

# IEO

Independent Evaluation Office  
of the International Monetary Fund

## BACKGROUND PAPER

---



BP/21-01/04

### **Structural Conditions, Structural Reforms and Growth in IMF-Supported Programs**

Jung Yeon Kim and Kwang Yeon Lee

**IEO Background Paper**  
Independent Evaluation Office  
*of the* International Monetary Fund

Structural Conditions, Structural Reforms and Growth in IMF-Supported Programs

Prepared by Jung Yeon Kim\* and Kwang Yeon Lee†

June 30, 2021

The views expressed in this Background Paper are those of the authors and do not necessarily represent those of the IEO, the IMF or IMF policy. Background Papers report analyses related to the work of the IEO and are published to elicit comments and to further debate.

---

\* Senior Economist, Independent Evaluation of the IMF.

† Former Research Officer, Independent Evaluation Office of the IMF.

Contents	Page
Abbreviations _____	iv
Executive Summary _____	v
I. Introduction _____	1
II. IMF's Policy on Structural Conditionality _____	2
III. Data and Methodology _____	5
IV. Structural Conditions in Program Design _____	6
A. Volume of SCs _____	7
B. Depth and Content _____	7
V. Structural Conditions in Program Implementation _____	9
A. Key Features of Observed SCs _____	9
B. SC Volume, SC Quality, Country Capacity, and SC Implementation _____	11
C. TA and SC Implementation _____	12
D. Collaboration with Other International Development Institutions _____	16
VI. Structural Conditions and Structural Reforms _____	18
VII. Growth Benefits of SCs in the Medium Run _____	21
VIII. Conclusions and Lessons _____	23
 Figures	
1. Volume of SCs Per Program at Program Design: 2009–19 _____	7
2. Composition of All Structural Conditions _____	8
3. Composition of Observed Structural Conditions _____	10
4. Number of SC, SC Depth and SC Implementation _____	11
5. Country Capacity and SC Implementation _____	12
6. Focus of IMF TA and SC _____	13
7. Country Capacity and IMF TA Receipt _____	14
8. IMF TA and SC Implementation _____	14
9. TA and SC Implementation by Sector _____	16
10. Changes in Structural Reform Index: Pre- versus Post-Program _____	19
11. SC Implementation and Structural Reform _____	20
12. Post-Program Growth Impact of SCs: Completed Programs _____	22
 Tables	
1. Structural Conditions by Depth, Content, and Sector _____	9
2a. SC Implementation Statistics _____	10
2b. SC Implementation Scores _____	11
3. SC Implementation: Fractional Logit Regressions _____	15
4. Average SC Scores: Core versus Shared/Non-Core Expertise _____	17
Reference _____	25

**ABBREVIATIONS**

ASCI	Average Structural Conditions Implementation Score
ASCID	Average Structural Conditions Implementation and Depth Score
ASCIDG	Average Structural Conditions Implementation, Depth and Growth Orientation Score
DPB	Differences in Primary Balance
DPG	Differences in Potential Growths
DSA	Debt Sustainability Analysis
ECF	Extended Credit Facility
EFF	Extended Fund Financing
ESAF	Enhanced Structural Adjustment Facility
FSAP	Financial Sector Assessment Program
FTE	Full-Time Equivalent
GEI	Government Effectiveness Index
GRA	General Resources Account
HIPC	Highly Indebted Poor Countries
IDI	International Development Institutions
IEO	Independent Evaluation Office
LIC	Low-Income Country
MONA	Monitoring of Fund Arrangements
PRGT	Poverty Reduction and Growth Trust
QPC	Quantitative Performance Criteria
RES	Research Department
ROC	Review of Conditionality
SAF	Structural Adjustment Facility
SBA	Stand-by Arrangement
SC	Structural Conditions
SCI	Structural Conditions Implementation
SCID	Structural Conditions Implementation and Depth
SCIDG	Structural Conditions Implementation Depth and Growth Orientation
SCIS	Structural Condition Implementation Score
SPC	Structural Performance Criteria
SR	Structural Reform
SRI	Structural Reform Index
TA	Technical Assistance
WGI	Worldwide Governance Indicator

## EXECUTIVE SUMMARY

This background paper assesses the composition, quality and implementation of structural conditions (SC) in IMF-supported programs over the period of 2008–19, as well as their role in promoting structural reforms and growth in the medium run. To this end, it develops score indices of SCs for each program included in the evaluation sample. This allows for consistent cross-country comparisons and quantitative analysis of the impact of SCs on structural reforms and ultimately, growth.

Key findings are broadly supportive of the role played by SCs in promoting structural reforms and growth after the program, but also suggest some areas for attention in the design and implementation of SCs. Five findings are notable in this regard:

First, SC implementation is strong but the average quality of SCs is relatively low in terms of depth and growth orientation. SC implementation is stronger in General Resource Account (GRA)-supported programs than in Poverty Reduction and Growth Trust (PRGT)-supported programs while the quality of SCs is broadly similar between the two. The PRGT-supported programs focused more on fiscal issues and demand management and less on enhancing growth and efficiency than the GRA-supported programs.

Second, the parsimony and macro-criticality principle seems to have worked relatively well to induce streamlining of SCs but less so for providing incentives to increase the quality of SCs.

Third, the study does not find statistical relationships between IMF's delivery of technical assistance (TA) and either countries' need for technical support or record of SC implementation.

Fourth, only a small fraction of SCs have explicitly recognized the need for collaboration with other partner institutions in shared/non-core areas of IMF's expertise, and the implementation of such SCs has been weaker than that of other SCs.

Fifth, SC implementation was positively associated with progress on structural reforms, especially when the depth and growth orientation of SCs are accounted for. For growth in the medium run, a mere implementation of SCs does not seem to have delivered the desired results: the quality of SCs has mattered significantly for such benefits.

Based on these findings, following broad lessons are drawn for the evaluation:

- First, increasing the depth and growth orientation of SCs would help significantly to promote medium-run growth. Strengthened staff guidance on the design of SCs could contribute to this goal.
- Second, attention needs to be paid to increasing the impact of TA on country's ability to implement SCs.

- Third, given the significant share of SCs in shared/non-core areas of Fund expertise, more effective and structured collaboration with partner institutions, including on providing technical support, seems particularly important.
- Fourth, efforts should continue to improve the quality and user-friendliness of the MONA database.

## I. INTRODUCTION

1. This background paper assesses how well structural conditions (SCs) in IMF-supported programs helped to promote structural reforms and growth in the medium run over the period of 2008–19 as part of the IEO’s broader evaluation on growth and adjustment in IMF-supported programs. Specifically, it examines the implementation status and quality of SCs measured in terms of their depth and growth orientation. It evaluates the role SCs have played in facilitating structural reforms and ultimately, growth.

2. SCs are an important component of IMF-supported programs. Consisting of prior actions (PAs) and structural benchmarks (SBs), SCs comprise about 40 percent of total conditionalities associated with IMF-supported programs and are universal in both GRA and PRGT programs. SCs function as a tool to monitor the progress of policy implementation agreed between the country authorities and the IMF. Unlike quantitative performance criteria, SCs extend beyond macroeconomic measures and include broader reforms, legislation and social policies that would help facilitate macroeconomic adjustment and enhance growth.

3. After a period in which structural conditionality was deliberately being reduced, SCs have gained greater prominence in IMF-supported programs in recent years.<sup>1</sup> During the last decade, prolonged weak growth has become an increasingly serious concern in many countries as global economic environment remained persistently weak. Many countries, as well as the IMF, turned to macro-structural measures in search for growth payoffs, and as a result, SCs became increasingly used in IMF-supported programs. Increasing attention to SCs has accompanied the rising use of the Extended Fund Facility (EFF), a type of arrangement designed to support countries addressing protracted and structural BOP needs among GRA-supported programs.

4. The 2018 Review of Program Design and Conditionality (ROC) noted the increase in number of SCs. The 2018 ROC recommended that SCs be better prioritized, that Fund should continue building expertise in shared areas of responsibility, and that Fund should use more realistic conditionality implementation timetables

5. This paper presents data on the key characteristics of SCs across programs and provides a quantitative analysis of the role of SCs in achieving successful program outcomes. To this end, it develops a new methodology to quantify progress on implementation and quality of SCs. Based on this methodology, each SC of 131 sample programs are assessed, classified, and numerically scored. This enables cross-country comparison and econometric analysis to examine the benefits of SCs on achieving structural reforms and post program growth. It also assesses the contribution of IMF technical assistance (TA) and inter-agency collaboration in the design and implementation of SCs.

---

<sup>1</sup> See *Structural Conditionality in IMF-Supported Programs – Evaluation Update* (IEO, 2018).

6. The paper is organized as follows. Section II reviews briefly the IMF's policies on SCs and how they have evolved over time. Section III develops a SC scoring methodology which allows for numerical representation of qualitative characteristics of each SC such as their depth, growth orientation, and implementation status.<sup>2</sup> An index of SC implementation scores (SCIS) constructed using this methodology enables consistent cross-country analyses and comparisons. Section IV assesses key features of SCs in program design, and Section V discusses SCs in program implementation as well as the underlying conditions conducive for stronger SC implementation. The role of IMF TA and collaboration with other international development institutions (IDIs) are also discussed here. Section VI examines the linkages between SC implementation and progress on structural reforms to assess how SCs have translated into progress in structural reforms. Specifically, it focuses on the relationship between SCIS and the Structural Reform Index (SRI) developed by IMF's Research Department. This is followed by Section VII which discusses the impact of SC on post-program growth. Section VIII concludes with policy implications.

## **II. IMF'S POLICY ON STRUCTURAL CONDITIONALITY**

7. The use of SCs in IMF-supported programs has evolved through time. Structural Performance Criteria (SPC) first appeared with the introduction of the Extended Fund Facility (EFF) in 1974, which was intended to support reforms for medium-term adjustment. However, until the early 1980s, IMF conditionalities were dominated by quantitative performance criteria on fiscal, monetary and external policy outcomes. The 1979 Conditionality Guidelines stipulated that performance criteria should relate to other (non-macroeconomic) variables only in "exceptional cases." Formal conditionality related to structural issues remained limited.

8. Since the mid-1980s, the scope and complexity of program conditionality has expanded to areas well beyond the IMF's traditional macroeconomic focus, to structural issues. This reflected various developments. Fund-supported programs were increasingly criticized for excessive "austerity" and "demand restraint" at the cost of growth. Lack of growth in low-income countries (LICs) in 1980s became a central topic, and many agreed that the revival of growth was essential for LICs to grow out of debt problems and gain market access. At the same time, there was also growing interest in policies reducing the role of the state in economic activity and shifting toward market-based economies. Conditionalities such as those related to reducing subsidies, market liberalization and privatization of public enterprises became widely used. The Executive Board, at the time of 1987 conditionality review, agreed that "a growth-oriented strategy called for a greater emphasis on stronger and more specific structural measures." They also indicated that conditionality should be applied to structural policies where they were critical for a program's objectives.

---

<sup>2</sup> The term "content," "growth orientation," and "growth/efficiency orientation" are used interchangeably in this paper.

9. During this period, the Fund expanded its involvement in LICs and transition economies, contributing to the increasing use of SCs. The Fund was seeking ways to provide financial support to LICs with greater emphasis on reforms needed for stronger medium-term performance. The Structural Adjustment Facility (SAF) and the Enhanced Structural Adjustment Facility (ESAF) were set up in 1986 and 1987, financed from special trust fund resources, and offered highly concessional loans to support macroeconomic adjustments and structural reforms in LICs. All SAF- and ESAF-supported arrangements included SCs from the beginning, and the average number of SCs in these programs increased substantially through the late 1990s. The accession of IMF membership by transition economies also contributed significantly to the growing role of SCs in IMF-supported programs. Twenty-two centrally planned economies sought IMF programs to assist their transition into market-based economies in the late 1980s and early 1990s. Initially, not all these programs involved SCs, but by 1994, SCs became universal in all arrangements. In 1987, Fund-supported programs contained, on average, two SCs per program year. By 1994, this number had increased to 7, rising further to an average of 14 SCs per program year in 1997–99.<sup>3</sup>

10. The Asian crisis in the late 1990s was a turning point for IMF's policy on SC. SCs proliferated in Korea, Thailand, and Indonesia's programs. In part this reflected the complexity of the financial sector restructuring that required corporate sector restructuring and related reforms in bankruptcy and other relevant legislation. The number of SCs in Asian crisis programs averaged over 22 per program year. However, the experience drew wide criticism of the effectiveness of these measures and their adverse impact on the country ownership of programs. The external as well as internal criticism on excessive use of SC prompted efforts to begin refocusing SC in 2000. An Interim Guidance Note on *Streamlining Structural Conditionality* was issued that year which called for a more focused and parsimonious application of conditionality for structural reforms, discouraging SC that may be relevant but not critical to the program's objectives. In 2002, the Executive Board adopted new *Guidelines on Conditionality* (IMF, 2002a), the first revision of the Fund's conditionality guidelines since 1979. The 2002 Guidelines emphasized national ownership of policies, parsimony in conditions, tailoring policies to circumstances, coordination with other multilateral institutions, and clarity in the specification of conditions.

11. Since the 2002 Guidelines were issued, it has been revised four times, generally to encourage more focused use of SCs. A revision following the 2004–05 ROC underlined the importance of program ownership. In 2008, the IEO evaluation of *Structural Conditionality in IMF-supported Programs* (IEO, 2008) found that SCs were used extensively and that program documents were not sufficiently clear about linkages to program goals, notwithstanding the 2002 Conditionality Guidelines' emphasis on "parsimony" and "criticality." It also concluded that most SCs had little structural depth, only half were implemented on time, and compliance was only weakly correlated with subsequent progress in structural reform. This led to the 2008 revision of Conditionality Guidelines which stressed the need to strengthen parsimony by emphasizing criticality as well as requiring rigorous justification of conditionality. The 2010

---

<sup>3</sup> See *Structural Conditionalities in Fund-Supported Programs* (IMF, 2001).

revision reflected the Board's decision to reform the Fund's conditionality framework in 2009 to discontinue structural performance criteria (SPC) altogether and to shift to structural benchmarks (SBs), placing greater reliance on a review-based approach to monitor structural reforms in Fund-supported programs. Finally, following the 2011 ROC, the 2014 revision primarily focused on incorporating conditionality related to macro-social (also called jobs and growth) issues, better leveraging surveillance and TA in program design, and improving partnerships with other institutions, particularly regional financing arrangements. In addition, the revision added guidance on the review-based approach to monetary policy conditionality in countries with inflation-targeting frameworks or evolving monetary policy regimes.

12. The 2011 ROC found that during 2002–11, SCs became more parsimonious (IMF, 2012a). SCs were also generally well-focused on macro-critical areas of core Fund responsibility. However, the 2018 ROC reported that during 2011–17, the number of SCs increased, reflecting that programs during this period were increasingly dealing with protracted structural challenges in a weak global environment (IMF, 2019b). Conditionality remained largely focused on the Fund's core areas of responsibility, even though critical reforms were needed in shared (e.g., labor and product market reforms) and non-core areas. There were also increased delays in the implementation of program conditionalities. The 2018 ROC recommended that SCs be better prioritized, that Fund should continue building expertise in shared areas of responsibility and that Fund should use more realistic conditionality implementation timetables. The 2018 IEO Update found that the progress was limited in increasing program ownership or reducing stigma. It also found that Bank-Fund cooperation on SCs needed to be strengthened to make collaboration less personality driven and more substantive and systematically effective. Program documents could also do a better job at explaining the link between SC and achievement of program goals.

13. There are some important facts and practices to note about SCs. First, while both quantitative performance criteria (QPC) and SCs are program conditionalities, they do not carry the same weight. If a country fails to meet a QPC, an explicit IMF Executive Board waiver is required to allow disbursement of Fund resources.<sup>4</sup> By contrast, SBs are generally not directly linked to continuation of disbursement of Fund resources. Delays in the implementation of a specific SB does not necessarily hold up the completion of a program review but delays in implementation would figure importantly in deliberations to complete a review. Second, SCs do not necessarily deal with macro-structural issues. In fact, about two-third of all SCs are in the fiscal sector, and the macro-structural issues cover only about 10 percent of total SCs. Third, in many countries, SCs are accompanied by TA because the authorities often need support to design and implement these measures. Since SCs often go beyond core topics of IMF's expertise, monitoring and TA delivery need to be well coordinated with other development partners.

---

<sup>4</sup> The waiver may be issued if the Board is satisfied that the program will still succeed. This may be because the deviation was minor or temporary or because national authorities are taking corrective actions. SPCs, like QPCs, required waivers for non-compliance when they were used in IMF arrangements prior to 2010.

### III. DATA AND METHODOLOGY

14. This study covers SCs included in 131 Fund arrangements—55 GRA and 76 PRGT and GRA-PRGT blended arrangements for 74 countries in total—approved and scheduled for completion between September 2008 and March 2020. Data are mostly taken from the MONA and World Economic Outlook (WEO) databases.<sup>5</sup> Other data used are taken from the Structural Reform Index (SRI) database compiled by the IMF’s Research Department, Travel Information Management System (TIMS) by the IMF, and the World Development Indicator and Worldwide Governance Indicators published by the World Bank.

15. One of the challenges in assessing structural conditions in IMF-supported programs is that SCs are stated in qualitative terms and thus hard to quantify. It is thus more difficult to measure and compare SCs across programs and assess their impact on macro variables than a QPC. SCs are also more challenging to monitor, especially in countries where data transparency, reporting and capacity constraints are major concerns. To overcome these challenges, this paper has adopted a new methodology to assess, classify and score numerically the quality and implementation of SCs. Specifically, it adopts the methodology developed in the 2008 IEO evaluation of SCs to measure the depth of a SC.<sup>6</sup> Then this paper adds a new dimension, the growth/efficiency orientation (or content) of SCs. Finally, it also incorporates a third dimension which reflects the SC implementation status. Using these three dimensions that reflect both the quality and quantity aspect of SC implementation, a Structural Conditionality Implementation Score (SCIS), or SC index, is assigned to each SC of every IMF-supported programs included in this evaluation.<sup>7</sup>

16. The details of the three aspects of SCs, i.e., implementation status, depth, and content, are as follow.

- *Implementation status* indicates whether SCs were met, met with delay, or not met. Numerical values of 1.0, 0.5, and 0.0 are assigned to each category, respectively. The underlying data for implementation status are taken from the MONA database which tracks the performance of countries in terms of scheduled purchases and reviews, quantitative and structural conditionality, and macroeconomic indicators.
- *Depth* refers to the degree of structural change that a SC would bring about if implemented, and its effectiveness in bringing about subsequent reforms. In general, SCs with higher depth are likely to be more challenging to implement. Numerical values of 1.0, 0.66, and 0.33 are assigned to SCs that were assessed as high, medium and low

---

<sup>5</sup> However, when program reviews are delayed or program goes off track, the progress of SC implementation is not tracked or recorded in IMF’s Monitoring of Fund Arrangement (MONA) database.

<sup>6</sup> See *Structural Conditionality in IMF-Supported Programs* (IEO, 2008).

<sup>7</sup> The terms “score” and “index” are used interchangeably in this paper.

depths, respectively.<sup>8</sup> The depth scores are mostly based on data put together in the 2018 ROC which also relied on the methodology developed by the IEO's 2008 evaluation on structural conditionality.

- *Content (or growth orientation)* is a classification in the MONA database which indicates whether the SC is intended primarily for: (a) enhancing growth and economic efficiency that would help the economy adapt better to changes in economic conditions (e.g., trade reforms or pricing policies in factor markets); (b) managing vulnerabilities and addressing actual and potential balance sheet risks (e.g., measures for strengthening prudential regulations and financial sector supervisory capabilities); or (c) facilitating adjustments in aggregate demand and underpinning the stabilization efforts by enhancing the functioning of fiscal, monetary, and exchange rate policies (e.g., improving tax structure).<sup>9</sup> Numerical values of 1.0, 0.66 and 0.33 are assigned to categories (a), (b), and (c), respectively. Given the higher values assigned to SCs that are growth and efficiency oriented, the terms "content," "growth/efficiency orientation," and "growth orientation" are used interchangeably.

17. Based on the score for each aspect of SCs, three aggregate SC score indices, denoted by SCI, SCID and SCIDG, are developed. Specifically,  $SCI_j$  refers to the score for implementation status only for structural condition  $j$ .  $SCID_j$  is constructed by multiplying  $SCI_j$  by the depth score  $D_j$  (i.e.,  $SCID_j = SCI_j \times D_j$ ). Similarly,  $SCIDG_j$  is constructed by multiplying  $SCID_j$  by the growth orientation score  $G_j$  (i.e.,  $SCIDG_j = SCID_j \times G_j = SCI_j \times D_j \times G_j$ ). The aggregate score indices, SCI, SCID and SCIDG, are then constructed by aggregating  $SCI_j$ ,  $SCID_j$  and  $SCIDG_j$  across all structural conditions, respectively, for a particular program. Finally, ASCI, ASCID and ASCIDG stand for the average score indices for a program, which are constructed by dividing the corresponding aggregate score indices by the total number of structural conditions.

#### IV. STRUCTURAL CONDITIONS IN PROGRAM DESIGN

18. This section presents cross-country analysis on various aspects of the use of SCs in program design. This analysis distinguishes between two data sets: a total data set that covers all SCs included in program design regardless of whether implementation status is known, and an observed data set that include only SCs whose implementation is known and classified as "met," "met with delay," or "not met." About 28 percent of SCs are excluded from observed SCs because their implementation status remains uncertain because the program is cancelled or terminated

---

<sup>8</sup> 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" (Gambia 2012 ECF). Assessment of depth of SCs that were not covered in the 2018 ROC were undertaken by the IEO staff.

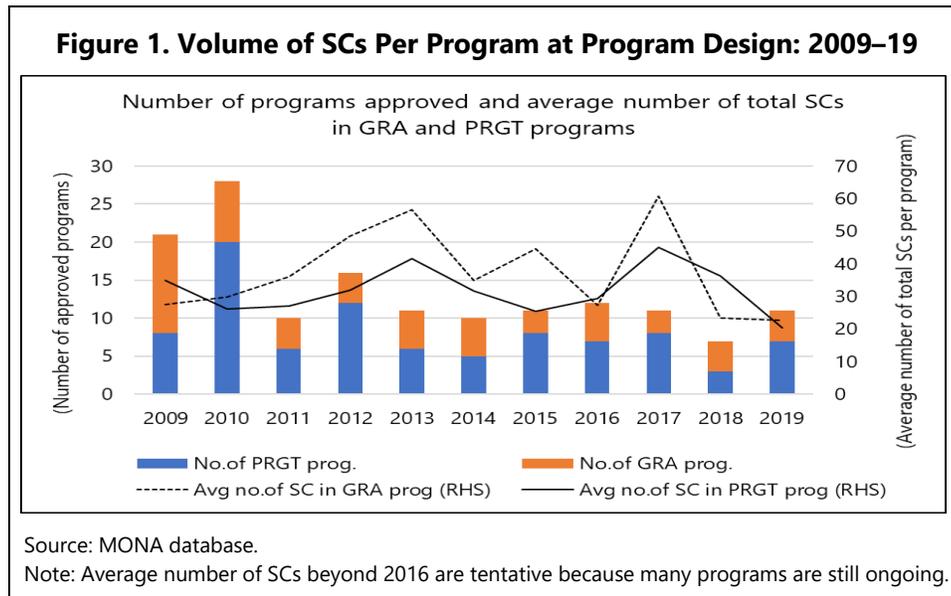
<sup>9</sup> This classification follows the methodology used in the 2004 ROC and accompanying Occasional Paper. See 2005 IMF Occasional Paper, *The Design of IMF-Supported Programs*.

without completion of the corresponding review. The universal set of SCs is used to assess program design and observed SCs are used to assess SC implementation and the impact of implemented SCs on structural reforms and growth.

### A. Volume of SCs

19. SCs constitute 42 percent of total program conditionality. The 2018 ROC, which covers the period of 2012–17, found that compared to the 2011 ROC sample period (2007–11), the volume of SC increased significantly, reflecting in large part a gradual shift in GRA programs from SBA to EFF arrangements which focus more on tackling structural problems than SBAs and make greater use of SCs.

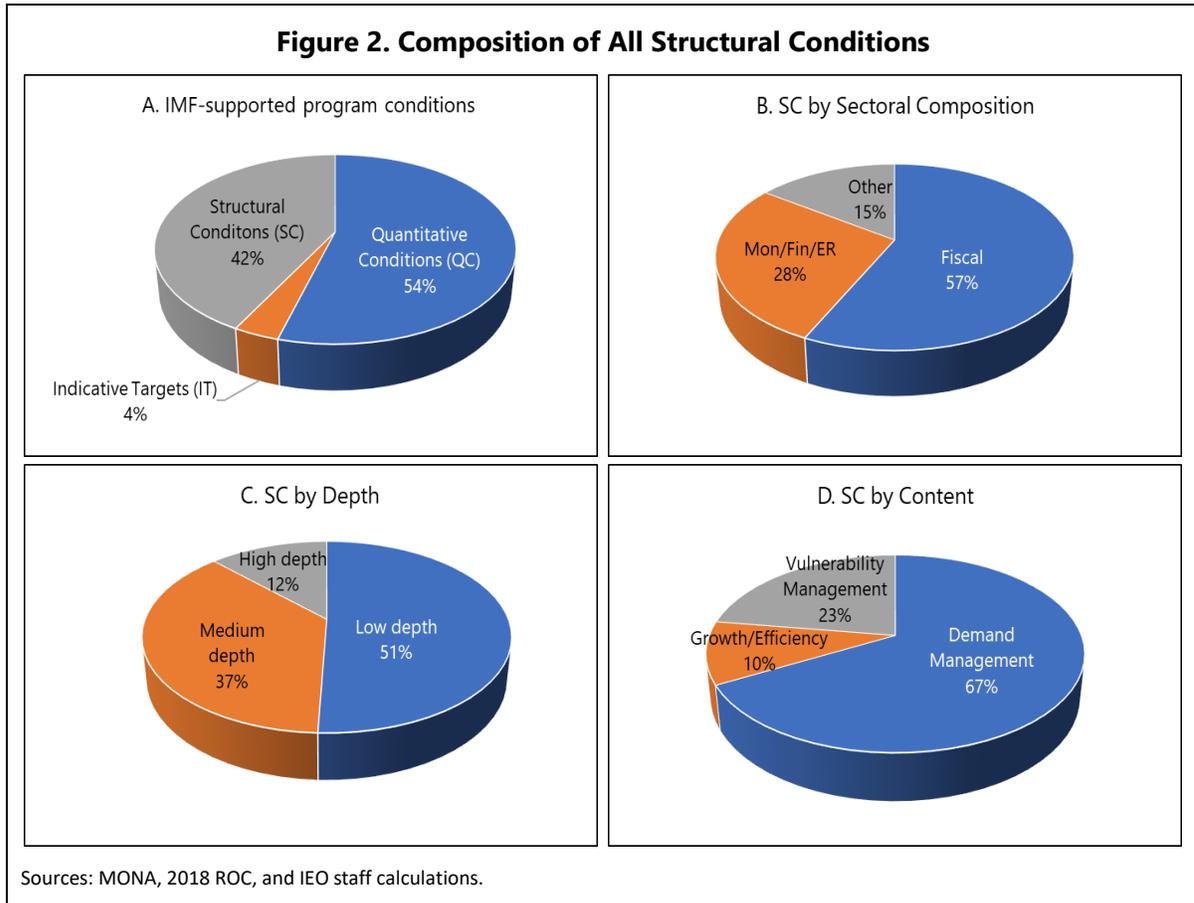
20. Figure 1 shows that the number of new programs approved fell sharply after 2010 for both GRA and PRGT programs, but that average number of SCs increased rapidly, peaking in 2013. There was a high prevalence of programs with 50 SCs or more, particularly during 2011–14. For example, Greece (2012 SBA) had 97 SCs, Pakistan (2013 SBA) 82 SCs, Cote d’Ivoire (2011 ECF) 72 SCs, Liberia (2012 ECF) 66 SCs, Jamaica (2013 EFF) and Albania (2014 EFF) 65 SCs, and Guinea (2012 ECF) had 61 SCs. GRA programs had a higher average number of SCs during most of the period. The average number of SCs per program returned to the 2009 level by 2016.



### B. Depth and Content

21. In practice, SCs are mostly of low to medium depth, related to demand management, and in the fiscal area (Figure 2). On average, 51 percent of SCs are of low depth, about 37 percent are of medium depth, and only 12 percent of SCs are of high depth. In terms of content, about two-thirds of the SCs are for demand management, about a quarter are for vulnerability management, and the remaining 10 percent are for growth and efficiency. In terms

of economic sector, SCs are highly concentrated in the fiscal area accounting on average for about 60 percent of total SCs.<sup>10</sup> About 30 percent are on monetary/financial/exchange rate issues and the remainder are on other structural issues. About 70 percent of SCs are in the area of the IMF's core expertise and the remaining 30 percent are in non-core areas or areas of shared expertise with other international development institutions (IDI).<sup>11</sup>



22. There are clear differences between GRA and PRGT-supported programs in terms of depth, content, and sectoral composition of SCs (Table 1). GRA-supported programs have a significantly higher share of high depth SCs (16 percent) than PRGT programs (9 percent). They also have a higher share of SCs related to both vulnerability management and growth/efficiency (40 percent in combined total) than PRGT programs where the corresponding share is only around 27 percent. In particular, PRGT programs have about 40 percent less growth/efficiency oriented SCs than GRA programs. The sectoral composition of SCs broadly mirrors the content composition. GRA programs have a significantly higher share of SCs covering non-fiscal

<sup>10</sup> Includes SCs in civil service and pension reform.

<sup>11</sup> Indeed, 9 out of 17 country case studies covered in this evaluation (i.e., Ghana, Grenada, Jamaica, Jordan, Latvia, Malawi, Pakistan, Romania, and Ukraine) explained that from the start, the IMF-supported programs did not include reforms outside its core areas and/or focused mostly on fiscal issues.

macroeconomic policies and structural issues (51 percent) than PRGT programs (36 percent). In PRGT programs, fiscal SCs dominate (64 percent), which are mostly for demand management in content as well as actions related to revenue collection, debt management, civil service reform, and fiscal transparency. It is notable that GRA programs account for 46 percent of total SCs (and 51 percent of SCs including those that are not observed SCs).

**Table 1. Structural Conditions by Depth, Content, and Sector  
(In percent of total)**

	Depth			Content			Sector		
	High	Medium	Low	Demand Control	Growth/ Efficiency	Vulnerability Mngmt	Fiscal	Mon/Fin/ER	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	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

Sources: MONA, 2018 ROC, and IEO staff calculations.

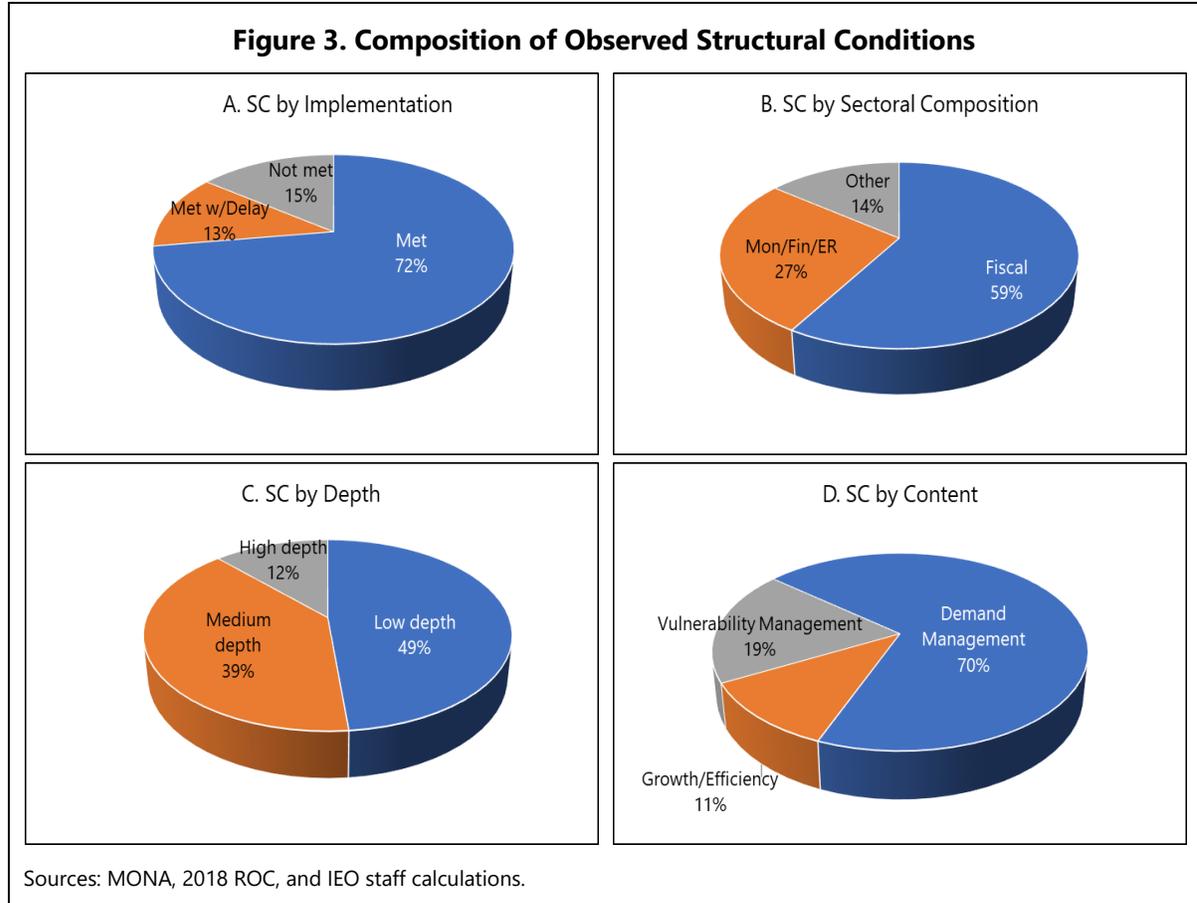
## V. STRUCTURAL CONDITIONS IN PROGRAM IMPLEMENTATION

23. This section focuses on the set of SCs whose implementation status is known based on the MONA database (Figure 3). It presents a broad picture of outcomes of SC implementation by program type (or level of income) and region. It also presents the volume and the composition of observed SCs by depth, content and sector, and discusses features related to successful implementation of SCs. There is a sizable overlap in the SCs assessed in this section with those considered in the 2018 ROC since the latter also focused only on observed SCs. This section draws on the 2018 ROC in many respects and further expands the analysis to assess how SC implementation has been influenced by country capacity and supporting policies including provision of TA and the IMF's collaboration with other IDIs.

### A. Key Features of Observed SCs

24. Based on observed SCs, GRA programs have on average strong implementation records with over 90 percent of SCs being met or met with delay. The corresponding figure for PRGT programs is 80 percent (Tables 2a and 2b). By region, implementation of SCs is strongest in countries in WHD and EUR. In terms of depth, it is noteworthy that the share of high-depth SCs is *lower* in GRA programs (11 percent) than PRGT programs (13 percent). Given that the opposite was true for the total set of SCs (see Table 1), this result implies that a higher percentage of high-depth SCs have remained incomplete in GRA programs that were terminated without the completion of all reviews. For growth orientation and sectoral composition, the results are similar to those of the full set of SCs. In particular, PRGT programs have much lower share of SCs aimed

at enhancing growth than GRA programs, and fiscal sector SCs clearly dominate.<sup>12</sup> It is also notable that despite the increased attention to inclusive growth and poverty reduction, less than 1 percent of total observed SCs (21 and 6 SCs in GRA and PRGT programs, respectively) are related to such objectives.



**Table 2a. SC Implementation Statistics  
(In percent of total)**

Program Type (No. of programs)	Implementation			Depth			Content			Sector		
	Met	Met w/ Delay	Not Met	High	Medium	Low	Demand Mngmt	Growth/ Efficiency	Vulnerability Mngmt	Fiscal	Mon/Fin/ ER	Other
GRA (53)	77	14	10	11	39	50	63	14	23	50	32	18
PRGT (74)	69	12	19	13	40	47	75	9	16	65	23	11

Sources: MONA, 2018 ROC, and IEO staff calculations.  
Note: Total number of programs are fewer because in some programs, all SCs were incomplete/outstanding.

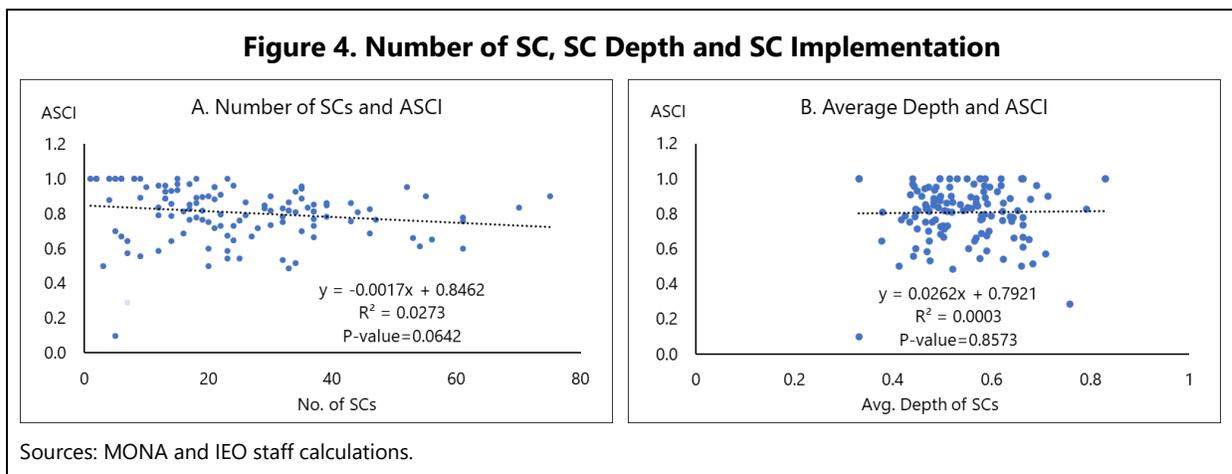
<sup>12</sup> Data showed that the fiscal SCs and SCs for demand management dominated the PRGT-supported programs throughout most of the evaluation period.

	<b>Implementation</b>	<b>Depth</b>	<b>Content</b>
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

Sources: MONA, 2018 ROC and IEO staff calculation.  
 Note: Numbers in parenthesis present the total number of programs for relevant category. Implementation, Depth and Content figures are average scores normalized by number of SCs.

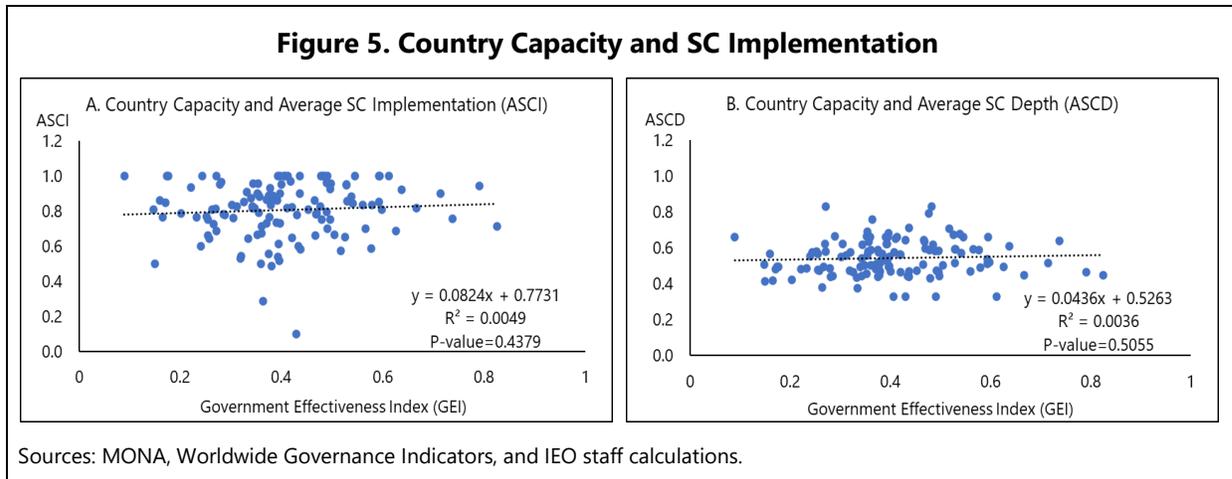
## B. SC Volume, SC Quality, Country Capacity, and SC Implementation

25. Bivariate regression analysis shows that a higher volume of SCs is negatively and statistically significantly associated with the average SC implementation scores (ASCI), supporting the merit of the parsimony and macro-criticality principle (Figure 4A). Interestingly, no statistically significant relationship is detected between the average depth of SCs and the average implementation scores (Figure 4B).<sup>13</sup> This lack of relationship between average depth and implementation does not seem to be an artificial artefact arising from non-observed SCs being excluded from the data set because the composition of outstanding SCs by depth is broadly identical to that of observed SCs (see Table 1). Finally, no statistically significant relationship is detected between the average scores for growth orientation and implementation.



<sup>13</sup> SCs with higher (lower) depth would normally be more (less) challenging to implement than lower (higher) depth SCs because they require more (less) planning, stronger (less) capacity and stronger (less) political will.

26. This study also does not find a statistically significant bivariate relationship between countries' institutional/human capacity and SC implementation. In the analysis, capacity is measured by the level of the Government Effectiveness Index (GEI) published by the World Bank.<sup>14</sup> A reasonable prior would be that all else equal, countries with higher (lower) capacity would be expected to perform better (worse) in SC implementation. However, while the bivariate relationship between average SC implementation score (ASCI) and capacity has the expected sign, it is not statistically significant (Figure 5A). Similarly, no statistically significant bivariate relationship is observed between capacity and the average depth of SCs (ASCD) (Figure 5B).



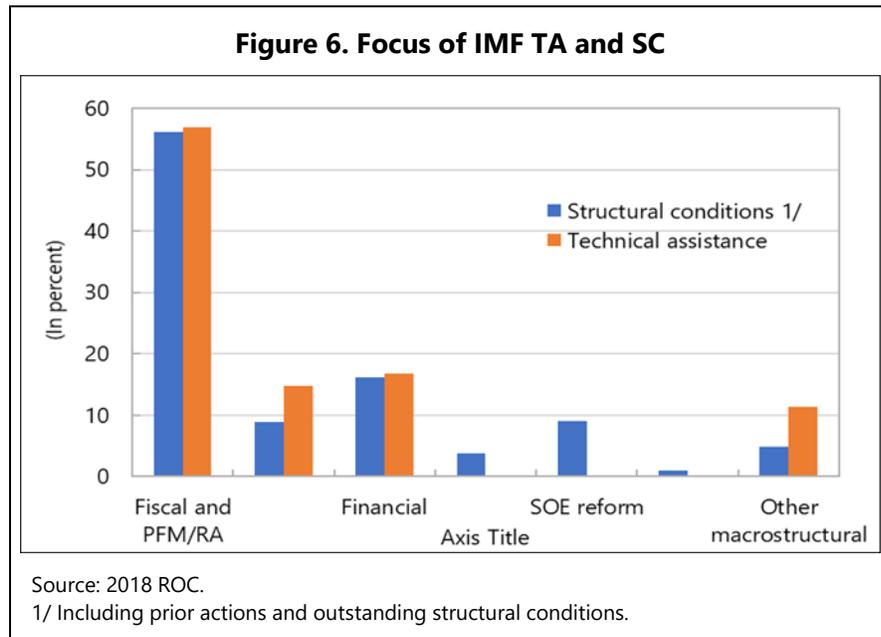
### C. TA and SC Implementation

27. Achieving greater integration among TA, surveillance and lending operations has been an important policy agenda for the IMF. Provision of TA should help countries to build needed human and institutional capacities which in turn contribute to program success. Conversely, IMF-supported programs help to identify the countries and the areas which provision of TA can have the biggest impact.

28. This section assesses how provision of IMF TA has been aligned with country need in the program context and how it has affected SC implementation. Specifically, it compares the sectors/topics of SCs with that of IMF TA delivered to these countries. It also provides evidence on to what extent the countries who need TA the most received more TA than others. Finally, it assesses whether IMF TA was positively associated with SC implementation. Data on TA are based on the IMF's Travel Information Management System (TIMS) and reflect TA delivered in the field.

<sup>14</sup> The Government Effectiveness Index (GEI) is one of the six dimensions of governance that constitutes the Worldwide Governance Indicators (WGI) developed by the World Bank. GEI captures perceptions of the quality of public services and civil service, the independence of the government from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.

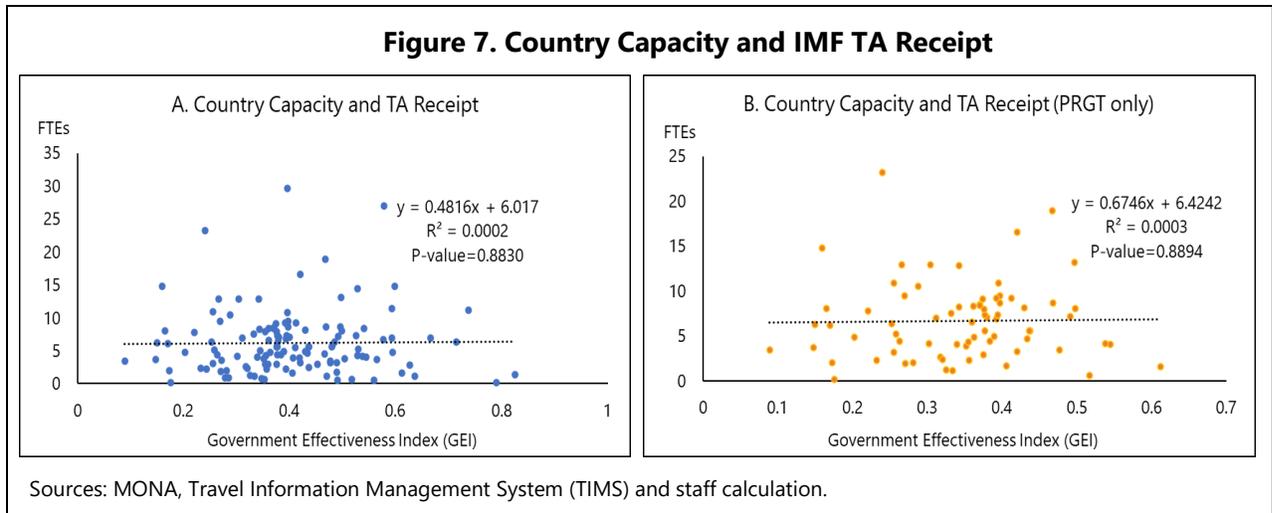
29. The 2018 ROC found a high degree of overlap between SCs and IMF TA, especially in the areas of the IMF's core expertise, although not in areas outside the Fund's core expertise (e.g., pension, SOE reform, and other macro-structural areas). This suggests that the focus of SC and TA have been broadly consistent (Figure 6). Relatedly, the IMF's 2018 Capacity Development (CD) Strategy Review stated that integration between IMF's capacity development (i.e., TA and training) and lending are relatively strong, and that CD is often integral to program framework.



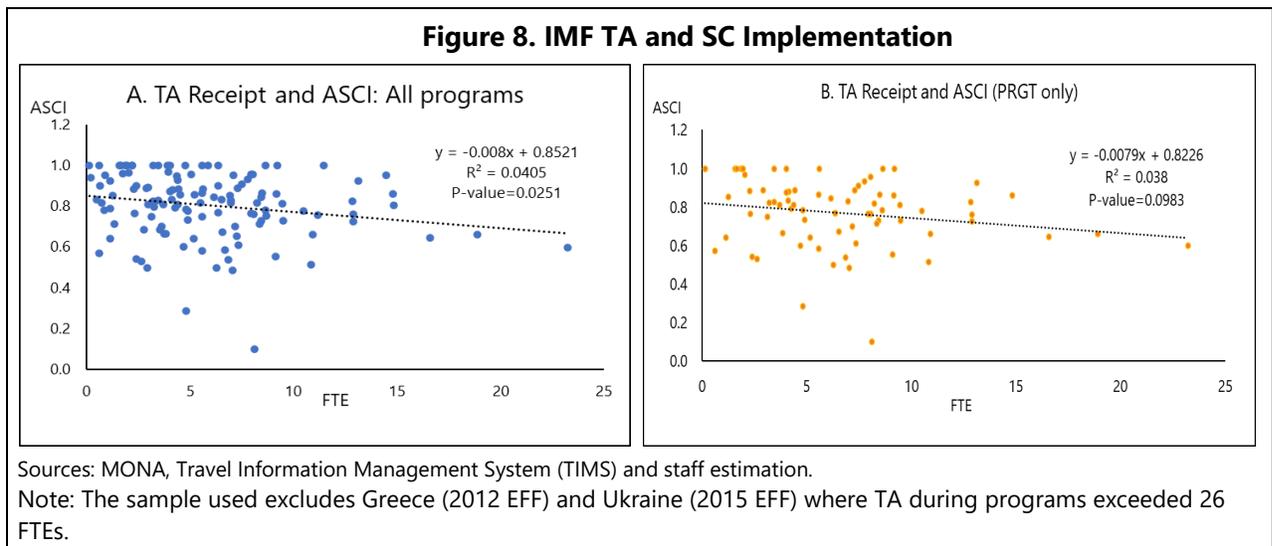
30. Nevertheless, the data for this evaluation do not show a close relationship between TA and either countries' needs or success in implementing SCs. First, the bivariate relationship between country capacity (measured by GEI) and the amount of IMF TA (received during the program period, measured in FTE terms) is weakly positive and not statistically significant (Figures 7A and 7B). This result suggests that lower capacity countries have on average not received more TA than countries with higher capacity.<sup>15</sup> Data shows that among the PRGT programs, the frontier LICs and countries that had issued Eurobonds at least once accounted for the majority of the top recipients of IMF TA.<sup>16</sup>

<sup>15</sup> The data for TA provision do not include other CD modalities, such as classroom and online training. The allocation of Fund CD resources appears to have been guided by multiple considerations and not just country needs or capacity. The 2019 Board document, *IMF Policies and Practices on Capacity Development (CD)* states "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." 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.

<sup>16</sup> Frontier LICs include Bangladesh, Bolivia, Cote 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 Republic of Congo, Ethiopia, Honduras, and Rwanda. See IMF (2015b).



31. Second, the bivariate relationship between total IMF TA during the program and the average SC implementation score is negative and clearly statistically significant (Figure 8A). Similar regression was conducted only for the PRGT programs where the TA needs are arguably greater than in GRA programs (Figure 8B). The results were similar, showing a negative relation albeit at a higher significance level (10 percent). Analysis also showed that there was no statistically significant bivariate relationship between IMF TA and the depth of SCs.



32. While this result may reflect in part that TA is more likely to be provided in areas where the issues are most pressing and where SCs are more challenging to implement, the lack of positive and statistically significant relationship between IMF TA and SC implementation continues even after controlling for other factors that could affect SC implementation, such as the average depth of SC, the recipient country's implementation capacity (measured by the government effectiveness index published by the World Bank), and the total volume of SCs. Specifically, the results of multivariate fractional logit analysis show that the relationship between

IMF TA and SC implementation remains statistically insignificant in both GRA and PRGT programs and continues to be negative in the latter (Table 3).<sup>17</sup> Overall, this finding is suggestive of the need to revisit the effectiveness of TA in supporting program implementation.<sup>18</sup>

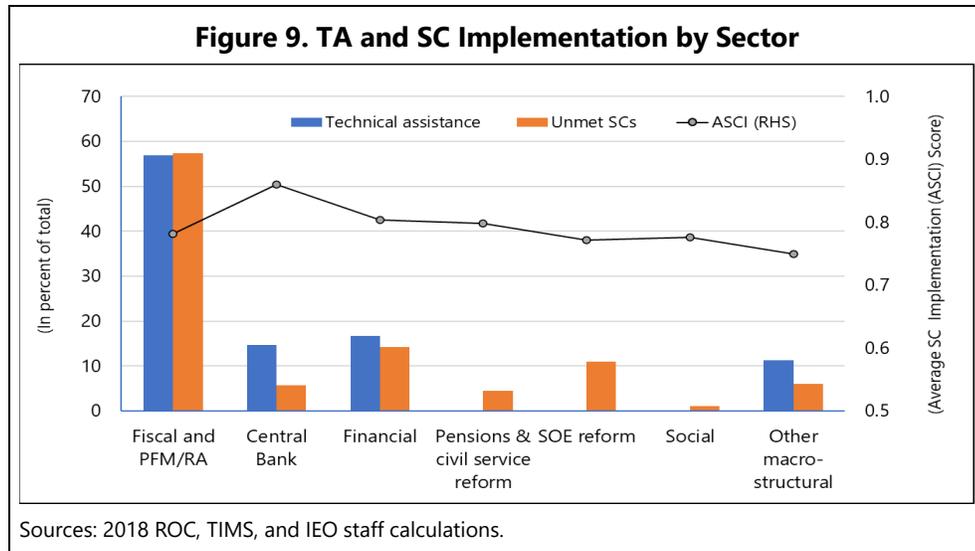
	All Programs			Completed Programs		
	Full (1) ASCI	GRA (2) ASCI	PRGT (3) ASCI	Full (4) ASCI	GRA (5) ASCI	PRGT (6) ASCI
Total TA	-0.014 (0.008)	0.024 (0.006)	-0.033 (0.012)	-0.01 (0.009)	0.021 (0.017)	-0.013 (0.013)
Avg Depth	0.072 (0.549)	-1.006 (0.626)	0.386 (0.752)	-2.384* (0.535)	-2.676 (0.804)	-2.63* (0.709)
GEI	0.152 (0.051)	-0.126 (0.074)	-0.088 (0.12)	0.229* (0.054)	-0.099 (0.08)	0.324 (0.117)
No. of SCs	-0.009 (0.003)	-0.017** (0.003)	-0.006 (0.005)	0.001 (0.003)	-0.005 (0.004)	0.008 (0.006)
Constant	1.78** (0.321)	2.641*** (0.356)	1.3 (0.457)	2.649*** (0.299)	3.025*** (0.421)	2.588*** (0.416)
N	126	53	73	82	32	50
Pseudo-R <sup>2</sup>	0.037	0.082	0.04	0.09	0.059	0.095

Sources: MONA database, Worldwide Governance Indicators and IEO staff estimates.  
 Note: ASCI and Avg Depth denote the average SC implementation and average depth scores of SCs, respectively. TA is measured in FTE units. GEI refers to the 3-year average prior to the program of the government effectiveness index published by the World Bank. Robust standard errors are in parentheses. \*p<0.1, \*\*p<0.05, \*\*\*p < 0.01.

33. The sectoral breakdown of TA and SC implementation shows that fiscal sector, which dominates program SCs and IMF TA allocations, also dominates the group of unmet SCs, while SC implementation is the strongest in the central bank and monetary sector (Figure 9). It is noteworthy that the SC implementation scores are broadly similar across sectors except for the central bank (and monetary policy) sector. In contrast, there is stark difference in the amount of TA allocated to different sectors, with TA being heavily tilted towards fiscal area. This raises concerns on how well TA has been integrated with IMF-supported programs beyond the program design stage, i.e., in program implementation and monitoring. While sectoral composition of SCs and the TA appear to be consistent at the time of program design (see Figure 6), the accompanying background paper “Fiscal Adjustment and Growth in IMF-Supported Programs” (Gupta, 2021) explains that the IMF’s Fiscal Affairs Department (FAD), the department responsible for fiscal TA, has little role in assessing compliance with fiscal SCs. The paper explains that based on the review of back-to-office reports for the 17 case study countries in this evaluation, IMF’s fiscal TA missions do not discuss the status of fiscal reforms in IMF-supported programs that are increasingly drawn from TA advice, pointing to a gap in integration of IMF TA with programs.

<sup>17</sup> Gupta (2021) reports the similar results for the relationship between fiscal TA and fiscal SC implementation.

<sup>18</sup> The upcoming IEO evaluation on the IMF and Capacity Development will take up these issues in greater detail.



34. These results also raise questions on whether the TA is being delivered in the most effective way in the program context. The mode of delivery is especially important for lower-income countries with weak capacities because these countries often operate on already thinly stretched human capital. This requires the Fund to explore ways to deliver TA in a way most accommodating for the authorities. In this regard, the IEO evaluation on *The IMF and Fragile States* (IEO, 2018), as well as the country case study on Malawi in this evaluation found that TA provided through a long-term resident advisor was more effective and preferred by the authorities than visits by short-term experts or remote learning. This highlights the importance of not only the quantity but also how TA is planned, delivered and followed up. Increased interaction between TA experts, country teams and authorities should aid in better design and implementation of SCs.

35. Findings from country case studies also provide further insights on how the integration of TA and programs can be improved. For example, studies on Grenada, Jamaica, Jordan, Pakistan, and Tunisia indicated that the provision of TA was not a substitute for capacity, and that even with extensive TA, implementation of structural reforms required domestic capacity and political and social support to generate the desired outcomes. They stressed that program reforms should avoid excessively accelerated timelines, and called for greater selectivity in program design, better contingency planning and more cautious assumptions on capacity constraints and feasibility of structural reforms. They called for more patience for setting timelines for in improving capacity and longer duration of programs.

#### **D. Collaboration with Other International Development Institutions**

36. Where SCs are being set in areas outside of IMF's core competence, inter-organizational collaboration becomes relevant for program success. About two-third of SCs are in core areas of IMF's expertise and the remaining one-third are in shared and non-core areas where other international development institutions (IDIs) may lead in terms of knowledge and experience.

Table 4 shows that the average SC implementation score is lower in the shared and non-core areas than in core areas of IMF's expertise. The opposite holds for depth and growth orientation, which are modestly higher in shared/non-core areas in all programs combined. In particular, the growth orientation is significantly higher in shared/non-core areas than in core areas.

		Implementation		Depth		Growth orientation	
Core	GRA	0.85	(0.80)	0.53	(0.54)	0.4	(0.39)
	PRGT	0.76		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, 2018 ROC, and IEO staff calculations.  
Note: Figures in parentheses are the averages of GRA and PRGT programs.

37. Focusing on the country case studies undertaken for the evaluation, program documentations for the majority of programs studied, especially PRGT programs, mentioned the need for and/or the ongoing collaboration with other IDIs (e.g., the World Bank) for structural reforms efforts in areas such as social safety net, energy sector reform, financial inclusion, etc.<sup>19</sup> However, fewer than 2 percent of total SCs in the entire evaluation sample were explicit about the collaboration with other IDIs. In this small sub-sample of SCs, the implementation rate was even lower, especially among PRGT programs, suggesting that collaboration has on average not worked well in supporting implementation of SCs in shared/non-core areas, particularly in PRGT programs (Table 4).

38. The need for stronger engagement with other IDIs in the areas outside the Fund's core expertise was a common theme in the majority of country case studies. For example, the case studies reported that the IMF paid insufficient attention to shared and non-core areas in Ghana (SOE reform), Latvia (SOE reforms), Malawi (governance and business climate), Pakistan (power sector reform), Jordan (business climate) and Ukraine (governance and business climate). In the case of Romania, while the importance of reforms in non-core areas was discussed in program documents, they were not included as SCs. The studies for Grenada and Jamaica also highlighted the importance of intervention by other IDIs in non-core macro-critical issues such as crime, labor market distortions and energy sector weaknesses where the problems were deep rooted but the Fund was not equipped to address them.

39. Collaboration with IDIs is important, but ultimately the setting, monitoring and follow-up of SCs remain the full responsibility of the IMF. Greater focus on growth enhancing SCs may require the IMF to be more proactively involved in critical areas outside of the Fund's core expertise while developing in-house expertise and/or consolidating effective and structured

<sup>19</sup> Documents include in the Memorandum of Economic and Financial Policies (MEFP) and Letter of Intent (LOI).

cooperation with other IDIs. In the case study for Pakistan (2008 SBA), the authorities criticized the IMF's "hands-off approach" as regards the power sector reforms and the overreliance upon others (e.g., the World Bank and Asian Development Bank) in alleviating the supply-side constraints on real growth. They suggested that the Fund should have advised based on best international practices rather than adopting a "hands-off" approach. In Ukraine, the authorities explained that the IMF should have paid more attention to areas typically reserved for the World Bank (e.g., reforms to improve business climate), given the IMF's "greater leverage" stemming from its larger financial support. Relatedly, the 2020 IEO evaluation on *IMF Collaboration with Bank on Macro-Structural Issues* (IEO, 2020), which focused on inter-agency collaboration in surveillance, finds that inter-agency collaboration is effective when there is a well-defined roles and framework in place, such as exist for Debt Sustainability Analysis and in the Financial Sector Assessment Program (FSAP). For SCs in the program context, however, there is no framework that guides the appropriate role of IDIs even in the areas where the IMF does not have the relevant expertise.

## VI. STRUCTURAL CONDITIONS AND STRUCTURAL REFORMS

40. This section focuses on assessing the impact of IMF's SCs on progress with structural reforms, based on regression analysis of the relationship between SC implementation and an index of progress on structural reform. SC implementation is represented by observed SC scores described above, incorporating both quantity (implementation) and quality (depth and growth/efficiency orientation) aspects, that is the SCIDG. Progress on structural reform is measured by the changes in the Structural Reform Index (SRI) compiled and developed by the IMF's Research Department (RES).

41. The RES's SRI is constructed based on detailed information on regulatory stances and reform episodes in both real and financial sectors. It covers assessment of reforms in domestic finance (regulation and supervision); external finance (capital account openness); trade (tariffs); product market (regulation in electricity and telecommunication, two large network industries); labor market (job protection legislation); and composite worldwide governance indicator. Each sector contains multiple sub-indicators. Aggregate reform index of each sector is an average of its sub-indicator scores, reflecting the overall regulatory stance and liberalization status of the sector.<sup>20</sup> The SRI covers a total of 90 countries including emerging market countries and LICs for the period of 1973–2014.

42. Using the SRI dataset, the connection between structural reform and growth is examined in detail in the 2019 October WEO (IMF, 2019a). The WEO study reported that reforms in such areas as governance, domestic and external finance, and product and labor markets can deliver sizable output gains in the medium term. The results suggest that a major and comprehensive reform package may raise output in the average emerging market and developing economy by more than 7 percent over a six-year period. It may double the speed of convergence of the average emerging

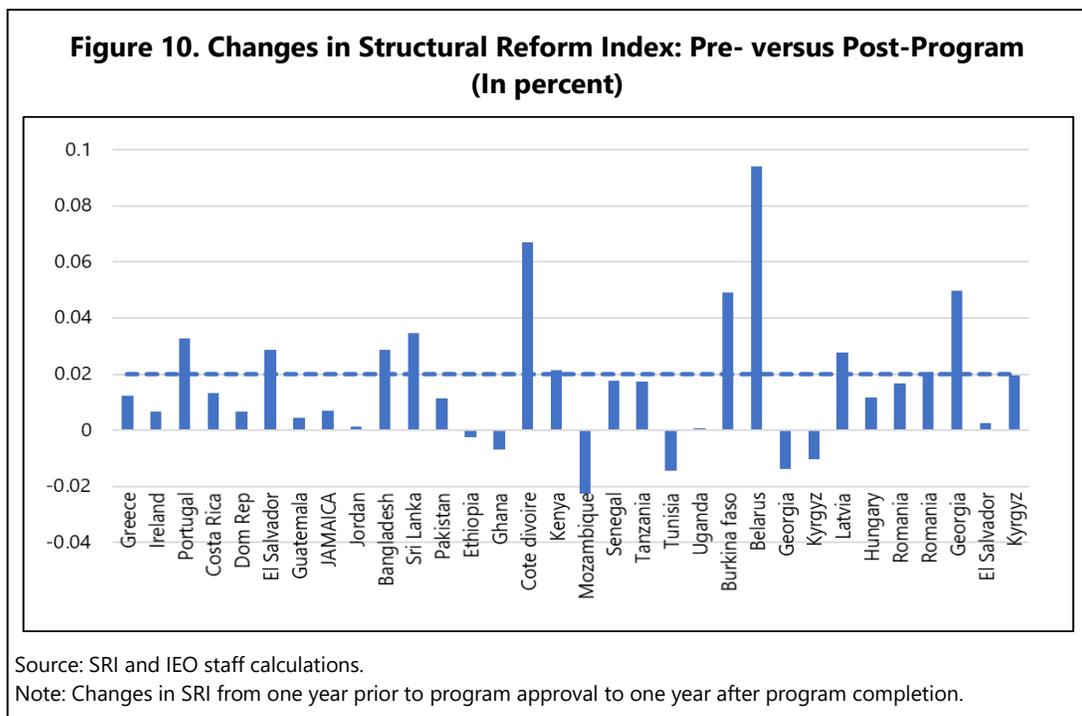
---

<sup>20</sup> See *World Economic Outlook, October 2019*, Chapter 3 "Reigniting Growth in Low-income and Emerging market Economies: What Role can Structural Reforms Play?"

market and developing economy, raising annual GDP growth by about 1 percent for some time. At the same time, reforms may take several years to be completed, and some may entail short-term costs for activity (e.g., policies that ease employee hiring/firing and liberalize domestic finance). Such reform measures are most likely to succeed if they are implemented under favorable economic conditions and early in authorities' electoral mandate. The research also found that reform gains are greater when governance and access to credit are strong, and where labor market informality is initially higher because reforms help reduce it. It highlighted the importance of carefully tailoring the reforms to country circumstances to maximize their benefits.

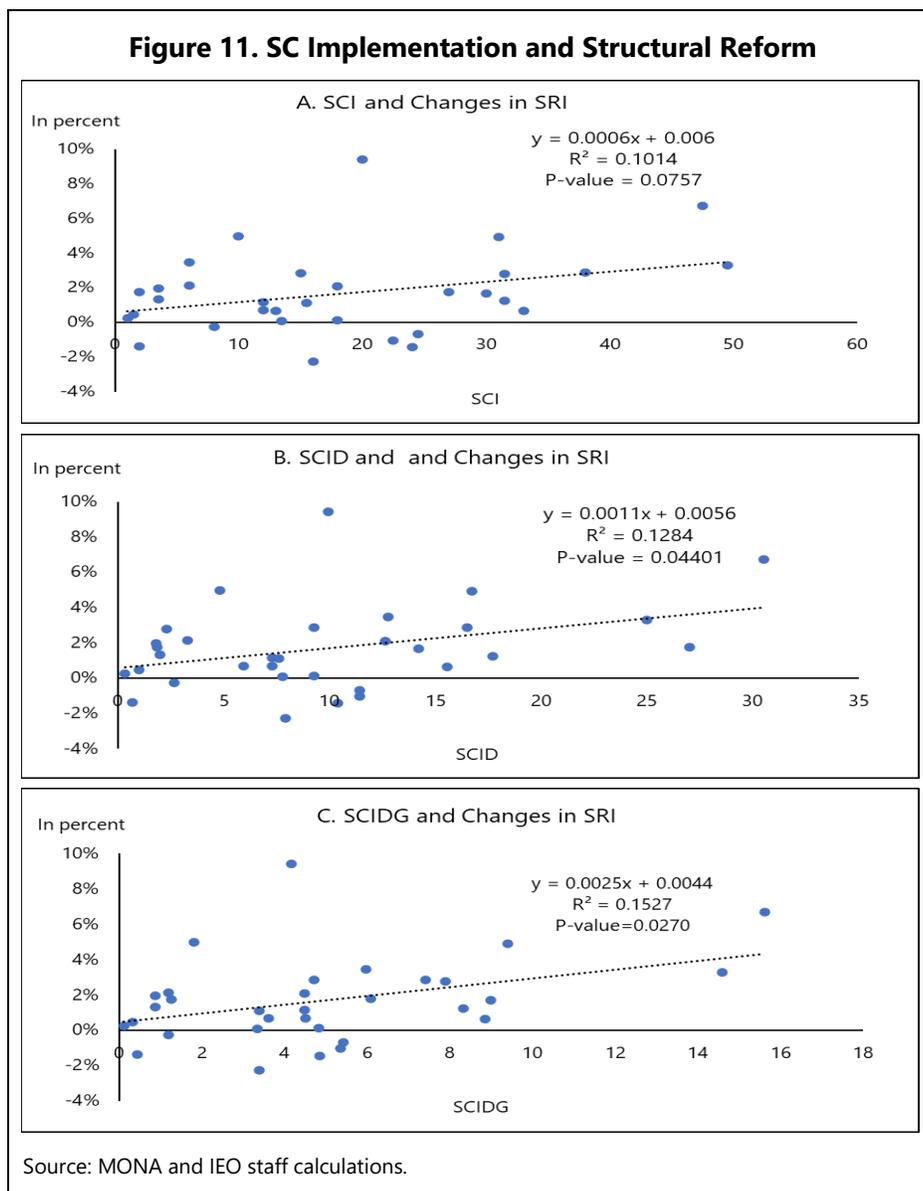
43. For our purpose, the last category of SRI, "composite worldwide governance indicator," is replaced with the World Bank's Government Effectiveness Index (GEI) in order to focus only on the institutional quality of the government sector, i.e., the country authorities who are responsible for implementation of SCs. Thirty-two programs in our dataset overlapped with the sample in this evaluation.<sup>21</sup> To estimate the progress in structural reform for these programs, this paper calculated the cumulative percentage change in SRI between T-1 and T+E+1 where T and T+E refer to the first and last year of the program, respectively.

44. In most cases in our dataset, the SRI improved during the program period. Figure 10 shows that in most cases where SRI is available, the SRI one-year post-program improved compared to SRI one year prior to program approval, with the average increase of about 2 percentage points. This suggests a positive role of IMF programs in advancing the structural reform agenda.



<sup>21</sup> Ukraine was excluded from the analysis due to multiple back-to-back programs most of which went off track.

45. Looking particularly at the link between structural conditions and progress on structural reforms, regression analysis indicates a positive relationship between SC implementation and structural reforms, especially when the depth and growth orientation (or content) of SCs are accounted for (Figure 11). The results of the bivariate regression show that SRI is positively related to SCI and the positive relationship is statistically significant at 8 percent (Panel A). The positive relationship becomes stronger (with steeper slope) and more robust with greater statistical significance when the depth and growth orientation are accounted for (Panels B and C), suggesting that the quality of SCs matters for successful structural reforms. Given the evidence on the growth impact of SRI, these results provide some preliminary and indirect support for growth benefits of SCs implemented in the program context, which is discussed in greater detail in the next section.



## VII. GROWTH BENEFITS OF SCs IN THE MEDIUM RUN

46. This section turns to growth benefits of SCs in the medium run. Discussions on growth benefits draw on the findings from accompanying thematic background paper by Kim and others (2021) which presents an in-depth analysis of medium-run growth benefits of SCs implemented during the program. This paper presents further regression analyses for sensitivity check.

47. Kim and others (2021) find that both stabilizations and reforms implemented during the program have affected post-program potential growth positively and that the quality of structural conditionality (measured by the depth and growth-orientation of SCs) contributed significantly to growth benefits. Specifically, SCI by itself is not found to have statistically significant impact on post-program potential growth (measured relative to a benchmark estimated based on external factors alone). However, SCI is found to affect post-program potential growth positively and significantly if interacted with the average depth and growth orientation scores, underscoring the importance of the quality of SCs in producing growth benefits.

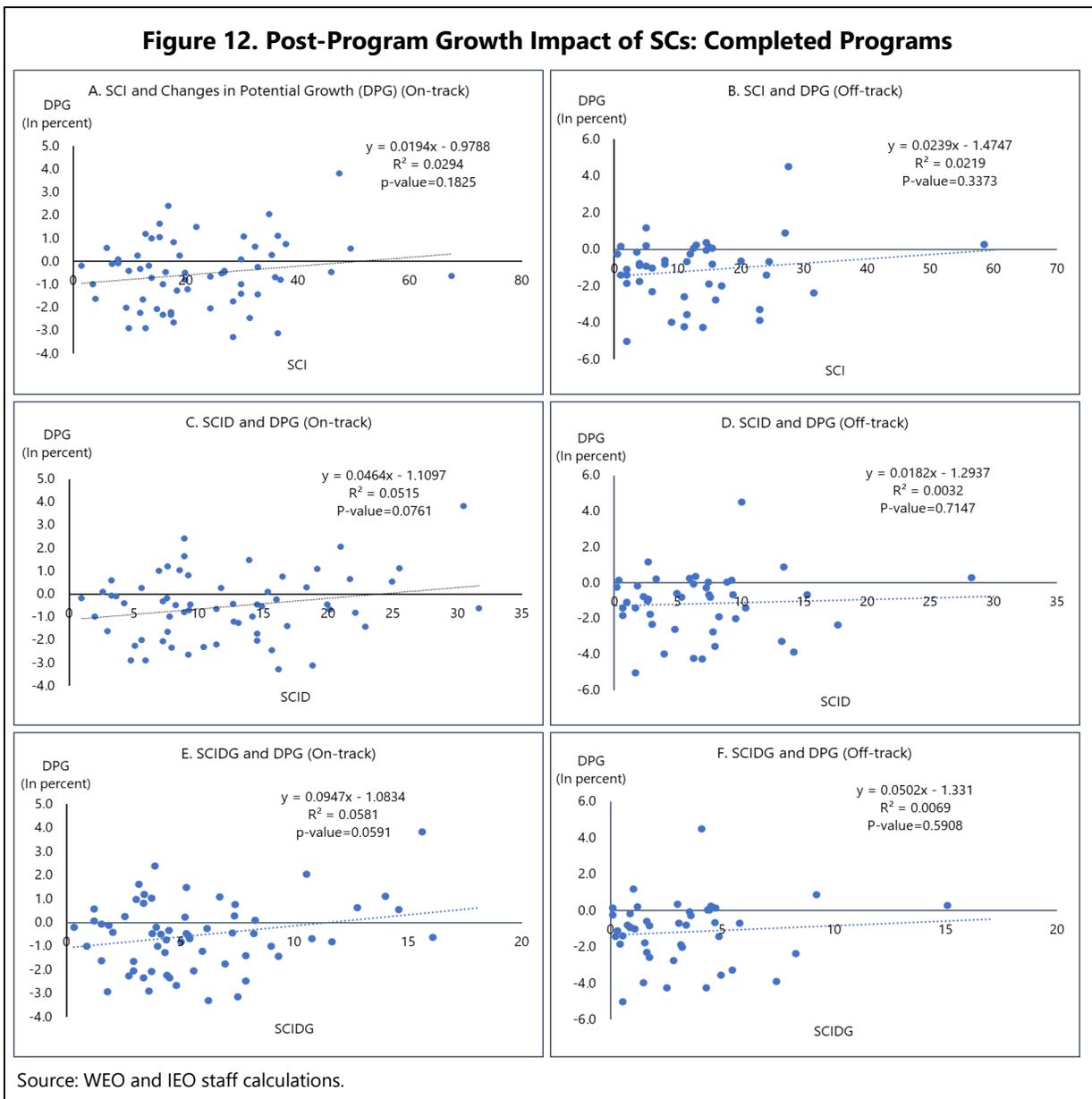
48. Above background paper recognizes that the positive findings on growth benefits of SCs are suggestive rather than conclusive because of the relatively small sample size, which in turn restricts the duration of the post-program period to three years. For this reason, several additional sensitivity checks are performed here based on simple bivariate regressions in which post-program growth performance is measured differently and longer duration is considered for the post-program period by using WEO projections for 2020–25.<sup>22</sup> Specifically, post-program growth performance is measured by the difference in the average potential growth rate (denoted as DPG) between the 5-year post- and pre-program periods using the HP filter. The sample covers 62 completed programs where completed programs are identified using the 2018 ROC definition of program completion.<sup>23</sup> In case of multiple back-to-back programs, they are treated as a (long) single program.

---

<sup>22</sup> In Kim and others (2021), the multivariate regression samples are small because they are based on actual data only (so that even the three-year post-program period is not well defined for many programs completed after 2016) and because of missing data in other control variables included in the regressions. In the simple bivariate regressions discussed here, data are extended to include WEO projections for 2020–25 (taken from the 2020 January WEO), which are not affected by the outbreak of the COVID-19 crisis. This allows recently completed programs to be included in the regression sample. A caveat is that use of WEO growth projections could introduce optimism bias into the regression analysis as found by IMF (2019b) and IEO (2014).

<sup>23</sup> The 2018 ROC defined “completed” programs as those with all reviews completed during the program period, including if they were completed with delays, after rephasing, or during a program extension. “Off-track” programs are those with at least two reviews completed and at least two reviews not completed at the end of the program. “Quickly off-track” programs are those with at most one review completed and at least two reviews not completed at the end of the program. The sample excludes four outliers with  $DPG > +4.0$  or  $DPG < -4.0$ , Central African Republic (2016 ECF), Armenia (2014 EFF), Belarus (2009 SBA), and Angola (2010 SBA).

49. The bivariate regression results provide good support for the medium-run growth benefits of SCs, consistent with the findings from multivariate regressions in Kim and others (2021). The relationships between SCI and DPG is positive but not statistically significant, suggesting that an implementation of SCs by itself does not necessarily deliver growth benefits (Figure 12, Panel A). Once depth and growth-orientation are taken into account (i.e., SCID and SCIDG), however, the relationship between SC implementation and DPG is positive and statistically significant, underscoring the importance of the quality dimension of SCs in producing growth benefits (Figure 12, Panels C and E). These results are in sharp contrast to the results for 44 off-track programs where no significant relationships are detected between DPG and any of the three score indices suggesting that the program completion status also matters for growth benefits of SCs (Figure 12; Panels B, D, and F).



## VIII. CONCLUSIONS AND LESSONS

50. This thematic background paper assesses the composition and quality of structural conditions in IMF-supported programs over the period of September 2008–March 2020 and their role in promoting structural reforms as well as growth and fiscal stability in the medium run. To this end and for consistent cross-country comparisons, it develops structural conditions score indices for each program included in the evaluation sample and uses them to assess quantitatively the quality and the medium-term growth and fiscal impacts of SCs. The analysis distinguishes between SCs in program design and SCs in program implementation where appropriate.

51. Key findings are broadly supportive of the role played by SCs in promoting growth and sustaining fiscal strength after the program, but also suggest some areas for attention in the design and implementation of SCs. Six findings are notable in this regard.

- First, SC implementation is strong but the average quality of SCs is relatively low in terms of depth and growth orientation. SC implementation is stronger in GRA programs than in PRGT programs while the quality (in terms of depth and growth orientation) of SCs is broadly similar between the two. PRGT programs have higher share of fiscal SCs and lower share of growth oriented SCs than GRA programs.
- Second, the parsimony and macro-criticality principle seems to have worked relatively well to induce streamlining of SCs but less so for providing incentives to increase the quality of SCs.
- Third, the IMF's delivery of technical assistance (TA) does not seem to have been targeted at countries with the lowest capacity and does not appear to have been effective in supporting SC implementation.
- Fourth, only a small fraction of SCs have explicitly recognized the need for collaboration with other partner institutions, and the implementation of such SCs has been weaker than that of other SCs.
- Fifth, a positive relationship was observed between SC implementation and progress on structural reforms, especially when the depth and growth orientation (or content) of SCs are accounted for. For the growth in the medium run, a mere implementation of SCs by itself is not found to have delivered significant growth: greater depth and growth orientation of SCs have mattered significantly for such benefits.

52. Based on these findings, following broad lessons are drawn for the evaluation:

- First, increasing the depth and growth orientation of SCs while streamlining the number of SCs overall would help to promote medium-run growth while respecting the parsimony principle.

- Second, more attention is warranted as to the provision of TA and collaboration with partner institutions in the program context. The result that Fund TA has not significantly contributed to improved implementation of structural conditions suggest the need for further consideration of how best to support countries' capacity to implement reforms.
- Third, particular attention needs to be given to effective and structured collaboration with partner institutions given the large share of SCs in shared/non-core areas of Fund expertise and particularly weak implementation in these areas. In this regard, a new framework with clearly defined responsibilities would help to strengthen collaboration with IDIs to foster stronger progress with growth-related reforms.
- Fourth, efforts should continue to improve the quality and user-friendliness of the MONA database. Records of the content and implementation status of SCs need to be expanded and updated regularly and more frequently.

## REFERENCE

- Andone, I., and Beatrice Scheubel, 2019, "Once Bitten: New Evidence on the Link between IMF Conditionality and IMF Stigma," European Central Bank Working Paper No. 2262, April.
- Bird Graham, and Dane Rowlands, 2017, "The Effect of IMF Programs on Economic Growth in Low Income Countries: An Empirical Analysis," *Journal of Development Studies*, 53:12, 2179-2196.
- Buira, Ariel, 2003, "An Analysis of IMF Conditionality," G24 Discussion Paper No. 22, August.
- Clements, Benedict, Sanjeev Gupta, and Masahiro Nozaki, 2013, "What Happens to Social Spending in IMF-Supported Programmes?" *Applied Economics*, Vol. 45 (28), pp. 4022–33.
- Furceri, Davide, João Tovar Jalles, and Prakash Loungani, 2015, "Fiscal Consolidation and Inequality in Advanced Economies: How Robust Is the Link?" *Inequality and Fiscal Policy*, edited by B. Clements and others (Washington: International Monetary Fund).
- Gupta, Sanjeev, 2021, "Fiscal Adjustment and Growth in IMF-Supported Programs," IEO Background Paper No. BP/21-01/03 (Washington: International Monetary Fund).
- Independent Evaluation Office of the International Monetary Fund (IEO), 2008, *Structural Conditionality in IMF-Supported Programs* (Washington: International Monetary Fund).
- \_\_\_\_\_, 2014, *IMF Forecasts: Process, Quality, and Country Perspectives* (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).
- \_\_\_\_\_, 2020, *IMF Collaboration with the World Bank on Macro-Structural Issues* (Washington: International Monetary Fund).
- International Monetary Fund (IMF), 2001, "Structural Conditionalities in Fund-supported Programs," (Washington).
- \_\_\_\_\_, 2002a, "Guidelines on Conditionality" (Washington).
- \_\_\_\_\_, 2002b, "Revised Staff Statement: Principles Underlying the Guidelines on Conditionality," Published as an attachment to the Guidelines on Conditionality (Washington).
- \_\_\_\_\_, 2002c, "Review of the Key Features of the Poverty Reduction and Growth Facility—Staff Analyses" (Washington).

- \_\_\_\_\_, 2005, "The Design of IMF-supported Programs," Occasional Paper No. 241 (Washington).
- \_\_\_\_\_, 2009, "Conditionality in Fund-Supported Programs – Purposes, Modalities, and Options for Reform" (Washington).
- \_\_\_\_\_, 2012a, "2011 Review of Conditionality—Overview Paper" (Washington).
- \_\_\_\_\_, 2012b, "Staff Guidance Note on the Fund's Engagement with Countries in Fragile Situations" (Washington).
- \_\_\_\_\_, 2012c, "2011 Review of Conditionality—Background Paper 3: Outcomes of Fund-Supported Programs" (Washington).
- \_\_\_\_\_, 2013, "Jobs and growth: Analytical and operational considerations for the Fund", (Washington).
- \_\_\_\_\_, 2014, "Revised Operational Guidance to IMF Staff on the 2002 Conditionality Guidelines" (Washington).
- \_\_\_\_\_, 2015a, "Crisis Program Review" (Washington).
- \_\_\_\_\_, 2015b, "Public Debt Vulnerabilities in Low-income Countries: The Evolving Landscape" (Washington).
- \_\_\_\_\_, 2016, "Staff Note for the G20—A Guiding Framework for Structural Reforms" (Washington).
- \_\_\_\_\_, 2017, "Social Safeguards and Program Design in PRGT and PSI-Supported Programs" (Washington).
- \_\_\_\_\_, 2018a, "Macroeconomic Developments and Prospects in Low-Income Developing Countries" (Washington).
- \_\_\_\_\_, 2018b, "IMF Fiscal Monitor: Capitalizing on Good Times," *World Economic and Financial Surveys* (Washington).
- \_\_\_\_\_, 2018c, "Capacity Development (CD) Strategy Review—Overview Paper" (Washington).
- \_\_\_\_\_, 2018d, "Capacity Development (CD) Strategy Review—Staff Background Studies and Short Notes" (Washington).
- \_\_\_\_\_, 2019a, "Reigniting Growth in Emerging Market and Low-Income Economies: What Role for Structural Reforms?", *World Economic Outlook* October 2019, Chapter 3 (Washington).

- \_\_\_\_\_, 2019b, "2018 Review of Program Design and Conditionality," IMF Policy Paper No. 19/012 (Washington).
- \_\_\_\_\_, 2019c, "IMF Policies and Practices on Capacity Development" (Washington).
- Kentikelenis, Alexander E., Thomas H. Stubbs, and Lawrence P. King, 2016, "IMF Conditionality and Development Policy Space, 1985–2014," *Review of International Political Economy* 23 (4): 543–82.
- Killick, Tony, 2002, "The Streamlining of IMF Conditionality: Aspirations, Reality and Repercussions," Overseas Development Institute (London).
- 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).
- Mercer-Blackman, Valerie, and Anna Unigovskaya, 2000, "Compliance with IMF Program Indicators and Growth in Transition Economies," IMF Working Paper 00/47 (Washington: International Monetary Fund).
- Przeworski, Adam, and James R. Vreeland, 2000, "The Effects of IMF Programs on Economic Growth," *Journal of Development Economics*, Vol. 62, 2000. 385-421.