

Self-Evaluation of IMF-Supported Programs

21. This section examines the self-evaluation activities and products in place at the IMF over the past decade to assess the successes and shortcomings of lending arrangements, namely ex post assessments (EPAs) conducted at the conclusion of a program in countries with prolonged use of IMF resources, and ex post evaluations (EPEs), conducted following programs supported by exceptionally large access to IMF resources.⁹

22. EPEs were introduced in 2002 to assess whether exceptional access programs were consistent with IMF policy and to review their performance.¹⁰ EPAs were initiated in 2003 to provide an opportunity for the IMF to step back from longer-term program engagement (LTPE) to analyze the economic problems facing a country, to provide a critical and frank review of progress during the period of IMF-supported programs, to conduct a forward-looking assessment that took into account lessons learned and presented a strategy for future IMF engagement, and where appropriate to present an explicit “exit strategy.”¹¹ EPEs and EPAs were undertaken by interdepartmental teams led by someone from outside the area department responsible for the program.

23. Self-evaluation of programs is important to learning and accountability in the IMF. By drawing lessons from past programs, such evaluations can help improve subsequent operations in the respective member

country or other member countries facing similar challenges. Additionally, evaluating individual programs is a key mechanism for the IMF to discharge its fiduciary responsibility to account to members for the use of their resources, by examining program design and assessing program success and IMF performance.¹² EPAs and EPEs typically included a statement from the authorities describing their views on the program and in some cases on IMF performance.

24. This evaluation finds that EPAs and EPEs have mostly fulfilled their learning and accountability roles. These assessments often distilled relevant lessons for IMF engagement in the countries concerned, and many of these lessons were taken into account in the design of follow-up operations. Assessments also offered a vehicle for authorities to share their own views on the program and the IMF’s work. However, the quality of lessons for improving future operations varied widely, and the assessments were less successful at identifying lessons with relevance across countries and for the institution as a whole.

25. In April 2015, as noted above, the IMF decided to discontinue EPAs and to shift the discussion of relevant LTPE lessons to the Staff report that would be prepared in the event of a subsequent program request. This decision was taken as part of a budget streamlining exercise, and not as part of a strategic review of the evaluation function or an in-depth assessment of EPAs. The implications of this decision are discussed at the end of this chapter and in [Chapter 7](#). The following discussion examines the experience with EPAs and EPEs in order to provide useful input for designing mechanisms for the self-evaluation of programs going

⁹This section draws on Goldsbrough (2015), which assesses the coverage, quality, and lessons in EPAs, EPEs, and Joint EPAs/EPEs completed between August 2005 and December 2013.

¹⁰IMF (2003a). In 2002, the Board established EPEs for exceptional access programs in the context of capital account crises; this requirement was expanded in 2003 to cover all exceptional access programs, including high-access precautionary programs (excluding the Flexible Credit Line). For operational guidance, see IMF (2005) and IMF (2010b).

¹¹IMF (2003b). According to the definition agreed in 2006, a country is determined to have LTPE if it has spent at least seven of the last ten years under IMF-supported programs (concessional or nonconcessional) excluding time spent under undrawn precautionary arrangements. For operational guidance see IMF (2003c) and IMF (2010a).

¹²To discharge its fiduciary responsibilities, the IMF also undertakes safeguards assessments of central bank governance and control frameworks for countries engaging in borrowing arrangements with the IMF and post-program monitoring of programs with large resources outstanding to provide early warning of policies that could jeopardize the country’s external viability and its capacity to repay the IMF.

Table 1. Self-Evaluation of IMF-Supported Programs¹*(In percent)*

Type of Assessment	Share of Program Arrangements Covered	Share of Resource Commitments Covered
EPA	39	3
EPE	10	53
Joint EPA/EPE	3	19
Total	52	75

¹For arrangements approved in 1999–2012 and expired by end-2013 for which an assessment was completed by end-2014. For EPAs, from 2006 onward, figures exclude precautionary programs unless drawn. Approved arrangements exclude the Flexible Credit Line.

forward, including informing the design of the instrument that will replace EPAs.

EPA/EPE Coverage

26. As of end-2014, the IMF had undertaken self-evaluation assessments for about half of all programs approved between 1999 and 2012 (see [Table 1](#)).¹³ Seventy-five percent of the volume of IMF resources committed over this period was covered by an assessment. With the drop in the number of member countries with LTPE over the period, EPAs played a decreasing role in the share of programs and commitments assessed. On the other hand, the increase in exceptional access programs during the financial and economic crisis led to a larger share of IMF commitments being covered by an EPE assessment.

27. This coverage seems appropriate from a systemic risk perspective, since scrutiny was focused on larger programs and on programs in member countries with LTPE, where the IMF was likely to have the greatest impact due to its continued presence. But there was no self-evaluation of programs within normal lending limits in support of countries that were not engaged with the IMF on a long-term basis.¹⁴ This gap in coverage had implications for transparency and accountability

¹³This includes assessments completed by end-2014 for programs approved during 1999–2012 and expired by end-2013. The review period was set beginning in 1999 in order to capture programs for which EPAs were undertaken once the EPA requirement was initiated in 2003. A list of countries for which assessments have been carried out can be found in Goldsbrough (2015), [Annex 2](#).

¹⁴Civil society organizations and other stakeholders interviewed for this evaluation indicated that the IMF has a fiduciary duty to examine and report on the results, quality of design, and implementation of all the programs it finances, as is the case with most other IFIs. The concern about partial coverage of programs is exacerbated by the fact that at the IMF there is no other mechanism for assessing completed programs, whereas at other IFIs at least a sample of lending operations undergoes independent evaluation.

and for the IMF's ability to draw lessons from experience with certain types of programs, some of which could be seen as having systemic lessons.¹⁵

Quality of EPAs and EPEs

28. The evaluation found that these assessments were generally effective tools for reflecting on experience, although there was room for improvement in their candor, in their focus on IMF and Staff performance, and in drawing lessons that could be applied across the institution. Goldsbrough (2015) found that the overall quality of a large majority of EPAs and EPEs was adequate or very good (78 percent of EPAs and 87 percent of EPEs), but that a sizable minority of EPAs, as well as most EPA Updates, were weak. Overall these reports were strongest with respect to assessing program involvement, program design, and what worked and what did not work in programs (see [Figure 1](#)). They did less well in probing why a program succeeded or failed and the broader rationale for IMF engagement. Many of them presented evidence and set the stage for pointing to and analyzing weaknesses in the program, but stopped short of drawing conclusions or identifying useful lessons for improving IMF operations more broadly.

29. The quality of lessons offered for improving IMF operations varied widely; nearly half of EPAs and one-third of EPEs failed to identify clear lessons. Only 20 percent of the assessments identified explicit forward-looking lessons for IMF operations in general. The weaker reports identified no lessons—or lessons that were too general to offer much help for future operational engagement with the program country or

¹⁵Examples of programs that were not evaluated included the 2010 Iraq Stand-By Arrangement, the programs between 2003 and 2012 in the Dominican Republic, and the 2002 and 2009 programs in the Democratic Republic of the Congo.

Figure I. Overall Quality of Self-Evaluation Reports
(In percent of type of report, excludes EPA Updates)



Source: IEO analysis.

a roadmap for an exit strategy. The Kenya 2008 EPA was found by the review to be an example of a high-quality assessment (see [Box 2](#)), and a few of the more recent EPEs (e.g., on Greece, Ireland, and Ukraine) presented particularly candid assessments.

30. A key factor determining the quality of assessments and lessons seems to have been the experience of

the evaluation team leader.¹⁶ Goldsbrough (2015) found that assessments prepared by senior Staff (B-level)

¹⁶Goldsbrough (2015) examined other potential influences on the quality of assessments and lessons: size of country involved, type of economy, geographic region, type of IMF lending facility involved, whether a follow-up program was likely. None seemed to be a major determining factor.

Box 2. Kenya Ex Post Assessment, 2008: An Example of Good Practice

The EPA report completed for Kenya in August 2008 provides a good example of self-evaluation of lending. It laid out a frank and critical analysis of the IMF's engagement through three programs (1993, 1997, and 2003) and provided clear lessons for future engagement. The report discussed the rationale for IMF involvement, the pressures leading to longer-term program engagement, and the strengths and weaknesses of program design. Among the issues it addressed were the impact of IMF engagement on domestic institutions, the effectiveness of IMF technical assistance, and the appropriateness of conditionality. While noting the centrality of governance issues, the report faulted the IMF for imposing conditions that were not macro-critical (e.g., the prosecution of former high-level officials) and for shifting its focus too late to institutional strengthening.

The report tackled the question of why programs had gone off-track despite the fact that macroeconomic policy design and its implementation were broadly appropriate. While acknowledging the difficulties of engagement in

Kenya during this period, the report identified a number of ways in which the IMF could have done better, and clear lessons for future engagement in Kenya, including:

- the importance of focusing conditionality on macro-critical issues;
- the need to consider an instrument with shorter duration and lower access to IMF resources in cases in which there are serious governance concerns and an unsound track record; and
- the need for enhanced transparency and a communications strategy when programs languish off-track for long periods.

Finally, the report integrated into its analysis the views of country authorities and other stakeholders, including the authorities' perspective that Kenya had been singled out and treated in a less than evenhanded manner. These views fed into the lessons that the assessment distilled for the IMF.

team leaders were better in overall quality and in the clarity of lessons identified than reports prepared by more junior team leaders. This underscores that experience is critical when trying to determine what could have worked better and identify alternative policies and strategies. In interviews, Staff also pointed out that senior Staff could act more independently and be more candid in assessments.¹⁷

31. Most of the assessment reports consigned the country authority's statement on the program to an annex, rather than taking it into account in the analysis of what had transpired during the program. Only a few reports discussed criticisms raised by external stakeholders.

32. The assessment reports were reviewed by SPR and the respective area department. This review and the internal clearance process limit the room of the EPA/EPE team to provide candid and independent assessments of programs. However, this evaluation acknowledges that it would be difficult to design an alternative "more independent" review process. To address this concern, other IFIs have a system of independent

validation of assessments, usually conducted by their independent evaluation office, designed to strengthen the independence and candor of these reports.

Utilization of Lessons from EPAs and EPEs

33. Ultimately, the main value of EPAs and EPEs lies in their contribution to improving future programs. Beyond their quality, discussed above, this depends on their dissemination and the extent to which any lessons they offer were learned by Staff and taken into account in the design and implementation of future programs. The evaluation found that country-specific lessons from EPAs and EPEs were utilized in the design of successor programs and that Executive Directors relied on these assessments when considering new programs.¹⁸ Indeed, Directors often referred to EPA and EPE findings and lessons in Board discussions of successor programs.¹⁹

¹⁷Nevertheless, leadership of these assessments was often assigned to more junior Staff. It seems that the selection of the project leader was often guided by the goal of allowing junior Staff to gain the experience needed to be considered for promotion. Some of these team leaders reported that they were concerned by the prospect of having to criticize a mission chief who could impact promotion decisions.

¹⁸This conclusion was based on a review of a sample of 15 successor programs for countries for which EPAs and EPEs had identified clear and actionable lessons.

¹⁹The IEO reviewed minutes of Board discussions on EPAs and EPEs, and on the request for successor programs for the respective countries. It found that one or more Directors made substantive references to the EPA or EPE lessons in discussing successor program

On the other hand, the initial Board discussions of these assessments themselves were brief and their Summings Up were usually bland.

34. Dissemination consisted almost exclusively of issuing the reports to the Board and disclosing most of them to the public.²⁰ Few efforts were made to call attention to the results or to disseminate lessons beyond the teams working on the respective country.²¹ The evaluation found no mechanism that highlighted or disseminated lessons that might be widely applicable. A large majority of Staff reported having read at least one EPA or EPE, and most found them useful, as did the vast majority of Staff who had participated in these exercises.²² But in interviews, most Staff indicated that they had only read assessments on countries they were working on or one of a handful of high-profile assessments, for example the EPE for the 2010 Greece Stand-By Arrangement. It would seem, therefore, that the learning benefits from these assessments flowed mostly to teams working on the country whose earlier programs were assessed.

35. While the 2011 Conditionality Review included an analysis of lessons for conditionality across programs, this evaluation found no attempts to distill lessons from

requests in more than half of the cases reviewed. At the same time, using the metric developed in Goldsbrough (2015), it found that more than half of the initial Board discussions of these assessments were weak, and only a quarter were very good. Indeed, in interviews Directors reported relying heavily on EPA and EPE lessons when considering new programs, but explained that discussions of the assessments were often sidetracked by the need to focus on forward-looking issues.

²⁰A few were not published: Argentina (Joint EPA/EPE, 2006); Brazil (EPE, 2006); Macedonia (EPE, 2014); Pakistan (EPE, 2012); and Turkey (Joint EPA/EPE, 2008).

²¹Operational guidance for Staff on conditionality mentions the potential to reflect on experience at the end of programs but does not require self-evaluation. The guidance states that “while the last program review may not always be the most opportune time to assess the outcome of a Fund-supported program, Staff could use such reviews for stocktaking and, where possible, provide a preliminary assessment of stated program goals and their achievement and of the experience with program implementation.”

²²Three-quarters of the respondents who had participated in EPAs/EPEs agreed or strongly agreed that the exercise was worthwhile, highlighting access to an insider perspective and information and the “unguarded opinions of colleagues.” On the other hand, one-fifth did not find the exercise worthwhile. Many questioned the incentives facing teams who prepared the assessments and the potential for impact.

clusters of EPAs and EPEs on a regular basis. This constituted a missed opportunity to promote learning across country teams and for the institution more broadly. In interviews, several Executive Directors and Staff indicated that they would have appreciated such cross-country lessons.

Recent Developments

36. In April 2015, EPAs were discontinued as part of cost-cutting efforts. Instead, the Board decided that for “. . . countries meeting the LTPE definition, country teams—most knowledgeable about the country situation—would conduct a post program assessment at the beginning of the successor program negotiation or during the Article IV consultation mission, whichever is earlier. The Staff report of a new program request will contain a succinct, peer-reviewed assessment of the previous program for Board consideration.” These assessments will examine the mode of IMF engagement, compliance with conditionality, and whether the ultimate objectives of the program were achieved. Staff estimated that this new modality would yield savings of \$0.2–0.3 million in 2016, due to efficiencies in travel costs and synergies with current surveillance and/or program work (IMF, 2015).

37. A key distinction of the new approach is that the assessment of past programs will be conducted by the country team, rather than by an inter-departmental team. Country teams are likely to be more knowledgeable of country circumstances and can readily integrate lessons into future programs and surveillance work. However, this approach also carries important risks. First, it may lead to less forthright and candid assessments since country teams may feel even more constrained in their independence by career concerns. Second, this new approach does not provide a vehicle for member country authorities to express their views on the program or on the IMF or Staff’s performance. Finally, this new approach will eliminate the opportunity for the Board to reflect on lessons from past programs in time to provide guidance for new requests.