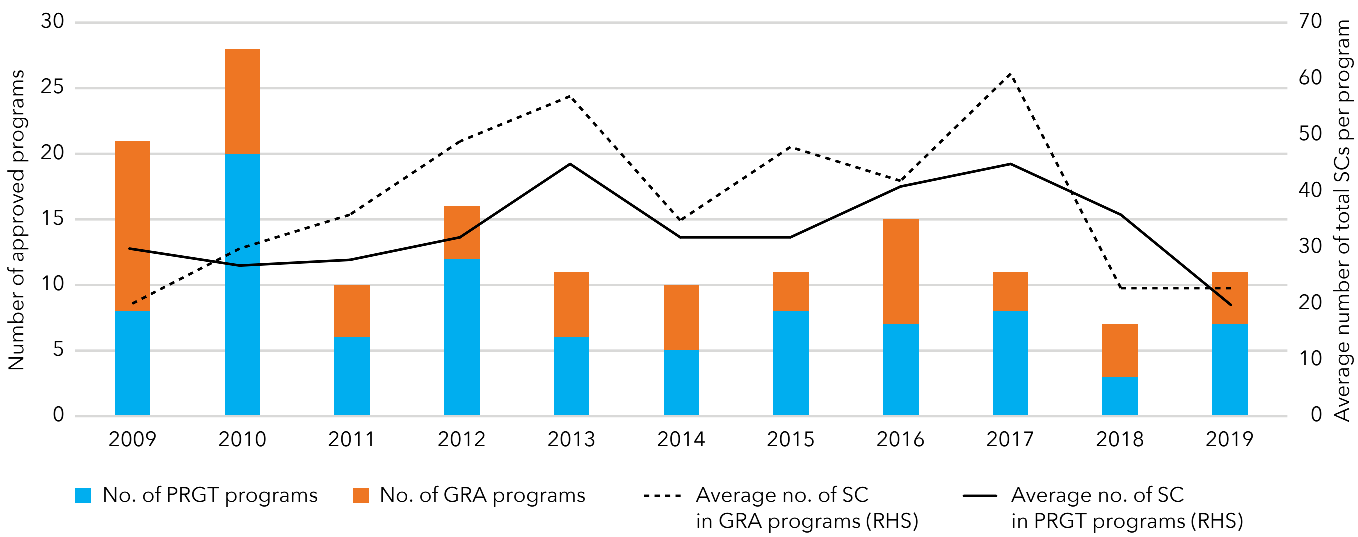
STRUCTURAL CONDITIONS, STRUCTURAL REFORMS, AND GROWTH

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.

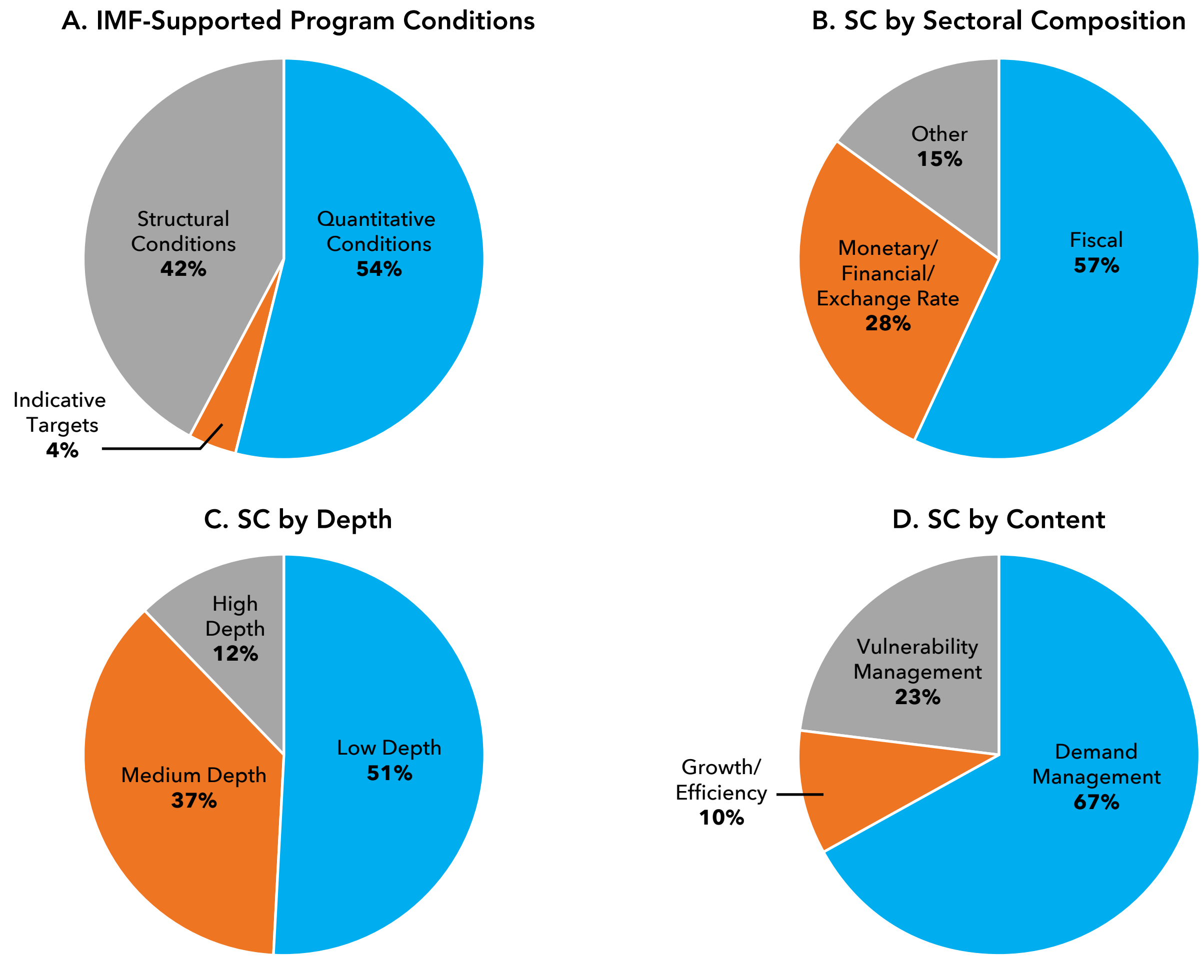
[42] The spike in 2017 in the average number of SCs in GRA programs is explained mainly by the fact that one of the three GRA programs approved in 2017 had an exceptionally large number of SCs (80 in total).

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



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.

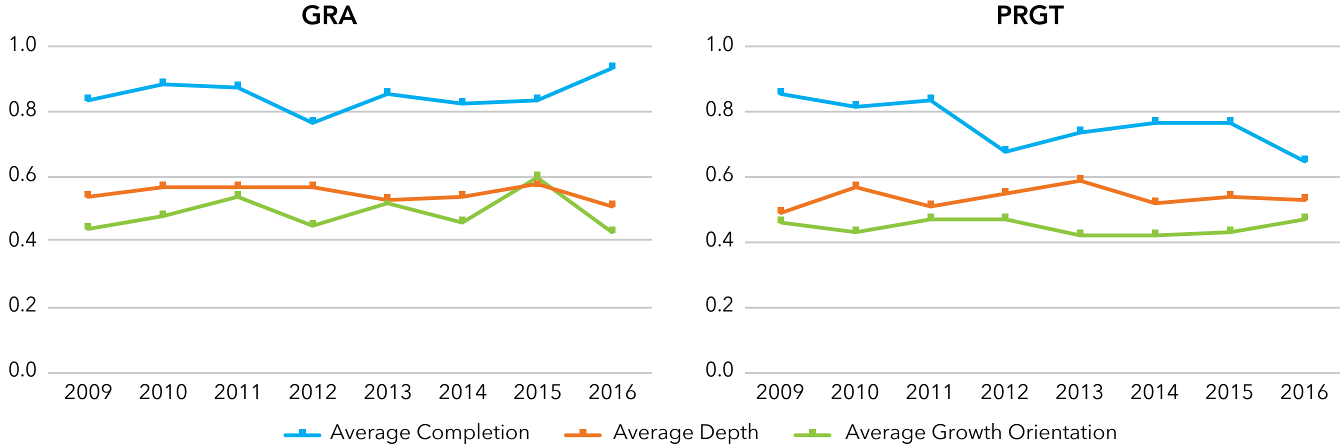
SC implementation was on average somewhat weaker for programs with a higher volume of SCs (Figure 28). SCs with higher depth would normally be considered more challenging to implement than lower depth SCs because the former require more technical input and stronger political commitment. However, no statistically 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

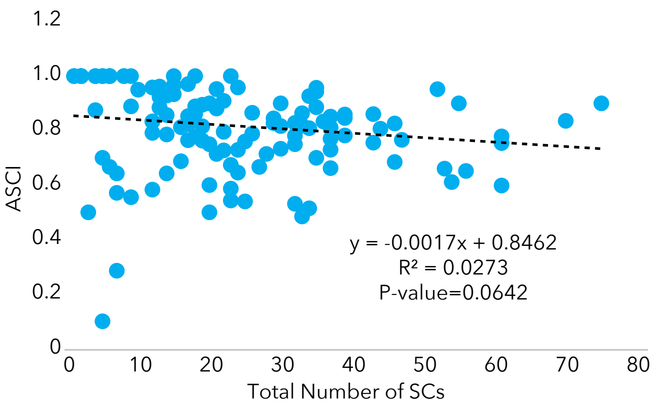
[44] Implementation status is not provided in the MONA database for SCs in program reviews that were never completed.

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



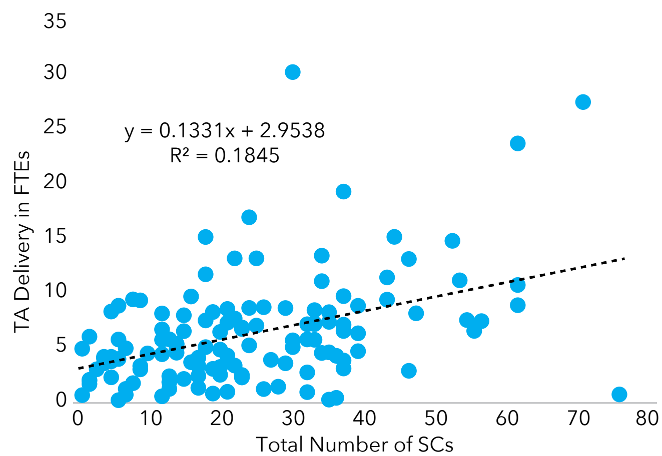
Sources: MONA database; IEO staff calculations.
Note: The year on the horizontal axis represents the year of program approval.

FIGURE 28. VOLUME OF SCs AND SC IMPLEMENTATION



Source: Kim and Lee (2021).

FIGURE 29. VOLUME OF SCs AND TA DELIVERY



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.

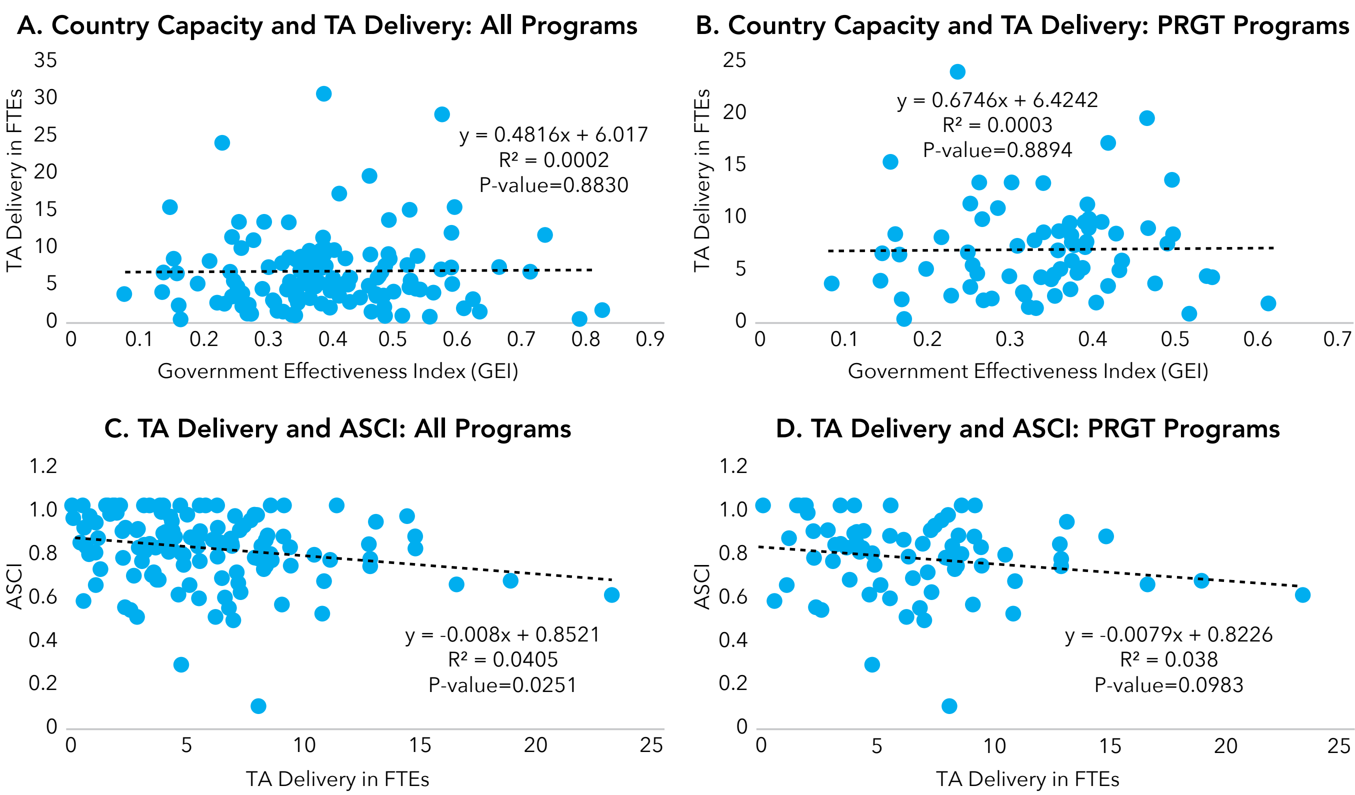
Two related questions are how well CD provision has been aligned with country need or capacity and how CD has affected the implementation of SCs. Cross-country evidence is less encouraging on these questions. Data on programs during the evaluation period suggest that more TA resources seem to have been 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

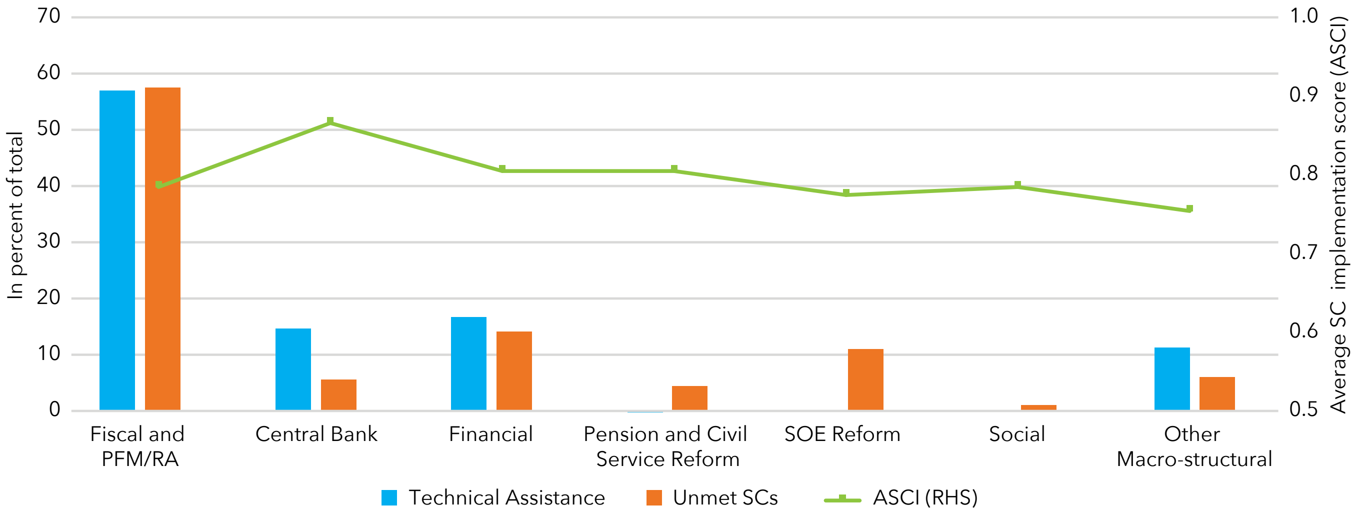
FIGURE 30. COUNTRY CAPACITY, IMF TA, AND SC IMPLEMENTATION



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.