign exchange market, a credible monetary policy regime to provide an alternative inflation anchor to limit exchange rate pass through to inflation and consistent macroeconomic policy settings. In all the cases with regime shifts, the Fund played a significant role in supporting more flexible exchange rate management through its capacity development work related to exchange rate management, market development, and the monetary policy framework, which was well appreciated. However, success clearly depended on strong buy-in and commitment from the country authorities.

ASSESSMENT

The experience assessed in this chapter suggests that greater exchange rate flexibility did help to improve competitiveness and support external adjustment and growth in a number of programs, particularly in GRA cases involving significant initial currency overvaluation. However, shifts in exchange rate regime and substantial depreciations in the real exchange rate were relatively infrequent, and were often not sustained post program. In a number of cases, the gravitational pull towards intermediate regimes reasserted itself after short episodes of floating, reflecting at least in part fear of floating. Moreover, a number of PRGT programs relying on heavily managed exchange regimes as an inflation anchor experienced quite significant real exchange rate appreciations, suggesting an over-reliance on the exchange rate as a tool to control inflation with adverse implications for growth.

This experience suggests that the exchange rate tool could be used more actively as a means to support growth outcomes while achieving external adjustment, subject to the principle that the exchange rate regime choice is ultimately the authorities' decision and to members' obligations under Article IV to avoid manipulating exchange rates in order to prevent effective BOP adjustment or to

---

[59] The staff's exchange rate assessments are taken from the IMF country reports at program approval or the latest Article IV consultation report prior to program approval.

gain an unfair competitive advantage. This said, the ground for more active use of the exchange rate would need to be well prepared. Policies to contain the inflationary impact of depreciation and a monetary policy framework that anchors inflation expectations are required to ensure the effectiveness of the exchange rate as a policy tool to improve competitiveness, especially in PRGT programs. Where they have worked well, they have followed various paths depending on country circumstances—for example, regime transition towards greater flexibility; more aggressive use of flexibility available under intermediate regime; and internal devaluation. To address obstacles to greater use of the exchange rate tool, there needs to be early attention to: (i) technical issues and related TA support to build functioning foreign exchange markets; (ii) the need to ensure adequate anchors (e.g., build a credible inflation targeting monetary policy regime) to limit exchange rate pass through to prices; and (iii) reducing non-price obstacles to improved export performance and currency mismatches in the financial sector balance sheets.

MARKET DEBT OPERATIONS AND GROWTH

This chapter examines the experience with *market* debt operations undertaken in IMF-supported programs over the evaluation period.[60] After providing an overview of these debt operations, it examines how programs with debt operations have fared in terms of growth and adjustment during and after the program. The assessment focuses primarily on 12 market debt operations, most of which were in the context of GRA programs.

MARKET DEBT OPERATIONS IN IMF-SUPPORTED PROGRAMS

The basic principle underlying the IMF's role in supporting debt operations is that it is for the member country to decide whether and how to restructure its debt and to manage the whole restructuring process. At the same time, the IMF must ensure that any IMF-supported program can successfully achieve its objectives while safeguarding the revolving character of IMF resources, which requires a satisfactory judgment on debt sustainability as a basic prerequisite for access. The financing assurances policy requires that a program should have adequate external financing and, in cases where the financing gap cannot be filled by other means, explicitly encourages debt restructuring operations on terms compatible with balance of payments viability (IMF, 2013b).

The lending into arrears policies (LIA/LIOA) require that a country under an IMF-supported program be making good faith efforts to negotiate the restructuring of the debt in default with its private or official creditors. These policies are supported by two carefully developed frameworks for debt sustainability analysis, for market access countries (MAC DSA) and for LICs (LIC-DSF), the latter prepared jointly with the World Bank. The LIC-DSF was last modified in 2017 and the MAC DSA in early 2021.[61] Under these policies, where debt is assessed as unsustainable or even sustainable but not with high probability (in exceptional access cases), debt operations have been required as a condition for access to Fund resources. Beyond this requirement, debt operations can contribute to support long-term growth by relieving the burden of future debt service, making more fiscal resources available for productive public investment, and improving incentives for private investors.

Following this framework, the IMF has played an active role in supporting market debt operations while adhering to the neutrality principle. The IMF's financing assurance and lending into arrears policies as well as the Fund's catalytic role have provided balanced incentives for debt restructuring on both debtor and creditor sides. The IMF has also provided technical advice in identifying resources available for debt servicing under

---

[60] This chapter draws on Erce (2021) and country case studies prepared for the evaluation.

[61] The recent modifications of the MAC DSA framework expanded the battery of analytical tools to increase the robustness of sovereign risk analysis with broader debt coverage, improve the framework's capacity to predict sovereign stress, and enhance transparency in exercising judgment (IMF, 2021).

alternative scenarios and assessing the restructuring envelope, primarily based on the DSA framework. In addition, Fund staff has often played a role in encouraging creditor participation by communicating with stakeholders about a country's economic policies and prospects and debt servicing potential.

The Fund's DSA framework has been used extensively in the program context to make difficult judgments on whether to insist on debt operations as a condition for access to IMF financing and whether the extent of debt relief in packages under negotiation would be sufficient to achieve debt sustainability. In some cases (e.g., Belize), the Fund's DSA has also played a crucial role in the surveillance context by serving as an independent baseline for negotiations between authorities and creditors.

While the MAC DSA is a highly sophisticated framework that has evolved over time, it does not include a systematic analysis of the impact of debt restructuring on growth and the timing of market access. While DSA was used to evaluate the effect of different debt operations on the gross financing need and debt path, the endogenous linkage between debt and growth was not always clearly grounded. In practice, Fund staff generally took an ad hoc approach to assess market access and the debt-growth nexus used in program design, typically embedding the effects of debt operations in the form of lower primary balance and reduced interest payments.[62] In some cases, growth projections were adjusted to reflect the design of the debt operation, particularly the extent to which the burden of debt restructuring falls on the domestic creditors. This lack of firm foundations raises concerns since unless a program provides a path for a country to regain market access, the IMF does not have a strong basis to conclude that the program is addressing the underlying problems (Hagan and others, 2017; Guscina and others, 2017).

EXPERIENCE WITH DEBT OPERATIONS

Although limited in numbers, the 12 market debt operations in the evaluation period have been diverse in terms of modality and coverage of debt, reflecting a variety of country-specific factors related to creditor participation, cross-border spillovers, domestic financial stability, and social consequences. Preserving the health of the financial sector received significant attention when debt operations involved debt owed to domestic financial institutions. For financial stability concerns and other strategic reasons (e.g., political concern on distributional consequences for domestic creditors), differential treatment of creditors and debt instruments was common. As arrears existed in all cases and were very large in some cases, program conditionality generally targeted the clearance of arrears as part of overall debt restructuring.

Experience shows that while completed debt operations ultimately brought significant debt relief, negotiations of debt restructuring packages can adversely affect credit availability and confidence, with an adverse short-term impact on growth. Negotiations can be particularly disruptive if extended by technical difficulties in reaching agreement with creditors, litigation, or political concerns (IMF, 2020a).[63] The increasingly heterogeneous creditor base has also complicated efforts to proceed promptly with debt negotiations. Prolonged litigation made the output costs of default felt for a longer period (e.g., in Antigua and Barbuda), while a confrontational stance with holdout creditors could compromise a country's ability to raise new financing in international capital markets. In some cases (e.g., Grenada), capacity constraints limited authorities' ability to develop and implement a debt restructuring strategy, delaying debt operations longer than expected. Most debt operations in the evaluation period involved default and accumulation of arrears with external creditors, adding to risks of delayed debt operations.

In this context, various legal approaches were used to facilitate creditor participation, limit holdouts, and avoid litigation. To increase creditor participation, for example,

---

[62] A notable exception in this regard is Jamaica (2013) where staff's analysis of direct and indirect (through the exchange rate) effects of the debt operation on growth contributed to a milder restructuring as being judged to be sufficient to restore debt sustainability.

[63] IMF (2020a) indicates that the increasingly diverse creditor base and debt instruments (especially collateralized debt) can complicate and lengthen the process of debt restructuring. Trebesch (2019) suggests that political instability, weak institutions, and strategic government behavior influence delays in completing restructurings more than creditor characteristics.

innovative contractual design and instruments were used including value recovery instruments and countercyclical and state contingent payouts (IMF, 2020b). In some cases, an aggregate collective action clause was retrofitted for restructuring of local-law debt (e.g., Barbados and Greece). In other cases, authorities modified financial regulations to make debt restructuring more palatable (e.g., Jamaica and St. Kitts and Nevis).

Experience with regaining market access was mixed. Greece, Grenada, and Jamaica all took longer than expected to regain market access, but for Côte d'Ivoire, Cyprus, and Ukraine regaining access was faster than expected. Despite successful implementation of debt restructuring, Grenada was judged to remain in debt distress due to the non-completion of all debt restructurings and the existence of arrears to official creditors, which delayed market access. In Jamaica (2010), the limited debt relief created a repayment bunching, resulting in delayed market access and leading to the need for a subsequent operation.

ADJUSTMENT AND GROWTH OUTCOMES IN PROGRAMS WITH DEBT OPERATIONS

The time profile of adjustment and growth was quite different between GRA programs with and without debt operations during the evaluation period.[64] Debt outcomes of the programs with debt operations were on average better than projected and those of other GRA programs (Figure 43, Panel C). While initial debt was far higher in programs with debt operations, debt ratios were on average put on a broadly declining trend while the opposite was the case for other GRA programs. Success in putting debt on a declining path in part reflected that in programs with debt operations, fiscal adjustment (measured as the change in the primary balance) was stronger and more front-loaded than in other GRA programs (Figure 43, Panel B).

As to the growth trajectory, growth on average rebounded sharply from a deep trough at T–1 in programs with debt operations, with the U-shaped pattern being much sharper than that of other GRA programs in both growth

projections and outcomes (Figure 43, Panel A). Growth outcomes of the programs with debt operations on average slightly underperformed initial projections in early years of the program (T and T+1) but exceeded projections in later years. As with debt and fiscal outcomes, growth outcomes ranged widely across programs with debt operations as indicated by the interquartile range in shade, which is significantly wider than that for other GRA programs.

Overall, IMF-supported programs with market debt operations had only mixed success in terms of debt sustainability and the BOP position after the program. Half were followed by successor programs. And in half the cases, either follow-on debt operations were needed or debt ended up in distress or at risk according to the DSA. Where the debt operations fell short of restoring debt sustainability, the underlying reasons varied, including insufficient debt relief, lack of technical expertise, shortfalls in fiscal adjustment, political pressures and, in some cases, the discovery of previously undisclosed debts.

In broad terms, debt operations with principal haircuts and upfront fiscal adjustment were more successful in reducing debt than those with just debt reprofiling or lowered coupons. Specifically, programs with debt operations were able to reduce debt on average by 14 percent of GDP over the 3-year horizon following program approval, with debt reduction of 21 percent of GDP in programs with principal-based operations compared to an average increase in debt of 3 percent of GDP in programs with debt reprofiling, although this is an unfair comparison because it does not take into account the impact of coupon reduction on the net present value of the debt (Figure 44).

Turning to growth outcomes, the evidence suggests that more effective debt operations have on average been associated with better growth outturns in both program and post-program periods (Figure 45). Specifically, both within-program and post-program growth outcomes (during the five years after the program ended) relative to the growth benchmark discussed in Chapter 3 have been superior for: (i) operations based on principal reduction rather than

---

[64] The analysis in this section is based on the data for 10 programs with debt operations, excluding Barbados (2018) which is an ongoing program and The Gambia (2017) where GDP rebasing in 2018 affected actual debt ratios significantly and thus skewed their comparison with program projections. Given the small sample size, evidence on adjustment and growth outcomes of the programs with debt operations could be sensitive to idiosyncratic outliers.

FIGURE 43. ADJUSTMENT AND GROWTH TRAJECTORIES: PROGRAMS WITH AND WITHOUT DEBT OPERATIONS
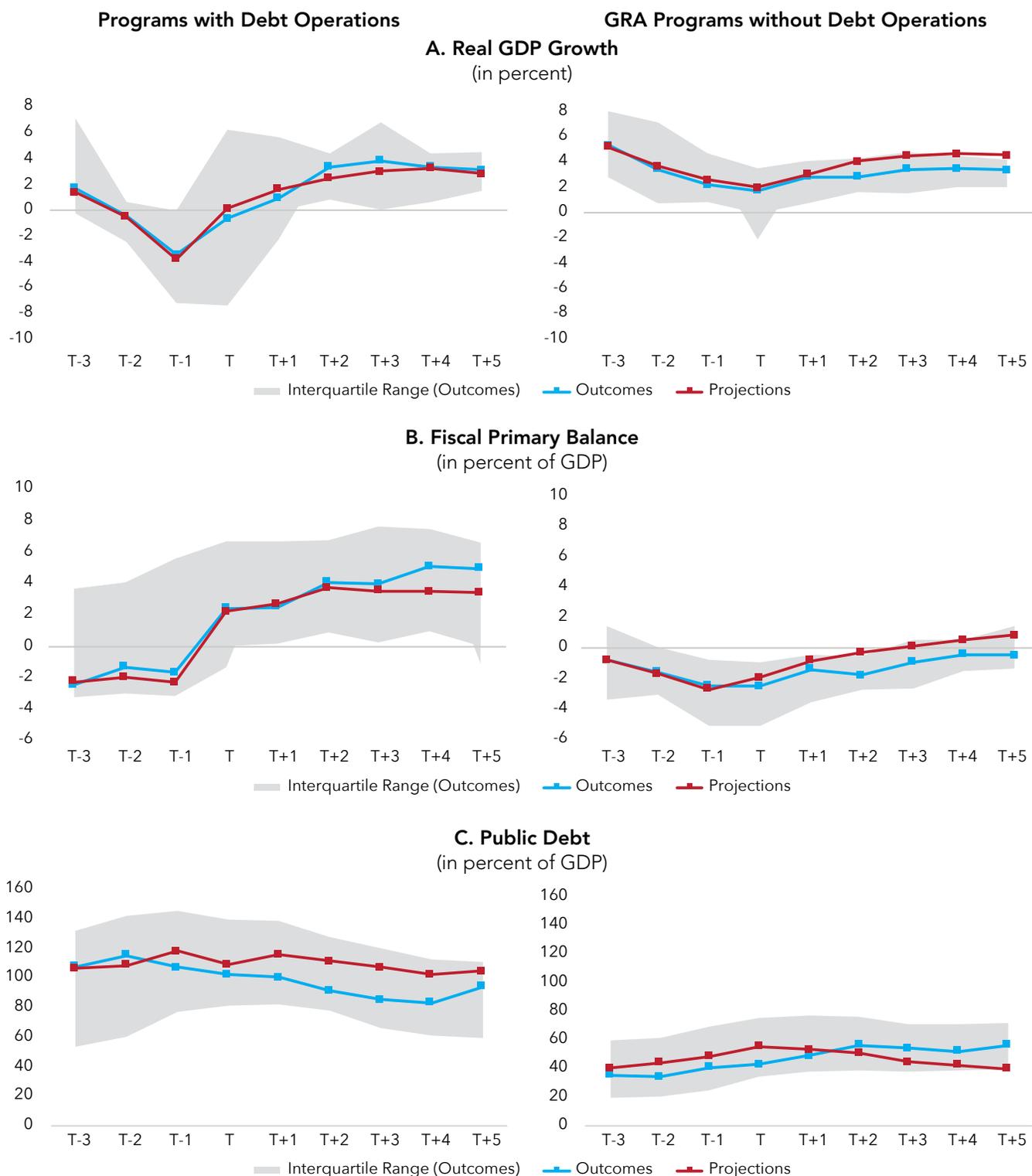
Programs with Debt Operations **GRA Programs without Debt Operations**

A. Real GDP Growth
(in percent)

B. Fiscal Primary Balance
(in percent of GDP)

C. Public Debt
(in percent of GDP)

Source: Erce (2021).
Note: Based on the data for 10 programs with market debt operations, excluding Barbados (2018) and The Gambia (2017), and 42 GRA programs without debt operations. Outcomes and projections represent cross-country medians. All projections are initial projections at program approval (T). Data availability is not uniform across periods. Due to the presence of successor programs for some countries in the sample, there is overlap in the data presented over the period and, therefore, the results are not always fully consistent with those based on program periods only.

FIGURE 44. DEBT OUTCOMES OF PROGRAMS WITH DEBT OPERATIONS

(Three-year cumulative change in debt from program approval; in percent of GDP)
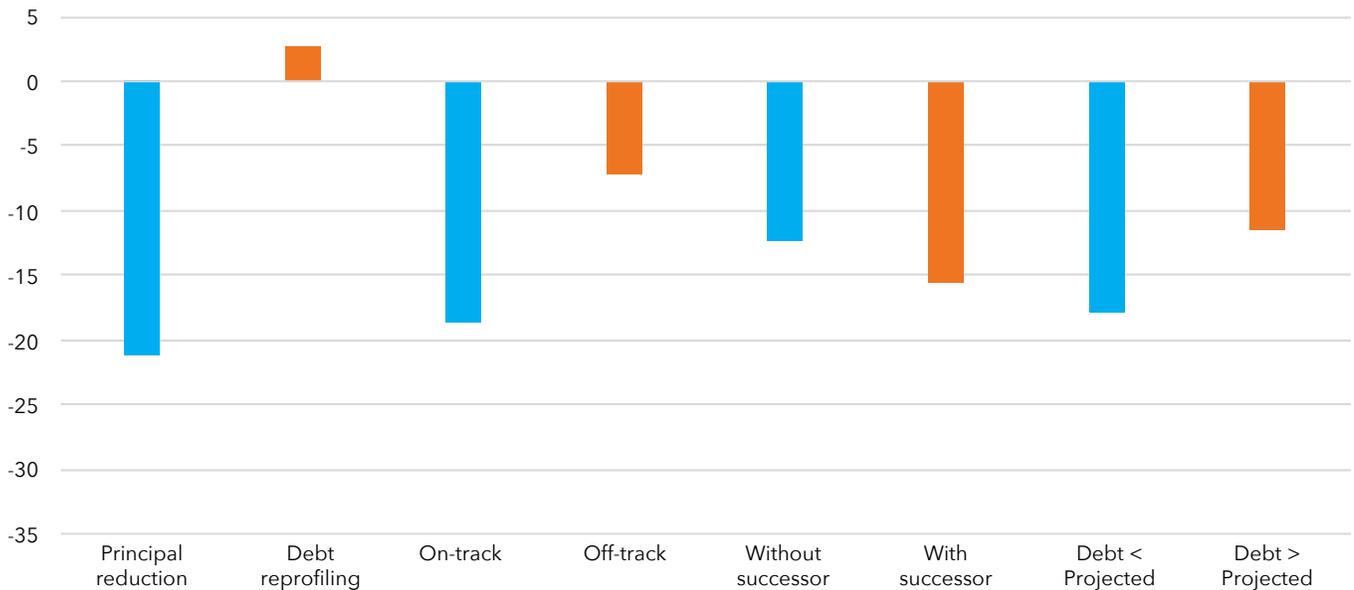
Source: Erce (2021).
Note: Based on the data for 10 programs with market debt operations, excluding Barbados (2018) and The Gambia (2017). "Without successor" stands for programs not followed by a successor program in less than 3 years; "Debt < projected" denotes programs where actual debt is less than projected; Public debt data represent the face value and do not take account of additional reduction in NPV terms.

FIGURE 45. GROWTH OUTCOMES OF PROGRAMS WITH DEBT OPERATIONS

(Average annual deviation from growth benchmark)
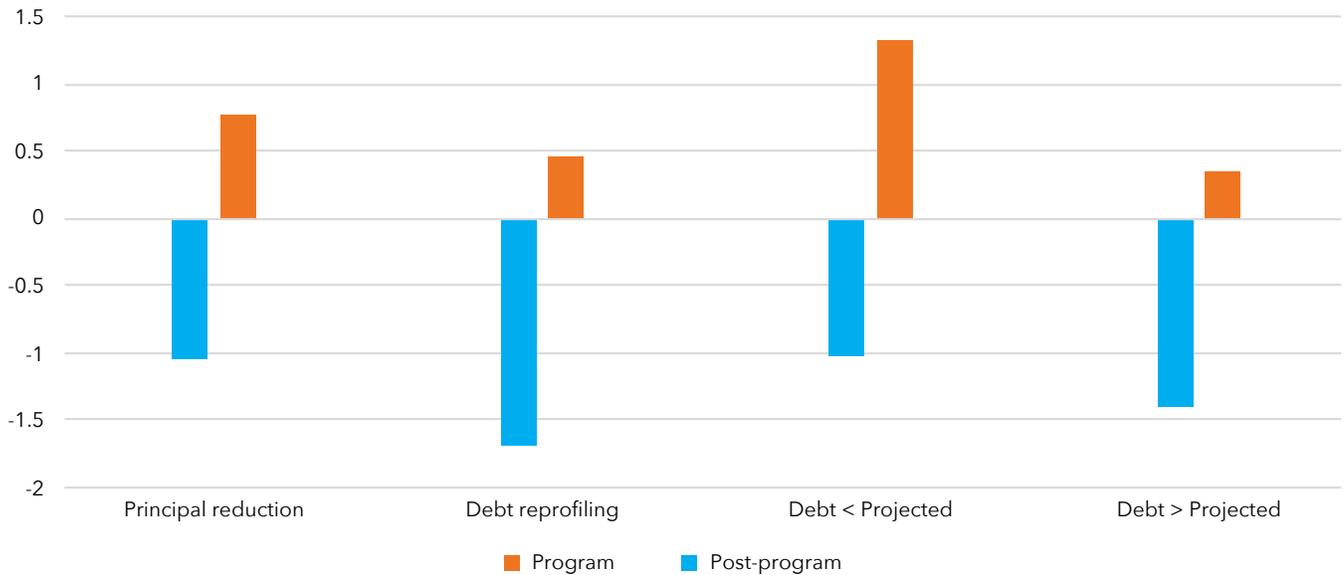


Source: Erce (2021).
Note: Based on the data for 10 programs with market debt operations, excluding Barbados (2018) and The Gambia (2017). Post-program period spans five years after the program ended. See Figure 44 for the definition of the categories on the horizontal axis.

reprofiling; and (ii) programs where debt trajectory has been better than projected during the program.[65]

Finally, it is worth recalling the results reported in Chapter 4 from growth regressions that while a reduction in the public debt-to-GDP ratio supports post-program growth, the debt operation itself can have a lingering adverse effect, presumably through the impact on borrowing costs and market access. This provides a reminder that the process of debt restructuring can affect growth outcomes. Some supporting evidence is provided in the literature, which suggests that pre-emptive negotiations, which avoid accumulating arrears, result in lower overall output costs from debt strains (Asonuma and Trebesch, 2016), and that hard defaults are more damaging for growth (Trebesch and Zabel, 2017).

LESSONS FROM COUNTRY EXPERIENCE

The country case studies provide more support for the view that debt operations that involved upfront principal reduction have led to more decisive and credible impact on debt sustainability than otherwise, with favorable growth implications. In Grenada, debt restructuring involved haircuts on a wide array of government debts as well as maturity extension which, taken together, resulted in the total NPV haircuts on the order of 50–60 percent, and allowed for a large reduction in the overall debt-to-GDP ratio. The successful operation helped to sustain domestic support for the program as well as reducing the debt burden on the economy. In Ukraine, debt restructuring involved principal reduction aimed at reducing the debt-to-GDP ratio by 20 percentage points, which helped the country to regain its access to international capital markets within two years at reasonable costs.

By contrast, in Jamaica where most of the debt was owed to domestic financial institutions, major debt restructuring with principal haircut was constrained because of concerns about potential financial stability risks. Debt operations were eventually undertaken in two rounds and focused on lowering interest rates and maturity extension in order to preserve the health of the domestic financial system, while excluding external debt from restructuring for concerns on insufficient creditor participation as well as future market access. With more limited reduction of the debt burden, sustained large primary surpluses have been necessary to restore debt sustainability.

Country experience illustrates how delays in debt negotiation or agreement on terms insufficient to restore debt sustainability can be detrimental in terms of regaining market access and restoring confidence and investment. Grenada (2010) and Ukraine (2014) did not include debt restructuring operations, even though debt vulnerabilities were acknowledged as rising, while Jamaica needed two rounds of restructuring as the first round was insufficiently aggressive and fiscal slippage occurred after the 2010 restructuring. Drawing on an earlier IEO evaluation (IEO, 2016), in Greece, pressures from European partners related to contagion concerns and the adverse impact on creditor balance sheets delayed debt operations and then limited the ability of the Fund and country authorities to set a realistic restructuring envelope based on the DSA framework. With Greece's commitment to membership of the euro area, delayed debt operations and related uncertainty imposed a high toll on growth and social inclusion.

Country case studies suggest that tailored design features can be helpful to ensure adequate creditor participation and to provide greater growth resilience in the face of adverse shocks. In Grenada, for instance, principal-based debt restructuring entailed two novel features: (i) a hurricane clause which provides for immediate debt moratorium in the event of another natural disaster, and (ii) a revenue-sharing clause in new bond contracts. These features helped to achieve wide creditor participation by providing creditors with upside potential while increasing resilience by making debt services contingent on shocks to growth.

ASSESSMENT

The IMF played a positive role in supporting debt operations in the program context during the evaluation period. Financing assurance and lending into arrears policies as well as the catalytic role of the Fund have provided

---

[65] These results are consistent with the literature. For example, Reinhart and Trebesch (2016) show that the macroeconomic situation of debtors improves significantly after debt relief operations, but only if these involve principal write-offs. Cheng and others (2018) find that more generous restructurings involving principal relief are associated with an acceleration of GDP growth, a reduction in poverty and inequality, and a drop of subsequent aid flows. See Erce (2021) for a more extensive literature review on the growth consequences of debt operations.

balanced incentives on both debtors and creditors to engage constructively in debt operations to restore debt sustainability. Beyond this, the IMF's main contribution has been technical advice based on its DSA frameworks, which have been used to determine the extent of debt relief needed to restore debt sustainability. In addition, Fund staff has often played a role in encouraging creditor participation in debt negotiation by communicating with stakeholders about a country's economic policies and prospects and debt servicing potential.

Reviews of experience suggest that debt operations with principal haircuts supported by upfront fiscal adjustment tended to be more successful in restoring debt sustainability and supporting growth than those with just debt reprofiling or lowered coupons, which delivered too limited NPV reduction. Although reprofiling operations can be effective if accompanied by sufficiently committed fiscal adjustment, this was clearly a harder route and took longer to demonstrate success.

Overall, IMF-supported programs with market debt operations included in the evaluation period have had only mixed success in strengthening debt sustainability and improving the BOP position over the medium-term. This experience in programs with debt operations confirms that successful debt operations can contribute to progress in lowering debt trajectory and restoring growth—but that debt operations that are "too little and too late" can fail to achieve these goals.[66]

This conclusion raises the question of whether the IMF should be more demanding in ensuring that debt operations in the program context achieve their objectives in terms of debt sustainability and providing a stronger basis for growth. In some cases, with hindsight, it seems that the IMF should have insisted that more ambitious debt operations were needed upfront to address debt sustainability concerns in order to qualify for financing. Steps to make the DSA frameworks more rigorous in the recent revisions to the MAC DSA and LIC-DSF may help to provide a more effective basis for the IMF to insist on more timely and adequate debt operations as a condition for access to Fund financing. However, further attention could be paid to reflecting more systematically how debt operations may affect market access and growth prospects, particularly if debt operations involve default on external debt and a restructuring of domestic debt owed to financial institutions and social security systems.

The potential growth consequences of specific design features of debt operations could also receive more attention. For example, while it may be helpful to secure high creditor participation by sharing some upside with creditors, if such features are too generous, they could backfire and make it more difficult to grow out of debt. Also, it can be helpful to introduce counter-cyclical features in restructured debt, including to automatically adjust debt service obligations in the event of natural disasters which can enhance growth resilience in the face of shocks.

---

[66] This finding is consistent with an earlier review of sovereign debt operations within IMF-supported programs which noted that they often took place long after Fund staff had assessed debt to be unsustainable and failed to durably re-establish market access (IMF, 2013b).

10

FINDINGS, LESSONS, AND RECOMMENDATIONS

This chapter summarizes the main findings of this evaluation, derives some broad lessons, and then recommends specific steps that the Fund could take towards fostering stronger growth-related outcomes in the program context. While the evaluation does not assess the experience during the COVID-19 pandemic, its lessons have become even more relevant as countries now face particularly strong headwinds to growth as they seek IMF support for achieving durable recoveries.

FINDINGS

Increasing attention to the growth consequences of IMF-supported programs seems to have delivered some positive results. The evaluation does not find evidence of a consistent bias towards excessive austerity in IMF-supported programs during the evaluation period (2008–19). IMF-supported programs during this period were in most cases (except in the crisis context) able to sustain output broadly in line with a growth benchmark that corrects for exogenous external factors, while still delivering needed adjustment. Indeed, cross-country evidence suggests that programs have yielded significant growth benefits relative to a counter-factual of no Fund program engagement and that stabilization and reforms implemented in the program context raised post-program growth. Historical data over a longer time horizon suggest a positive role of IMF-supported programs at initiating sustained growth surges.

Analysis of program design and adaptation shows that programmed fiscal policy incorporated both sustainability and growth considerations although less so in initial program design in the case of PRGT programs. In program reviews, fiscal adjustment targets tended to be revised downwards in response to interim growth shortfalls and upwards in response to adjustment slippages in both GRA and PRGT programs. However, very few programs included explicit contingencies for addressing adverse growth shocks.

Notwithstanding these positive findings, growth outcomes consistently fell short of projections incorporated in the program's macroeconomic framework, both during programs and in the post-program period, consistent with the findings of the 2018 ROC. Of the programs covered in the evaluation, around one-half experienced an average growth shortfall during the program period of ½ percentage points or more, while one-fourth had a growth shortfall of over 1.5 percentage points. Growth shortfalls were particularly marked in the first year of GRA programs in the crisis context, but were observed in PRGT projections too, particularly in the post-program period. Macro modeling errors, particularly those related to fiscal multiplier assumptions, seem to have been a significant source of such growth optimism, particularly in GRA programs outside of a crisis context. While fiscal multiplier assumptions seem to have been broadly in line with the "bucket approach" suggested by guidance given to staff, they were not discussed widely in program documents and their adaptation to country circumstances seems to have been limited. At the same time, case study evidence suggests that

political economy considerations in program negotiations that encouraged agreement on ambitious growth projections also played a significant role.

Persistent growth optimism raises serious concerns because growth outcomes below program projections in the macroeconomic framework imply slower than intended progress in increasing incomes and strengthening the public balance sheet, undercut program ownership, and fuel rising adjustment fatigue and public opposition to reforms. While greater scrutiny of the realism of program projections as recommended by the 2018 ROC could help to reduce growth optimism, it seems even more relevant to consider whether IMF-supported programs can achieve more robust growth outcomes more in line with the program's macroeconomic framework by paying greater attention to growth-friendly policies in program design and implementation.

To shed light on this question, the evaluation examined to what extent different policy instruments were used to support the program's growth-related outcomes and their impact. It found that fiscal policies typically incorporated growth-friendly measures but with mixed success. Tax mobilization improved in PRGT programs making space for higher capital spending than otherwise, while GRA programs were able to help encourage a more growth-promoting tax structure in the post-program period. However, GRA programs often relied heavily on spending cuts to achieve deficit reduction during the program, and there were no significant increases in health and education spending in either PRGT or GRA programs. A number of case studies raised concerns that growth benefits of higher public investment could be limited by poor project selection and wasteful implementation and that efforts to protect low-income and vulnerable groups often fell short of their goals.

Structural conditionality included in programs played a positive role in promoting structural reforms and growth, but the potential growth benefits of structural reforms were not fully realized. SC implementation was positively associated with independent measures of progress in structural reforms and helped to boost growth within and after the program, with stronger growth impact for higher-quality SCs. However, the bulk of SC was oriented to stabilization rather than promoting growth and the

average depth and growth orientation of SCs was relatively low. Fund CD assistance was actively provided to support reforms and associated SCs in the program context and was generally appreciated by country authorities. However, cross-country evidence suggests that CD support does not seem to have been delivered more to countries with weaker capacity nor consistently effective in strengthening SC implementation. Some country officials observed that SCs were often too numerous, going beyond a country's capacity to deliver even with CD support, and embodied unrealistic timetables. Moreover, in their view, Fund teams sometimes paid too little attention to growth-oriented reforms, relying too heavily on partner institutions, even for reforms crucial to program success. Implementation was significantly weaker for SCs in areas outside of Fund expertise and where collaboration with partners was sought.

The use of the exchange rate as a policy tool to support growth and external adjustment during programs was quite limited. Exchange rate regime transition was infrequent during the evaluation period, and more often toward greater fixity. Where more flexible regimes were introduced, progress was often at least partly reversed, in part because of volatile markets in the context of insufficiently supportive macroeconomic policies. Efforts were typically made to correct clear cases of overvaluation and were generally successful, although more generally the impact of nominal exchange rate movements on the REER were partially muted by pass-through to prices. There was also a tendency towards a loss of competitiveness in PRGT programs that relied on a heavily managed exchange rate as an anchor for inflation. Nevertheless, where significant REER depreciation did occur, it seems to have supported external adjustment and growth, particularly in PRGT programs, although there were also disappointments, particularly in the face of supply-side impediments to the export response.

In a number of cases, market debt operations were useful to restore debt sustainability and provide the basis for renewed market access, supporting a return to growth. However, the overall record was mixed, and there were examples in which debt operations were too little and too late, and thus had only limited impact in strengthening debt sustainability and improving growth prospects. Debt operations with principal haircuts and upfront fiscal adjustment were

more successful than those with just debt reprofiling and lower coupons.

LESSONS

While this evaluation acknowledges increased attention to growth in IMF-supported programs and finds that such programs have generally played a positive role in promoting growth, the fact that growth outcomes have typically not met growth projections embodied in program macroeconomic frameworks suggests a need for increased attention to growth-related aspects of program design and implementation. The aim should be to strengthen growth-related outcomes, both during programs and in the post-program period, while ensuring that needed external adjustment takes place to correct balance of payments problems.

The diverse experience in the case studies underlines that there is no simple recipe for delivering better growth-related outcomes in IMF-supported programs given the variety in country circumstances and preferences, the underlying causes and contexts of the BOP problems, and the potential scope for policy action. Moreover, the need for careful tailoring is underlined by clear experience that it is essential that the adjustment and growth strategy be fully owned by the government and broadly supported. Particularly in the context of a BOP crisis, ambitious upfront adjustment and reforms may quickly restore growth after an initial downturn by restoring confidence and market access. However, in other cases, more gradual adjustment and reform paths may be better suited to a country's limited capacity and fragile social tolerance for short-term economic stress. In some circumstances, stabilization may by itself be sufficient to restore a satisfactory growth path, while in other situations there may be greater need for deep reforms to raise a country's medium-term growth potential. Moreover, the approach taken to address social and distributional concerns, particularly to ensure adequate protection for the vulnerable and growth benefits for low-income groups will depend on country capabilities and national preferences.

In developing growth strategies, particular care should be paid to ensuring that macroeconomic frameworks used in program design incorporate realistic program assumptions and that program design pays more consistent attention to contingencies for growth shortfalls. Continued efforts should be made to developing and applying a suite of tractable models suitable for use in different country circumstances to analyze the growth impact of adjustment and reform policies. At a minimum, more attention is needed to ensure that fiscal multipliers are carefully tuned for country circumstances and that expectations for the pace and impact of reforms are not excessively sanguine. More explicit analysis of short-term fiscal multipliers in staff reports would enable a more realistic understanding of short-term growth consequences of fiscal adjustment and could help reduce optimism bias. Moreover, greater attention should be paid to program contingencies at the initial program design stage as well as during program reviews, particularly on how to respond to unexpected growth shortfalls. This early attention will help to not only guide subsequent program adaptation in a timely way but also promote country ownership and alleviate negative perception of the Fund's austerity bias.

The evaluation also provides lessons for how a broad spectrum of policy tools—fiscal policy, structural reforms, exchange rate policy, and debt operations—can be used to foster stronger growth outcomes in the program context.

In the area of fiscal policy, greater attention is warranted to ensure that fiscal adjustment and reforms are indeed growth friendly and inclusive. The apparent lack of progress on raising social spending, especially on education and health care, is disappointing and the limited monitoring of distributional impact limits the ability to make mid-course corrections. More granular approaches to conditionality and monitoring in this area could help to ensure that social spending to support low-income and vulnerable groups is at least protected during adjustment and raised in a durable manner over time. While it is encouraging that public investment has been boosted in a number of programs, the case studies demonstrate clearly that more attention is needed to maximize the growth impact of such investment and limit the risk of acquiring new debt without significantly raising debt-servicing capacity. This will require addressing transparency and governance issues especially to ensure a productive allocation of investment resources and limit leakages and corruption, building on the staff's continuing work in providing technical support on public financial management and applying the new governance framework introduced in 2018.

Greater focus on growth-enhancing structural reforms in IMF-supported programs would help to raise medium-term growth prospects given the clear evidence for the importance of the depth and growth orientation of SCs in determining the growth impact of reforms. At the same time, too many low-quality SCs should be avoided following the principle of parsimony and macro-criticality. Recognizing that higher quality SCs take more time to implement, Fund arrangements of longer duration could allow for a more realistic time frame for reform implementation. In addition, steps could be taken to foster more effective integration of CD support with program implementation, including to target more Fund CD resources at countries with limited capacity and giving CD experts more of a role in setting and monitoring program structural conditionality. More effective collaboration with partner institutions could produce greater synergy and traction in supporting reforms in areas with high growth impact that lie outside IMF core expertise

The limited use of exchange rate adjustment as a tool in the program context suggests that there could be greater scope to use exchange rate policy as a means to facilitate adjustment while supporting growth and resilience to adverse shocks, subject to the principle that the exchange rate regime choice is ultimately the authorities' decision. Cross-country evidence suggests that, depending on a country's economic structure, significant depreciation of the REER can help to boost exports and restrain imports, helping to shift the trade-off between external adjustment and growth. The case studies show that such depreciation can be achieved within different exchange rate regimes (including through internal devaluation under a currency union or peg) depending on country circumstances. The greatest and most effective route will be an upfront currency adjustment, although care will be needed to limit exchange rate pass through to inflation and ensure that any depreciation is consistent with members' obligations under Article IV to avoid manipulating exchange rates to prevent effective BOP adjustment or to gain an unfair competitive advantage. Use of the exchange rate as a policy tool would need to take due account of country circumstances, respect the member's right to choose their exchange rate regime, and address the concerns giving rise to a "fear of floating." Doing so will require assisting countries to build a supporting policy framework, including to securely anchor inflation expectations, to develop foreign exchange markets

with adequate depth and liquidity, to address foreign currency balance sheet mismatches and distributional consequences of exchange rate depreciation, and to alleviate supply-side impediments to export growth.

The experience of IMF-supported programs with debt operations suggests that the Fund should seek to make sure that where restructuring is needed to address debt sustainability concerns to qualify for access to IMF financing, it is not "too little and too late." While respecting the neutrality principle, applying a consistently careful approach to debt sustainability assessment would help ensure that where debt restructuring is needed, it is achieved in a timely and growth-friendly manner with adequate depth. Recent modifications to the LIC-DSF and MAC DSA should help in this respect. In addition, the potential growth and market access consequences of debt operations, including their specific design features, could receive more attention in analyzing the consequences of debt operations. Creative design may help facilitate debt negotiation and secure high creditor participation in debt exchanges, thus allowing for more rapid restoration of market access to new financing. However, it could also backfire if restructuring terms are too generous to creditors and discourage debtors' policy effort to grow out of debt if growth dividends to creditors are too great.

Finally, two more general lessons are worth emphasizing. First, in order to ensure that program design is well tailored to country needs and circumstances, the groundwork for a successful policy response to cushion the output and distributional consequences of an adverse exogenous shock should ideally be laid well in advance through surveillance and CD work. The case studies repeatedly show that meaningful reforms to strengthen such growth resilience take many years to put in place and become effective, even with strong capacity development support. In this respect, areas for attention include building an institutional structure for an effective social safety net, strengthening governance over public investment, establishing a workable framework for effective exchange rate management, and identifying structural impediments to investment, productivity and export gains.

Second, growth and reform strategies envisaged in program design should pay adequate attention to social and distributional consequences in line with country circumstances

and national objectives. While the focus in this evaluation has been largely on aggregate outcomes, fair distribution of the burden of adjustment and the rewards of recovery are important in their own right to meet national goals and to ensure continued public support for program implementation. Towards this end, there is a need to strengthen the analysis, monitoring and reporting of the social impact of the overall program and of the specific policies to protect vulnerable groups. The lack of a capacity to track effectiveness made it hard to track progress made in achieving inclusive growth, to identify emerging risks, and to assess the need for further reinforcing actions.

RECOMMENDATIONS

This section suggests specific actions that could be considered to strengthen growth-related outcomes in the program context both during the program period and in the medium term, while ensuring needed external adjustment. These actions are grouped into three umbrella recommendations: first, to increase the overall attention to growth-related implications in designing and implementing Fund-supported programs; second, to encourage deeper and more growth-oriented structural reforms; and third, to further develop the tools needed to support greater attention to sustainable and inclusive growth in program work.

> **Recommendation 1—Attention to growth implications of IMF-supported programs should become more thorough, systematic, realistic, and sensitive to social and distributional consequences.**

▶ Board papers supporting GRA as well as PRGT programs should clearly explain the program's growth implications, both during the program and over the medium term. They should discuss how program design reflects the country's growth strategy, including whether and how the program will help to protect activity during the program and help the country achieve sustainable medium-term growth while solving its balance of payments problems in a manner consistent with the Articles of Agreement. The relevant considerations will vary depending on country circumstances and national preferences, including the country's social and distributional goals.

▶ The discussion of growth implications in Board documents should provide a more thorough analysis of how growth has been taken into account in the design of the underlying macroeconomic framework of the program, including the interaction of different policy tools, ideally based on a well-calibrated country-specific model. Documents should provide more systematic coverage of the quality dimensions of growth, including distributional consequences of adjustment and reform policies, such as how low-income and vulnerable groups are affected during the program period and how they would share in growth over time.

▶ In discussing the macroeconomic framework, particular attention should be paid in program documents to discussion of fiscal multiplier assumptions, especially where available country-specific modeling is limited. While the bucket approach could continue to provide a useful starting point for fiscal multiplier discussion, multiplier assumptions should be further fine-tuned to country circumstances based on available evidence and informed judgement.

▶ Program design should pay more consistent attention to contingencies for growth shortfalls, based on scenario analysis, which should help better prepare to deal with adverse shocks and help fend off negative perceptions of the Fund's austerity bias. The appropriate approach would be determined case by case. In some situations, inclusion of explicit growth contingencies in the program may be helpful. In others, program adjustments may be best determined in the context of reviews but discussion of growth contingencies with authorities at the program design stage would still be desirable to foster ownership and preparedness to deal with adverse developments.

▶ Efforts to pay greater attention to distributional aspects related to growth may require more granular approaches to conditionality and monitoring. Conditionality could focus more on policies needed to achieve distributional outcomes

where they are of critical importance for achieving program goals, while strengthened monitoring of key social and distributional metrics would help to measure progress and signal emerging issues to be addressed in program reviews. This work would need to be adapted to data availability, which is likely to be quite limited in the context of many LICs.

▶ Revisions to the 2002 Guidelines on Conditionality and the 2014 Operational Guidance Note on Conditionality should be considered to give further guidance on the role of Fund-supported programs in fostering favorable growth outcomes while solving the member's balance of payments problems in a manner consistent with the Articles of Agreement. These revisions could elaborate further on the appropriate treatment in Fund-supported programs of a country's growth-related objectives and of considerations related to the quality of growth, including protecting vulnerable groups during the program period and encouraging inclusive and sustainable growth over the medium term, tailored to country circumstances and national preferences. They could also provide updated guidance on the use of contingencies for growth shortfalls and the application of structural conditionality (consistent with Recommendation 2). The update to the Guidance Note in response to the 2018 ROC now under way can provide an opportunity to advance this work. Revisions to the Guidelines on Conditionality could be considered in the next Review of Program Design and Conditionality, which would involve broad consultation and require eventual approval by the Executive Board.

Recommendation 2—IMF-supported programs should pay greater attention to supporting deep, more growth-oriented structural reforms with more effective capacity development support and more effective collaboration with partners in areas outside the Fund's core mandate and expertise.

▶ The program's structural reform strategy should be geared to what is important and not what is most easy to agree on or monitor or where the IMF has core expertise, subject to careful consideration of the country's implementation capacity and the program's adjustment and growth-related goals.

▶ Structural conditionality should be parsimonious enough to avoid overtaxing country capacity but also more focused on correcting underlying distortions and removing structural impediments critical to achieving sustained and inclusive growth even though this may require greater attention to areas outside the IMF's core competencies. Under such an approach, there would be less dependence on structural benchmarks that are relatively shallow and greater reliance on a review-based approach to assessing progress towards reforms critical to achieving the program's growth-related goals.

▶ Recognizing the limits on IMF expertise outside core areas, the Fund should seek ways to strengthen collaboration with the World Bank and other relevant partners in design and implementation of structural reforms in shared and non-core areas to foster an increased focus on and more effective delivery of growth-oriented reforms. These efforts would need to avoid cross-conditionality consistent with the principle that the Fund be fully responsible for setting and monitoring all conditions attached to use of its resources and protect against undue delays in completing reviews and making disbursements. A useful step could be preparation of a Board paper reviewing experience with Bank-Fund collaboration in Fund-supported programs.

▶ The Fund should revisit how CD support is integrated with program design and implementation, aimed at promoting deeper and more successful reform efforts in the program context.

For this purpose, CD experts could be involved more in program implementation and monitoring, which may be facilitated by greater use of virtual or hybrid meetings. The ongoing IEO evaluation of IMF capacity development can contribute to a reassessment in this area, ahead of the next strategic review of IMF CD work in 2023.

Recommendation 3—The Fund should continue to invest in building a toolkit of models and monitors that can be applied as a basis for analysis of the adjustment-growth relationship and assessing growth-related developments in the program context.

▶ Functional departments could continue to take the lead in developing a suite of models suitable for analyzing the adjustment-growth relationship that are tractable and easily accessible for use by country desks to calibrate and apply in their country context. Particular attention should be paid to developing small-scale, easy-to-adapt macro/growth models for LICs where data are limited.

▶ Country teams should be encouraged to apply the models now being developed to achieve greater realism in program projections, to explore trade-offs between alternative policy mixes, and explain baseline projections and associated risks to authorities, which should help promote country ownership and mitigate the tendency towards growth optimism. Teams would determine case by case the models best suited to country circumstances and needs. Area departments could also contribute by undertaking in-depth case studies on program successes and failures.

▶ The Fund should increase efforts to keep track of whether structural reforms included in programs were sustained after the program concludes. This initiative could involve investing more in the new Research Department structural reform database.

▶ Further attention should be given to developing and deploying monitors to help support country desks' capacity to track developments in key distributional indicators such as median incomes and poverty rates, to provide more current and granular information to gauge program impact on key social distributional dimensions of growth, as suggested under Recommendation 1. This work could be done in close collaboration with the World Bank and other agencies.

Budgetary Implications

It should be recognized that full implementation of these recommendations would have significant resource costs. Most significantly, the recommendations to take a fuller and more rigorous approach to analyzing and supporting program growth strategies with greater attention in program documents could add considerably to the time needed for program work (including for effective collaboration with the World Bank and other partners). More extensive coverage of reforms that are important for growth but not in the core of IMF expertise would require additional efforts at strengthening collaboration with development partners and additional specialized resources in-house (including to support effective collaboration). The research work to build a set of useful macroeconomic models to underpin these efforts would require substantial continued investment. And greater efforts at monitoring and reporting on the social and distributional consequences of policies would require a sustained effort across multiple agencies in which the IMF would be just one player.

At the same time, much of this work is already well under way or at least anticipated in the Fund's work program. New tools have been developed for use in debt sustainability assessment and to guide work on social spending and governance issues. Considerable efforts are already underway to develop models that could be used in the program context, which will help deliver on the commitment to improve the realism of program projections as part of the follow-up to the 2018 ROC. Taking on the additional commitments required would depend on a broader strategic decision to increase attention in the program context to ensure that IMF-supported programs not only deliver sufficient adjustment but also contribute in a more thorough way to sustainable and inclusive growth.

ANNEX

TABLE A1. IMF LENDING ARRANGEMENTS: SEPTEMBER 2008–MARCH 2020

DEPT.	COUNTRY	GRA	APPROVED	DEPT.	COUNTRY	PRGT	APPROVED
AFR	Angola[1]	SBA	2009	AFR	Benin[1]	ECF	2010
AFR	Seychelles[2]	SBA	2008	AFR	Burkina Faso[1]	ECF	2010
AFR	Seychelles[1]	EFF	2009	AFR	Burkina Faso[1]	ECF	2013
AFR	Seychelles	EFF	2014	AFR	Burundi[1]	ECF	2012
APD	Mongolia	SBA	2009	AFR	Cape Verde	PSI	2010
APD	Sri Lanka[1]	SBA	2009	AFR	Central African Republic[2]	ECF	2012
EUR	Albania	EFF	2014	AFR	Central African Republic	ECF	2016
EUR	Belarus[2]	SBA	2009	AFR	Chad[1, 2]	ECF	2014
EUR	Bosnia and Herzegovina[1]	SBA	2009	AFR	Comoros[1]	PRGF	2009
EUR	Bosnia and Herzegovina[1]	SBA	2012	AFR	Congo, Democratic Republic of	PRGF	2009
EUR	Cyprus	EFF	2013	AFR	Congo, Republic of	PRGF	2008
EUR	Greece[2]	SBA	2010	AFR	Côte d'Ivoire[2]	PRGF	2009
EUR	Greece[2]	EFF	2012	AFR	Côte d'Ivoire[1]	ECF	2011
EUR	Hungary[1]	SBA	2008	AFR	Ethiopia[1]	ESF	2009
EUR	Iceland[1]	SBA	2008	AFR	The Gambia[2]	ECF	2012
EUR	Ireland	EFF	2010	AFR	Ghana	PRGF	2009
EUR	Kosovo, Republic of	SBA	2010	AFR	Ghana[1]	ECF	2015
EUR	Kosovo, Republic of	SBA	2012	AFR	Guinea[1]	ECF	2012
EUR	Kosovo, Republic of[1]	SBA	2015	AFR	Guinea-Bissau	ECF	2010
EUR	Latvia[1]	SBA	2008	AFR	Guinea-Bissau[1]	ECF	2015
EUR	Portugal[1]	EFF	2011	AFR	Kenya	ECF	2011
EUR	Romania[2]	SBA	2009	AFR	Kenya*[1, 2]	SBA-SCF	2015
EUR	Romania[1]	SBA	2011	AFR	Kenya*[1]	SBA-SCF	2016
EUR	Romania	SBA	2013	AFR	Lesotho[1]	ECF	2010
EUR	Serbia, Republic of[1]	SBA	2009	AFR	Liberia[1]	ECF	2012
EUR	Serbia, Republic of	SBA	2011	AFR	Madagascar	ECF	2016
EUR	Serbia, Republic of	SBA	2015	AFR	Malawi	ESF	2008
EUR	Ukraine[2]	SBA	2008	AFR	Malawi[2]	ECF	2010

DEPT.	COUNTRY	GRA	APPROVED	DEPT.	COUNTRY	PRGT	APPROVED
EUR	Ukraine	SBA	2010	AFR	Malawi[1]	ECF	2012
EUR	Ukraine[2]	SBA	2014	AFR	Mali[2]	ECF	2011
EUR	Ukraine[2]	EFF	2015	AFR	Mali[1]	ECF	2013
MCD	Armenia[2]	SBA	2009	AFR	Mauritania[1]	ECF	2010
MCD	Armenia[1]	EFF	2014	AFR	Mozambique[1, 2]	PSI	2010
MCD	Egypt	EFF	2016	AFR	Mozambique	PSI	2013
MCD	Georgia[1]	SBA	2008	AFR	Mozambique	SCF	2015
MCD	Georgia[2]	SBA	2014	AFR	Niger[1]	ECF	2012
MCD	Iraq[1]	SBA	2010	AFR	Rwanda[1]	PSI	2010
MCD	Iraq	SBA	2016	AFR	Rwanda[1]	PSI	2013
MCD	Jordan	SBA	2012	AFR	Rwanda[1]	SCF	2016
MCD	Jordan[1]	EFF	2016	AFR	São Tomé & Principe	PRGF	2009
MCD	Pakistan[1]	SBA	2008	AFR	São Tomé & Principe	ECF	2012
MCD	Pakistan[1]	EFF	2013	AFR	São Tomé & Principe[1]	ECF	2015
MCD	Tunisia[1]	SBA	2013	AFR	Senegal[1]	PSI-ESF	2008
WHD	Antigua and Barbuda	SBA	2010	AFR	Senegal[1]	PSI	2010
WHD	Costa Rica	SBA	2009	AFR	Senegal[1]	PSI	2015
WHD	Dominican Republic	SBA	2009	AFR	Sierra Leone	ECF	2010
WHD	El Salvador[2]	SBA	2009	AFR	Sierra Leone[1]	ECF	2013
WHD	El Salvador	SBA	2010	AFR	Tanzania	PSI	2010
WHD	Guatemala	SBA	2009	AFR	Tanzania[1]	SCF	2012
WHD	Jamaica	SBA	2010	AFR	Tanzania[1]	PSI	2014
WHD	Jamaica[2]	EFF	2013	AFR	Uganda[1]	PSI	2010
WHD	Jamaica	SBA	2016	AFR	Uganda[1]	PSI	2013
WHD	St. Kitts and Nevis	SBA	2011	APD	Bangladesh[1]	ECF	2012
WHD	Suriname[2]	SBA	2016	APD	Maldives*	SBA-ESF	2009
				APD	Solomon Islands	SCF	2010
				APD	Solomon Islands	SCF	2011
				APD	Solomon Islands[1]	ECF	2012
				EUR	Moldova*[1]	EFF-ECF	2010
				EUR	Moldova*	EFF-ECF	2016
				MCD	Tajikistan[1]	PRGF	2009
				MCD	Afghanistan, I. R. of	ECF	2011
				MCD	Afghanistan, I. R. of[1]	ECF	2016
				MCD	Armenia[2]	PRGF	2008
				MCD	Armenia*[1]	EFF-ECF	2010
				MCD	Djibouti[1]	PRGF	2008
				MCD	Georgia*	SBA-SCF	2012
				MCD	Kyrgyz Republic	ESF	2008
				MCD	Kyrgyz Republic[1]	ECF	2011
				MCD	Kyrgyz Republic	ECF	2015
				MCD	Yemen, Republic of[2]	ECF	2010

DEPT.	COUNTRY	GRA	APPROVED	DEPT.	COUNTRY	PRGT	APPROVED
				MCD	Yemen, Republic of[2]	ECF	2014
				WHD	Grenada	ECF	2010
				WHD	Grenada	ECF	2014
				WHD	Haiti[1]	ECF	2010
				WHD	Haiti[2]	ECF	2015
				WHD	Honduras*	SBA-SCF	2010
				WHD	Honduras*	SBA-SCF	2014

* Indicates GRA-PRGT blended arrangements.
1 Completion delayed or program extended but not completed (Pakistan 2008).
2 Cancelled.

FIGURE A1. DISTRIBUTION OF PROGRAM APPROVALS: 2008–16
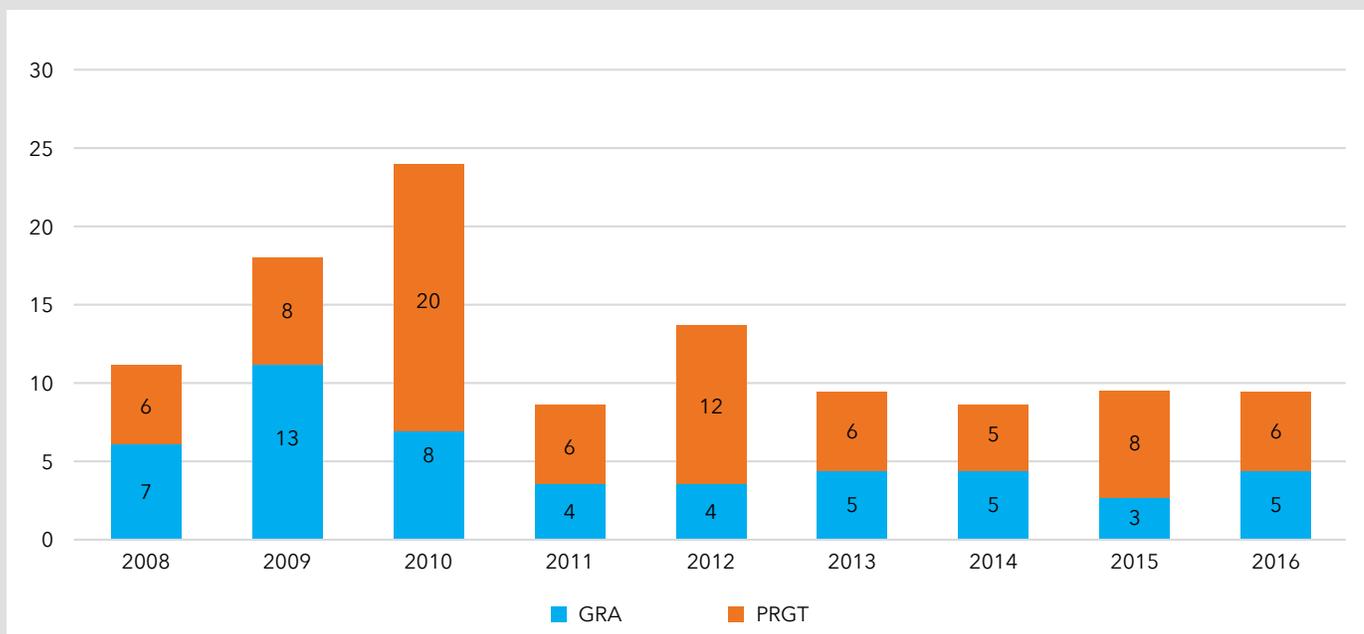
Sources: MONA database; IEO staff calculations.

TABLE A2. COMPOSITION OF THE EVALUATION SAMPLE

	FULL	GRA	PRGT
Number of countries[1]	75	33	44
Fragile states	25	2	23
Small states	11	4	7
Currency union members	20	5	15
Number of programs	131	54	77
Completed programs	82	32	50
Off-track programs[2]	27	15	13
Quickly off-track programs[2]	22	7	15
Precautionary programs	18	10	8
Exceptional access programs	26	26	0
Crisis programs[3]	23	23	0

[1] Armenia and Georgia have both GRA and PRGT programs.

[2] Following the definition used by the 2018 ROC, "off-track programs" refer to programs where at least two reviews were completed and at least two reviews were not completed at the end of the program and "quickly off-track programs" refer to programs where at most one review was completed and at least two reviews were not completed at the end of the program.

[3] Crisis programs include the programs approved during 2008–09 in response to the GFC (Angola, Armenia, Belarus, Bosnia and Herzegovina, Costa Rica, Dominican Republic, El Salvador, Georgia, Guatemala, Hungary, Latvia, Mongolia, Pakistan, Romania, Serbia, Sri Lanka, and Ukraine) and five euro area programs arranged in response to the European debt crisis (Cyprus (2013), Greece (2010, 2012), Ireland (2010), Portugal (2011)).

DATA CONVENTIONS

For analytical purposes, the following conventions are used to define the program period in the cross-country analysis unless otherwise indicated:

▶ *Convention 1.* If the program is approved in the last quarter of year t, the following year t+1 is counted as the first year (T) of the program; otherwise year t is the first year.

▶ *Convention 2.* If the program is fully completed in the first quarter of year t, the previous year t-1 is considered as the last year (T+E) of the program; otherwise, year t is the last year.

▶ *Convention 3.* For off-track programs, the last year of the program is determined based on the date of the last completed program review while applying Convention 2 above.

REFERENCES

Adler, Gustavo, and others, 2017, "Gone with the Headwinds: Global Productivity," IMF Staff Discussion Note SDN/17/04 (Washington: International Monetary Fund).

Asonuma, Tamon, and others, 2019, "Costs of Sovereign Defaults: Restructuring Strategies, Bank Distress and the Capital Inflow-Credit Channel," IMF Working Paper No. WP/19/69 (Washington: International Monetary Fund).

Asonuma, Tamon, and Christoph Trebesch, 2016, "Sovereign Debt Restructurings: Pre-Emptive or Post-Default," *Journal of the European Economic Association*, 14(1), pp. 175–214.

Atsebi, Jean-Marc Bedhat, and Joshua Wojnilower, 2021, "Initiating Growth Surges: The Role of IMF-Supported Programs," IEO Background Paper No. BP/21-01/02 (Washington: International Monetary Fund).

Bal Gündüz, Yasemin, 2016, "Economic Impact of Short-Term IMF Engagement in Low-Income Countries," *World Development*, 87, pp. 30–49.

Bal Gündüz, Yasemin, and Reginald Darius, 2021, "Exchange Rate Adjustment and Growth in IMF-Supported Programs," IEO Background Paper No. BP/21-01/05 (Washington: International Monetary Fund).

Barro, Robert, and Jong-Wha Lee, 2005, "IMF Programs: Who Is Chosen and What Are the Effects?" *Journal of Monetary Economics*, 52(7), pp. 1245–69.

Bas, Muhammet A., and Randall W. Stone, 2014, "Adverse Selection and Growth Under IMF Programs," *The Review of International Organizations*, 9(1), pp. 1–28.

Batini, Nicoletta, and others, 2014, "Fiscal Multipliers: Size, Determinants, and Use in Macroeconomic Projections," IMF Technical Notes and Manuals No. 2014/04 (Washington: International Monetary Fund).

Bird, Graham, and Dane Rowlands, 2017, "The Effect of IMF Programs on Economic Growth in Low Income Countries: An Empirical Analysis," *The Journal of Development Studies*, 53(12), pp. 2179–96.

Blanchard, Olivier J., and Daniel Leigh, 2013, "Growth Forecast Errors and Fiscal Multipliers," *American Economic Review: Papers and Proceedings*, 103(3), pp. 117–20.

Bohn, Henning, 1998, "The Behavior of US Public Debt and Deficits," *Quarterly Journal of Economics*, 113(3), pp. 949–63.

_____, 2008, "The Sustainability of Fiscal Policy in the United States," in *Sustainability of Public Debt*, ed. by Reinhard Neck and Jan-Egbert Sturm (Cambridge, MA: MIT Press), pp. 15–49.

Bordo, Michael D., and Anna J. Schwartz, 2000, "Measuring Real Economic Effects of Bailouts: Historical Perspectives on How Countries in Financial Distress Have Fared with and without Bailouts," *Carnegie-Rochester Conference Series on Public Policy*, 53(1), pp. 81–167.

Butkiewicz, James L., and Halit Yanikkaya, 2005, "The Effects of IMF and World Bank Lending on Long-Run Economic Growth: An Empirical Analysis," *World Development*, 33(3), pp. 371–91.

Cheng, Gong, Javier Diaz-Cassou, and Aitor Erce, 2018, "Official Debt Restructuring and Development," *World Development*, 111, pp. 181–95.

Choudhri, Ehsan, and Dalia S. Hakura, 2006, "Exchange Rate Pass-Through to Domestic Prices: Does the Inflationary Environment Matter?" *Journal of International Money and Finance*, 25(4), pp. 614–39.

Clements, Benedict, Sanjeev Gupta, and Masahiro Nozaki, 2013, "What Happens to Social Spending in IMF-Supported Programs?" *Applied Economics*, 45(28), pp. 4022–33.

Conway, Patrick, 2009, "Participation in IMF Programs and Income Inequality," in *Finance, Development, and the IMF*, ed. by James Boughton and Dominico Lombardi (Oxford: Oxford University Press).

Crivelli, Ernesto, and Sanjeev Gupta, 2016, "Does Conditionality in IMF-Supported Programs Promote Revenue Reform?" *International Tax and Public Finance*, 23, pp. 550–79.

Dicks-Mireaux, Louis, Mauro Mecagni, and Susan Schadler, 2000, "Evaluating the Effect of IMF Lending to Low Income Countries," *Journal of Development Economics*, 61(2), pp. 495–526.

Dreher, Axel, 2006, "IMF and Economic Growth: The Effects of Programs, Loans and Compliance with Conditionality," *World Development*, 34(5), pp. 769–88.

Easterly, William, 2005, "What Did Structural Adjustment Adjust? The Association of Policies and Growth with Repeated IMF and World Bank Adjustment Loans," *Journal of Development Economics*, 76(1), pp. 1–22.

Eichengreen, Barry, Poonam Gupta, and Ashoka Mody, 2008, "Sudden Stops and IMF-Supported Programs," in *Financial Markets Volatility and Performance in Emerging Markets* (Chicago: University of Chicago Press), pp. 219–66.

Erce, Aitor, 2021, "Market Debt Operations and Growth in IMF-Supported Programs," IEO Background Paper No. BP/21-01/06 (Washington: International Monetary Fund).

Evrensel, Ayşe Y., 2002, "Effectiveness of IMF-Supported Stabilization Programs in Developing Countries," *Journal of International Money and Finance*, 21(5), pp. 565–87.

Garuda, Gopal, 2000, "The Distributional Effects of IMF Programs: A Cross-Country Analysis," *World Development*, 28(6), pp. 1031–51.

Ghosh, Atish R., and others, 2013, "Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies," *Economic Journal,* 123(566), pp. F4–F30.

Ghosh, Jayati, 2019, "The IMF's Latest Victims," Project Syndicate, August.

Gupta, Sanjeev, 2021, "Fiscal Adjustment and Growth in IMF-Supported Programs," IEO Background Paper No. BP/21-01/03 (Washington: International Monetary Fund).

Gupta, Sanjeev, and others, 2000, "Social Issues in IMF-Supported Programs," IMF Occasional Paper No. 191 (Washington: International Monetary Fund).

Guscina, Anastasia, Sheheryar Malik, and Michael G. Papaioannou, 2017, "Assessing Loss of Market Access: Conceptual and Operational Issues," IMF Working Paper No. WP/17/246 (Washington: International Monetary Fund).

Hagan, Sean, Maurice Obstfeld, and Poul M. Thomsen, 2017, "Dealing with Sovereign Debt —The IMF Perspective," IMF Blog, February 23, 2017 (Washington: International Monetary Fund).

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik, 2005, "Growth Accelerations," *Journal of Economic Growth*, 10(4), pp. 303–29.

Hutchison, Michael, 2004, "Selection Bias and the Output Costs of IMF Programs," Working Paper Series 04-15, Economic Policy Research Unit, Department of Economics (Copenhagen, Denmark).

Hutchison, Michael, and Ilan Noy, 2003, "Macroeconomic Effects of IMF-Sponsored Programs in Latin America: Output Costs, Program Recidivism and the Vicious Cycle of Failed Stabilization," *Journal of International Money and Finance*, 22(7), pp. 991–1014.

Independent Evaluation Office of the International Monetary Fund (IEO), 2002, *Evaluation of Prolonged Use of IMF Resources* (Washington: International Monetary Fund).

_____, 2003, *Fiscal Adjustment in IMF-Supported Programs* (Washington: International Monetary Fund)

_____, 2007, *Structural Conditionality in IMF-Supported Programs* (Washington: International Monetary Fund).

_____, 2014, *IMF Forecasts: Process, Quality, and Country Perspectives* (Washington: International Monetary Fund).

_____, 2016, *The IMF and the Crises in Greece, Ireland, and Portugal* (Washington: International Monetary Fund).

_____, 2017, *The IMF and Social Protection* (Washington: International Monetary Fund).

_____, 2018a, *The IMF and Fragile States* (Washington: International Monetary Fund).

_____, 2018b, *Structural Conditionality in IMF-Supported Programs: Evaluation Update* (Washington: International Monetary Fund).

International Monetary Fund, 1979, "Guidelines on Conditionality: Use of Fund Resources and Stand-by Arrangements," March (Washington).

_____, 2002a, "Review of the Poverty Reduction and Growth Facility: Issues and Options," February (Washington).

_____, 2002b, "Guidelines on Conditionality," September (Washington).

_____, 2003, "Operational Guidance on the New Conditionality Guidelines," May (Washington).

_____, 2005, "Review of the 2002 Conditionality Guidelines," March (Washington).

_____, 2009a, "Creating Policy Space—Responsive Design and Streamlined Conditionality in Recent Low-Income Country Programs," September (Washington).

_____, 2009b, "Review of Recent Crisis Programs," September (Washington).

_____, 2010, "Operational Guidance Note on Conditionality—December 2009 Revisions," January (Washington).

_____, 2012a, "2011 Review of Conditionality, Overview Paper," June (Washington).

_____, 2012b, "2011 Review of Conditionality, Background Paper 1: Content and Application of Conditionality," June (Washington).

_____, 2012c, "2011 Review of Conditionality, Background Paper 2: Design of Fund-Supported Programs," June (Washington).

_____, 2012d, "2011 Review of Conditionality, Background Paper 3: Outcomes of Fund-Supported Programs," June (Washington).

_____, 2013a, "Jobs and Growth: Analytical and Operational Considerations for the Fund," March (Washington).

_____, 2013b, "Sovereign Debt Restructuring—Recent Developments and Implications for the Fund's Legal and Policy Framework," April (Washington).

_____, 2013c, "Greece: Ex Post Evaluation of Exceptional Access under the 2010 Stand-By Arrangement," IMF Country Report No. 13/156 (Washington).

_____, 2014a, "The Fund's Lending Framework and Sovereign Debt—Annexes," May (Washington).

_____, 2014b, "Operational Guidance Note on Conditionality—July 2014 Revisions," July (Washington).

_____, 2014c, "Reform of the Policy on Public Debt Limits in Fund-Supported Programs," November (Washington).

_____, 2015a, "Where Are We Headed? Perspectives on Potential Output," *World Economic Outlook*, Chapter 3, April (Washington).

_____, 2015b, "Review of Crisis Programs," November (Washington).

_____, 2015c, "Public Debt Vulnerabilities in Low-income Countries: The Evolving Landscape," December (Washington).

_____, 2016, "Portugal: Ex Post Evaluation of Exceptional Access under the 2011 Extended Arrangement," IMF Country Report No. 16/302 (Washington).

_____, 2017, "Regional Economic Outlook: Sub-Saharan Africa," October (Washington).

_____, 2018a "CD Strategy Review—Staff Background Studies and Short Notes," October (Washington).

_____, 2018b, "2018 Review of the Fund's Capacity Development Strategy—Overview Paper," November (Washington).

_____, 2019a, "2018 Review of Program Design and Conditionality," May (Washington).

_____, 2019b, "2018 Review of Program Design and Conditionality—Supplementary Information," May (Washington).

_____, 2019c, "2018 Review of Program Design and Conditionality—Case Studies," May (Washington).

_____, 2019d, "A Strategy for IMF Engagement on Social Spending," June (Washington).

_____, 2019e, "Reigniting Growth in Low-income and Emerging Market Economies: What Role Can Structural Reforms Play?" *World Economic Outlook*, October, Chapter 3 (Washington).

_____, 2019f, "IMF Policies and Practices on Capacity Development," November (Washington).

_____, 2020a, "The Evolution of Public Debt Vulnerabilities in Lower Income Economies," IMF Policy Paper No. 20/003 (Washington).

_____, 2020b, "The Role of State-Contingent Debt Instruments in Sovereign Debt Restructurings," Staff Discussion Notes No. 2020/006 (Washington).

_____, 2021, "Review of the Debt Sustainability Framework for Market Access Countries," Policy Paper No. 2021/003, January (Washington).

Ismail, Kareem, Roberto Perrelli, and Jessie Yang, 2020, "Optimism Bias in Growth Forecasts—The Role of Planned Policy Adjustments," IMF Working Paper No. WP/20/229 (Washington: International Monetary Fund).

Kim, Jun, and others, 2021, "Cross-Country Analysis of Program Design and Growth Outcomes: 2008–19," IEO Background Paper No. BP/21-01/01 (Washington: International Monetary Fund).

Kim, Jung Yeon, and Kwang Yeon Lee, 2021, "Structural Conditions, Structural Reforms and Growth in IMF-Supported Programs," IEO Background Paper No. BP/21-01/04 (Washington: International Monetary Fund).

Kuvshinov, Dmitry, and Kaspar Zimmermann, 2019, "Sovereigns Going Bust: Estimating the Cost of Default," *European Economic Review*, 119, pp. 1–21.

Lambertini, Luisa, and Christian Proebsting, 2019, "Does Austerity Go Along with Internal Devaluations?" *IMF Economic Review*, 67(3), pp. 618–56.

Mendoza, Enrique G., and Jonathan D. Ostry, 2008, "International Evidence on Fiscal Solvency: Is Fiscal Policy "Responsible"?" *Journal of Monetary Economics*, 55(6), pp. 1081–93.

Mitra, Pritha, and Tigran Poghosyan, 2015, "Fiscal Multiplier in Ukraine," IMF Working Paper No. WP/15/71 (Washington; International Monetary Fund).

Mussa, Michael, and Miguel A. Savastano, 1999, "The IMF Approach to Economic Stabilization," IMF Working Paper No. 99/104 (Washington: International Monetary Fund).

Oberdabernig, Doris A., 2013, "Revisiting the Effects of IMF Programs on Poverty and Inequality," *World Development*, 46, pp. 113–42.

Ooms, Gorik, and Rachel Hammonds, 2009, "Scaling Up Global Social Health Protection: Prerequisite Reforms to the International Monetary Fund," *International Journal of Health Services,* 39(4), pp. 795–801.

Ostry, Jonathan D., and others, 2010, "Fiscal Space," IMF Staff Position Note No. 10/11 (Washington: International Monetary Fund).

Przeworski, Adam, and James R. Vreeland, 2000, "The Effect of IMF Programs on Economic Growth," *Journal of Development Economics*, 62(2), pp. 385–421.

Razafimahefa, Ivohasina F., 2012, "Exchange Rate Pass-Through in Sub-Saharan African Economies and its Determinants," IMF Working Paper No. 12/141 (Washington: International Monetary Fund).

Reinhart, Carmen M., and Christoph Trebesch, 2016, "Sovereign Debt Relief and Its Aftermath," *Journal of the European Economic Association*, 14(1), pp. 215–51.

Rodrik, Dani, 2008, "The Real Exchange Rate and Economic Growth," *Brookings Papers on Economic Activity*, 2008(2), pp. 365–412.

Rowden, Rick, 2009, "Concern Over IMF Impact on Health Spending," *IMF Survey Magazine: What Readers Say.*

Taylor, John B., 2000, "Low Inflation, Pass-Through, and the Pricing Power of Firms," *European Economic Review*, 44(7), pp. 1389–1408.

Terrones, Marco E., 2020, "Do Fixers Perform Worse Than Non-Fixers During Global Recessions and Recoveries?" *Journal of International Money and Finance*, 104, Article 102160.

Timmermann, Allan, 2007, "An Evaluation of the World Economic Outlook Forecasts," *IMF Staff Papers*, 54, pp. 1–33.

Trebesch, Christoph, 2019, "Resolving Sovereign Debt Crises: The Role of Political Risk," *Oxford Economic Papers*, 71(2), pp. 421–444.

Trebesch, Christoph, and Michael Zabel, 2017, "The Output Costs of Hard and Soft Sovereign Default," *European Economic Review*, 92, pp. 416–32.

Tsangarides, Charalambos G., 2012, "Crisis and Recovery: Role of the Exchange Rate Regime in Emerging Market Economies," *Journal of Macroeconomics*, 34(2), pp. 470–88.

Van Waeyenberge, Elisa, Hannah Bargawi, and Terry McKinley, 2010, *Standing in the Way of Development? A Critical Survey of the IMF's Crisis Response in Low Income Countries*, TWN Global Economy Series No. 32, European Network on Debt and Development (Eurodad) and Third World Network (TWN).

Vreeland, James R., 2003, *The IMF and Economic Development* (Cambridge: Cambridge University Press).

STATEMENT BY THE MANAGING DIRECTOR

ON THE INDEPENDENT EVALUATION OFFICE REPORT ON GROWTH AND ADJUSTMENT IN IMF-SUPPORTED PROGRAMS EXECUTIVE BOARD MEETING AUGUST 30, 2021

I welcome the report of the Independent Evaluation Office (IEO) on Growth and Adjustment in IMF-Supported Programs, including the background papers and country case studies. The report offers valuable analysis and recommendations and comes at a very opportune time, as Fund lending is at an all-time high. I agree with the overall assessment, which is broadly consistent with staff's 2018 Review of Program Design and Conditionality (ROC). I also broadly agree with the recommendations, which will be an important input into the ongoing review of the operational guidance note on conditionality, expected to be finalized in 2022.

I am pleased with the IEO conclusion that there is no evidence of a consistent bias towards excessive austerity in IMF-supported programs during the evaluation period. I also agree with the finding that Fund-supported programs have yielded growth benefits relative to a counterfactual of no Fund engagement and that stabilization and reforms implemented in the program context boosted post-program growth performance. I welcome the IEO's finding that fiscal multiplier assumptions were broadly in line with the guidance given to staff and the empirical literature. I should add in this respect that, as seen in some programs, macro-financial considerations may affect growth outcomes directly or condition the size of fiscal multipliers, and the composition of an adjustment path can affect growth outcomes. Thus, staff should continue striking a balance between the general guidance and consideration of country-specific policy circumstances.

I acknowledge the IEO's assessment that the limited use of exchange rate adjustment as a tool in a program context may suggest that there could be greater scope to use exchange rate policy as a means to facilitate adjustment while supporting growth and resilience to adverse shocks. I would like to emphasize nevertheless that there is no one-size-fits all approach, with the need for a case-by-case assessment of individual country circumstances. In some cases, the use of the exchange rate is not possible due to regional currency arrangements, large negative spillover risks, and ultimately national decisions on the exchange rate regime, which the IEO also acknowledges.

I also take note of the IEO's finding that successful debt operations can contribute to progress in lowering debt trajectory and restoring growth. While I agree in general with the need to avoid "too little and too late" debt restructurings, I would also like to emphasize that the Fund has a duty of neutrality, as also mentioned by the IEO, and has a limited role in debt restructuring operations. I am pleased to say that the impact and consequence of debt operations on growth and market access is already captured as one of the relevant economic risks in recent IMF/World Bank work, while the new MAC DSA guidance note would cover macro projections and realism issues.

RESPONSE TO IEO RECOMMENDATIONS

> **Recommendation 1. Attention to growth implications of IMF-supported programs should become more thorough, systematic, realistic, and sensitive to social and distributional consequences.**

I support this recommendation, with some qualifications.

I broadly concur with this recommendation. In particular, I agree with the need for more thorough discussion on growth issues in IMF-supported programs (both GRA and PRGT), and I would highlight that staff reports for program requests and reviews should—and many already do—discuss the growth strategy on a case-by-case basis, if critical to the program success.

In this context, I would echo the IEO's message that the ongoing revisions to the operational guidance note on conditionality will provide an opportunity to update guidance on program design and conditionality, based on the findings of the 2018 ROC regarding over-optimistic growth assumptions and of this IEO evaluation.

At the same time, while it is important to support economic adjustment and reforms for sustained growth over the medium term (e.g., fiscal balance targets or debt sustainability), growth cannot be placed above or on par with the core objective of Fund lending of helping members resolve their balance of payments problems, according to our Articles of Agreement. The focus needs to be on how we can provide further guidance on the role of Fund-supported programs in fostering sustainable growth, taking into account the quality of growth.

While I support the sub-recommendation to discuss fiscal multiplier assumptions, I would stress the need for a flexible application of tools, as fiscal multipliers are often difficult to assess ex ante, and even harder during crises.

I also concur with the sub-recommendation to pay more consistent attention to contingencies for growth shortfalls. In this regard, I would like to emphasize the importance of a case-by-case approach to developing contingency plans, depending on country circumstances. Importantly, program reviews already discuss key program risks, including growth shocks, and allow for program

modifications in response to unforeseen shocks. Finally, in some cases, contingency discussions with the authorities need to be confidential to avoid adverse market reactions.

Finally, I agree with strengthening the monitoring of key social and distributional metrics where feasible. I would, however, also note that limited data availability—particularly in countries with a large informal sector and weak household-level data—and possible budgetary implications for the Fund may limit a broader application to all programs. While some country teams already examine distributional consequences of adjustment and reform policies, a more systematic incorporation of such assessments in program design would require specialized staffing.

> **Recommendation 2. IMF-supported programs should pay greater attention to supporting deep, more growth-oriented structural reforms with more effective capacity development support and more effective collaboration with partners in areas outside the Fund's core mandate and expertise.**

I support this recommendation.

I agree with focusing more on growth-oriented structural reforms, effective capacity development (CD) and collaborating with partners such as the World Bank. I concur with a more effective collaboration with partners in areas outside the Fund's core mandate and expertise, including the World Bank. I would emphasize that staff—in program cases in particular—typically collaborate closely with World Bank counterparts on a range of topics, including structural fiscal issues and social support to financial development. I am pleased to state that avenues for further cooperation are being pursued in the context of the ongoing MIP for IMF Collaboration with the World Bank on Macro-Structural Issues. As the IEO correctly implies, expanding SC beyond the Fund's core expertise should recognize the need for Fund conditionality to be independently verifiable and monitorable by Fund staff, which also requires inhouse expertise, while cross-conditionality (linking the use of Fund resources to the rules and decisions of other organizations) is disallowed.

I would also note that the Board-endorsed recommendations of the 2018 ROC called for SC to be parsimonious, also reiterated by the IEO, and better prioritized in line with program objectives. The 2018 IEO evaluation update of SC in IMF-supported programs also found the SC in core Fund expertise to be associated with better implementation. I would thus caution against veering too far out of areas of the Fund's core mandate and expertise.

I agree that the Fund could revisit how CD support is integrated with program design and implementation, with the aim to promote deeper and more successful reform efforts in a program context. I look forward to the ongoing IEO evaluation of IMF CD, including how the coordination of Fund-supported programs and CD activities could be improved further, as well as the next strategic review of IMF CD work in 2023. I would note that the benefits of more effective CD in Fund-supported programs also depend on the country's implementation capacity. For a fuller picture, the costs associated with CD activities would also need to be fully reflected in the IEO report.

Recommendation 3—The Fund should continue to invest in building a toolkit of models and monitors that can be applied as a basis for analysis of the adjustment-growth relationship and assessing growth-related developments in the program context.

I support this recommendation.

I support further developing and refining our toolkit of models. A number of such models (DIG, DIGNAR, new RES model, and FAD's SDG financing model that "endogenizes" growth) are already available for both program and surveillance purposes. Moreover, an ongoing ICD initiative on Modernizing Financial Programming Toolkit ("FP 2.0") aims to build a tractable modeling toolkit and to design user-friendly ways to apply the modeling tools, calibrate baseline projections, reflect expected growth impacts of policy measures, and build risk scenarios. I agree that teams should be encouraged to make use of the available modeling resources and the new models on a case-by-case basis; further investment in the RES structural reform database could be indeed helpful.

TABLE 1. THE MANAGING DIRECTOR'S POSITION ON IEO RECOMMENDATIONS

RECOMMENDATION	POSITION
1. Attention to growth implications of IMF-supported programs should become more thorough, systematic, realistic and sensitive to social and distributional consequences.	QUALIFIED SUPPORT
2. IMF-supported programs should pay greater attention to supporting deep, more growth-oriented structural reforms with more effective capacity development support and more effective collaboration with partners in areas outside the Fund's core mandate and expertise.	SUPPORT
3. The Fund should continue to invest in building a toolkit of models and monitors that can be applied as a basis for analysis of the adjustment-growth relationship and assessing growth-related developments in the program context.	SUPPORT

THE ACTING CHAIR'S SUMMING UP

Executive Directors welcomed the comprehensive report of the Independent Evaluation Office (IEO) on Growth and Adjustment in IMF-Supported Programs, noting that it comes at an opportune time, as many member countries are seeking Fund support to close external gaps exacerbated by the pandemic, while building sustainable growth. They acknowledged that the overall assessment is broadly consistent with and complements staff's 2018 Review of Program Design and Conditionality (ROC). Directors welcomed that the IEO did not find evidence of a consistent bias toward excessive austerity in Fund-supported programs during the evaluation period and the finding that programs yielded growth benefits relative to a counterfactual of no Fund engagement. They also welcomed the Managing Director's broad support for the IEO findings and recommendations, while noting qualifications in some areas.

Directors broadly agreed with Recommendation 1 that attention to growth implications of Fund-supported programs should become more thorough, systematic, realistic, and sensitive to social and distributional consequences, while reiterating that the core objective of Fund lending is to help members resolve their balance of payment (BOP) problems without resorting to measures destructive of prosperity, as mandated by the Articles of Agreement. While it was also emphasized that growth is fundamental to sustainably resolving BOP problems, there was also recognition of macroeconomic adjustment and sustainable policies as a pre-condition for sustainable and balanced growth. Directors regretted that growth outcomes have often fallen short of program projections and concurred on the need to improve the realism of forecasts but also to pay greater attention to growth outcomes in IMF-supported program design. In this context, they considered that the findings of the IEO evaluation together with the ROC should provide important input to the ongoing review of the operational guidance note on conditionality. A number of Directors also saw merit in reviewing the 2002 Conditionality Guidelines to further clarify the balance between stabilization and growth considerations.

Directors agreed with the need to carefully discuss fiscal multiplier assumptions, while calling for a flexible application of tools, as multipliers are often difficult to estimate and depend on country-specific circumstances. They also recommended paying more attention to contingencies for growth shortfalls, taking into account country specificities and the potential need for confidentiality to avoid adverse market reactions. Directors agreed with strengthening the monitoring of key social and distributional aspects wherever possible, including by working with relevant partners such as the World Bank. Some Directors also encouraged a more systematic assessment of distributional considerations in programs.

Directors broadly concurred with Recommendation 2 that Fund-supported programs pay greater attention to supporting deep, more growth-oriented structural reforms, with more effective capacity development (CD) support and more effective collaboration with partners—such as the World Bank—in areas outside the Fund's core mandate and expertise. They reiterated the need to keep structural conditionality parsimonious and prioritized in line with

program objectives, and generally cautioned against veering too far out of core areas.

Directors concurred on the need to assess how CD and surveillance could be better integrated with program design and implementation and looked forward to the conclusions of the ongoing IEO evaluation of Fund CD. While looking forward to the Management Implementation Plan for IMF Collaboration with the World Bank on Macro-Structural Issues, a few Directors encouraged staff and management to propose concrete steps on this matter and to review the experience with World Bank-Fund collaboration in Fund-supported programs.

Directors agreed with Recommendation 3 that the Fund continue to invest in building a toolkit of models and monitors that can be applied as a basis for analysis of the adjustment-growth relationship and assessing growth-related developments in the program context. They welcomed the set of already available models and encouraged staff teams to make use of them on a case-by-case basis and adapt them to better reflect country-specific circumstances. Directors agreed that investing in the Research Department's structural reform database would be helpful.

Directors agreed with the need to avoid "too little and too late" debt restructurings, but they stressed that the Fund has a duty of neutrality, which leaves the design of the restructurings to debtors and their creditors. At the same time, some Directors noted that, in practice, the Fund's debt sustainability analysis provides the Fund with an analytical basis for timely and effective engagement. While a number of Directors considered that there could be greater scope to use exchange rate policy in Fund-supported programs to facilitate adjustment while supporting growth, Directors reiterated the need for a case-by-case assessment of individual country circumstances respecting national decisions on the exchange rate regime. A few Directors saw scope to use the Fund's Integrated Policy Framework to better inform exchange rate discussions between staff and the authorities in a program context.

In line with established practice, management and staff will carefully consider today's discussion when formulating the Management Implementation Plan for Board-endorsed recommendations, including approaches to monitoring progress.

GROWTH AND ADJUSTMENT IN IMF-SUPPORTED PROGRAMS
EVALUATION REPORT 2021