THE FOUR CRITERIA

In general, the EACs played a disciplining role by compelling staff and other stakeholders to deliberate carefully on each of the related requirements. Evidence on country cases over time—buttressed by the view of different stakeholders including Executive Directors, staff, and management—suggests the EAP constituted a helpful guardrail that contributed to systematic thinking, triggered discussions on the key program issues, and guided judgment. However, problems have emerged with respect to the EAC's formulation, clarity, and implementation that relate to the design of the criteria.[40] These concerns cover both the overall framework that the EACs provide, and the specific EACs themselves.

THE EAC FRAMEWORK

The EACs do not provide substantively stronger safeguards relative to NA programs. In principle, the EACs should provide higher substantive standards in order to be fully consistent with the spirit of the conditionality guidelines, which imply a need for stronger safeguards when access is higher. As noted above, however, in the EAP the substantive difference between the EACs and the requirements for NA programs is limited to debt sustainability. While the text, design, and interpretation of the other criteria are different from NA requirements, the differences do not represent a higher standard as such.[41] The criteria also do not seem significantly to affect program design. Most of the higher "evidentiary standard" in the EAP comes, therefore, from the higher level of scrutiny of the criteria and the procedures for reviewing their soundness.

The role of judgment in assessing the criteria generates questions. Each of the criteria involves a measure of staff judgment in assessing whether they are met, which is necessary for providing the Fund with appropriate flexibility to take into account country circumstances. The use of judgment by the Fund is not unique to the EAP; it applies across many Fund policies. However, a concern raised during the EAP evaluation was the extent to which the scope for judgment has diminished the EAP's effectiveness and left open the possibility that the Fund's assessments of the criteria may be unduly influenced at times by the Fund's strategic considerations—including the difficulty of denying financial assistance to a member in need, staff incentives, or external pressures and reduce the transparency of decisions.

The widespread perception of biased assessments in some EA cases is concerning. Outside the Fund, there is a strong perception of political pressures in some high-profile cases affecting the assessment of EACs. Internally, this perception is shared by many and the analysis for this evaluation confirms that pressures on staff and management, exerted directly or indirectly, were strong in high-stakes cases. The majority view among staff is

[40] Bal Gündüz (2024) and Erce (2024) provide key evidence for the discussion of the EACs.

[41] The EAC1 requirement of higher BOP needs than NA limits provide is a definitional statement. Reasonable prospects of program success (EAC4) and regaining market access (EAC3) are not explicit requirements for NA programs according to the conditionality guidelines, but it is difficult to see how a program could be approved without a reasonable prospect of success.

that the EACs have not sufficed to shield the Fund from the pressure in favor of lending when the fulfillment of the criteria is questionable and, therefore, the effectiveness of the framework hinges on staff and management's determination to apply it rigorously. These perceptions affect the credibility and reputation of the Fund, which is seen as being more flexible in some cases depending on the pressure exerted.

Notwithstanding pressures on the Fund, the evaluation did not find direct evidence of reverse engineering of EAC assessments. The evidence suggests that when staff and management presented programs to the Board for approval, they judged the EACs to be met, thought such programs had a reasonable chance of success, and presented the risks involved clearly and explicitly. In several cases, however, such as Argentina (2018)[42] and Greece (2010), there were strong dissenting views among staff about both the substantive compliance with the EAP and the way the decision-making processes were handled. Eventually, those views were resolved by management in exercise of their prerogatives.

In difficult cases where there may have been doubts about whether all of the EACs were met but the Fund had compelling broader reasons to proceed with an EA program, an EC clause might have been useful to consider. These cases were rare, typically involved concerns about the impact of a lack of Fund support on a swathe of the membership, or even on the Fund's standing in the international monetary system, and they involved the provision of assurances from sources unforeseen in the EAP (such as backstops to assure capacity to repay, or strong financing assurances from sections of the membership). Some interviewees felt that if program design was sound and basic principles of the Fund's lending safeguards were met (such as debt sustainability and capacity to repay), then an EC clause could have provided a useful "escape valve" in such circumstances by being more transparent and not stretching the EAP like the current approach may have done. Use of an EC clause would have also allowed the Fund to deal better with situations where IMF repayments were the main source of the BOP need, and to more clearly assign ownership of decisions to support members that may involve higher residual risks.

Alternative views were based on concerns about the implications of an EC clause for evenhandedness and safeguards. Concerns included the risk that problems associated with previous use of the EC clause may recur, including with respect to: evenhandedness (how to ensure that not only members favored by powerful shareholders were treated as exceptional); the risks of proliferation (how to avoid cases being unduly proposed to be "exceptional"); and the Fund's leverage to ensure sufficiently strong programs if EACs were not met. Finally, some suggested the problems with the EAP called for a move to a more "disclosure based" approach rather than use of an EC clause. Under such an approach, once debt sustainability was assessed positively, instead of having to make binary (yes/no) judgments related to criteria, the staff would instead present programs to the Board with a thorough disclosure of the trade-offs and risks associated with an EA program, its design, and the alternatives of providing normal access or of providing no program at all, and the related risks for the membership and the Fund.

A third concern relates to the assessment of EACs at program reviews. Before 2016, it was unclear whether the EACs needed to be assessed at each review. Staff reports for several early programs (for example, Latvia) did not present such assessments during reviews. Staff interviewees clarified that in fact the assessments were always checked, but not required to be explicitly discussed in the staff reports. Related, there is some evidence that EAC assessments may be less rigorous in reviews than at program approval, with assessments rarely changing once a program is underway and with changes rarely being so large as to change the thrust of the program, perhaps providing a false sense of security. The EPEs for the euro area programs questioned whether EAC2 was in fact met during program reviews that took place after firewalls had been established. But there are important qualifications, as in notable cases (including Argentina in 2019, Greece in 2013, Ukraine in 2014), programs were interrupted when staff assessed that one or more EACs were no longer met.

[42] De Las Casas and Pérez-Verdía (2024) analyze the 2018 SBA for Argentina from the perspective of the EAP.

EAC1

Program documents generally justified proposed access levels in terms of an "adding-up" exercise to fill residual financing needs rather than on an analytical or empirical basis. EAC1 does not clarify how access proposals are to be made beyond the principles noted in the Fund's conditionality guidelines that higher access would generally be associated with a stronger program, stronger track record of policy implementation, and stronger capacity to repay.[43] EA program documents do not clearly associate EA with these features.

It is unclear from the evidence whether access levels, which varied widely across EA programs, reflected differing circumstances or uneven treatment. Access levels varied from 103 percent of quota (Argentina, January 2003) to 3,212 percent of quota (Greece, 2010) and from 1.4 percent of GDP (Egypt, 2020) to 19 percent of GDP (Ukraine, 2015). While countries' circumstances were significantly different, cross-country analysis shows that among members with comparable BOP needs, some received EA support while others did not. All else equal, smaller and poorer countries were more likely to receive programs with NA rather than EA.[44] However, it is hard to ascertain if these facts point to uneven treatment as access levels depend on a variety of factors. Program documents do not regularly present cross-country comparisons of access levels. Area Departments do not maintain systematic information on cases where members expressed interest in an EA program but where an EA program was not agreed.

When debt is sustainable but not with high probability, that is, in the gray zone, program documents indicated no additional scrutiny of the implications of EA for future BOP stability. The empirical literature indicates weak, and sometimes negative, catalytic effects of EA when debt is in the gray zone. For example, while EA may attract private and official flows by signaling efforts to address macroeconomic problems, Bal Gündüz (2024) and Montiel, Cohen-Setton, and Li (2024) show that especially where debt is in the gray zone it may also deter private creditors whose claims are subordinate to the rising stock of preferred-creditor claims. However, access decisions require no additional scrutiny, such as realism checks, of the expected impact of EA on catalytic financing in such cases. Further, some empirical evidence suggests that larger outstanding IMF credit is correlated with a higher likelihood of successor programs, suggesting delays in problem resolution.

Conversely, when BOP conditions turned out stronger than the program envisaged, EAC1 usually was still deemed to be met. There was generally no mechanism other than reserve accumulation to capture the upside risk, such as exit strategies that considered reductions in access or a switch to precautionary programs supported by effective communications to avoid adverse market reactions. Only in a few cases—such as Brazil (2002), Hungary (2008), Latvia (2008), and Uruguay (2005)—did the authorities not draw fully on the approved access and treat the remainder of the program as precautionary. In most cases, EAC1 was considered met, disbursements continued as scheduled, and future repurchase obligations built up further. An alternative could be moving to clearer member statements about early repayment, in the event overperformance persists.

EAC2 AND EAC3

The design of EAC2 and EAC3 may not sufficiently recognize the strong links between them. Debt sustainability and market access prospects reinforce each other, and, in practice, their assessments have often been linked. These linkages were acknowledged when the EAP was formulated but the evaluation found that their significance has grown with the evolution of the EACs, and—especially since 2016—can imply a gap in the safeguard that they provide in the absence of an analytical framework for EAC3.

EAC2 is intended to support the EAP's objective in at least two important ways. First, it sets clear expectations on how public debt sustainability affects the IMF's lending decisions, including about when the Fund cannot proceed

[43] See IMF (2024a).

[44] Bal Gündüz (2024) substantiates this finding. With poorer countries often having relatively less institutional capacity, it might be harder for them to meet all of the EACs. If so, this would be an unintended consequence of the criteria, as the Board was clear in 2002 about the need to avoid a bias toward larger members that would be inconsistent with uniformity of treatment.

with EA if the proposed program is not designed to restore debt sustainability to the required standard under EAC2. Second, it seeks to ensure the risks are appropriately weighed with a view to designing a successful program and safeguarding Fund resources.

EAC2 represents a higher bar for debt sustainability in EA programs than in NA programs. EA programs require sustainability with "high probability" while for NA programs there is no such requirement (as long as the debt is not unsustainable). The assessments involve a measure of staff judgment, which has been supported by increasingly sophisticated tools, starting with basic deterministic DSAs focused on debt-GDP and gross financing needs thresholds, the MAC DSA (since 2012), the HP tool (an internal tool to assign probabilities to debt sustainability, since 2015), and the Sovereign Risk and Debt Sustainability Framework (SRDSF) to replace the MAC DSA in 2022. These tools contributed to an improvement in the rigor and consistency of the assessments over the evaluation period.

EAC2 has evolved in ways that can be seen to reduce the stringency of the safeguard it provides. In 2002, the DSA reflected an assessment of ex ante debt sustainability. In 2009, it became "forward-looking" by recognizing policy changes under the program that had yet to be implemented. In 2010, it provided for the systemic exemption. In 2016, while the systemic exemption was removed, other ways (such as re-profiling) were introduced for members to secure EA when debt was in the gray zone. The introduction of the gray zone recognized that in uncertain circumstances where it is hard to assess debt sustainability accurately, a reprofiling may be less costly than a deep debt operation aimed at restoring debt sustainability with high probability. Nevertheless, when debt operations are needed, the requirements for meeting EAC2 differ for red and gray zone cases: if debt is unsustainable (red zone), the debt operation must restore sustainability to "high probability;" while if it is in the gray zone there is no such requirement (unlike in the case of the PRGT-EAP during the evaluation period). Given that EAC2 is the only criterion that substantively differentiates EA from NA requirements, the above trend suggests a lowering of EAP's relative standards. The fact that debt can remain sustainable but not with high probability beyond the program period may reduce the strength of program adjustment and the urgency of debt restructuring even where it is warranted. At the same time, this feature provides the Fund with flexibility for continued program engagement in difficult and rare cases where the resolution of problems might take time.

The definition of the gray zone has been hard to communicate. While the term "sustainable but not with a high probability" has a logic within the framework of the EAP and DSA, it has proved confusing to many outside the Fund, including market participants. A particular point of confusion has concerned whether the point of the assessment is "sustainable" or "not with a high probability." While the term may have been intended to provide reassurance that debt was not unsustainable, in fact it often has had the opposite effect, generating uncertainty (see de Las Casas and Pérez-Verdía, 2024).

EAC2 does not fully recognize the differing risks posed by domestic versus external public debt nor by resident versus non-resident creditors, which reduces its clarity. While in 2009, the perimeter of the DSA was clarified as being public debt, both external and domestic, EAC2 itself makes little distinction between the differing risks associated with external versus domestic debt. The SRDSF—which applied to cases after the sample for this evaluation—makes some of these distinctions; and experience with the SRDSF may help to provide sharper guidance for how to apply the criterion going forward.[45] Further, although domestic debt is a major component of public debt in many countries, when DSAs are close—notably in gray zone cases—the staff has tended to focus on external debt when estimating the fraction that is "restructurable."

EAC3 was intended to ensure that Fund financing was temporary, and that the member used it to recover market access, return to external sustainability, and be able to repay the Fund. Members are required to have prospects to access private capital markets by the time their repurchases to the Fund start falling due and until the Fund has been repaid. For NA programs, by contrast, the requirement under the Fund's conditionality guidelines is that the program has prospects to strengthen the BOP (without specifying whether that occurs through the

[45] Erce (2024) provides some further discussion of the SRDSF in this context.

current account or capital market access). This may reflect the origins of the EAP in capital account crises.

EAC3 is ambiguous about whether it refers to international or also to domestic market access and is also silent about other relevant features. While EAC2 is about public debt (both external and domestic), EAC3 is about market access without specifying if that refers only to external market access or also domestic market access. Staff views have differed on this point. On one side, the logic of EAC3 is understood as regaining BOP stability and repaying the Fund, which requires access to international markets. Another view, informed by euro area programs and others, understands that domestic market access is also directly relevant for the member's capacity to repay the Fund when non-resident investors are involved in them. Indeed, the fact that the government had some form of domestic market access was used to justify EAC3 in several instances (such as Argentina, 2018). Given the linkages between EAC2 and EAC3 and given the explicit focus of EAC2 on both external and domestic public debt, the ambiguity in EAC3 bears clarification.

The Fund lacks an analytical framework to assess the fulfillment of EAC3. Unlike EAC2, staff lacks the guidance and tools to analyze market access prospects in a systematic, comparable, and sound way. Evidence shows that this has led to the adoption of a variety of inconsistent approaches and assessments of this criterion across cases. It has been hard to assess market access prospects for countries (such as Jordan) that previously did not have market access, and for those (such as Greece) that had reasonable expectations of long-term official support. In some cases, backward-looking assessments have sufficed for meeting EAC3. The criterion does not indicate on what terms market access should be considered (re)gained, for example, that access should be on sustainable terms in order to be consistent with BOP stability and the member's capacity to repay the Fund. Several authorities made the point that programs should seek to ensure that market access is regained in ways that manage the risks associated with the inflows before the crisis. They noted that private investors are often quick to come back to troubled environments in search of high returns, assuming that risks are safeguarded by the Fund program. EAC3 assessments could be informed by assessment of the member's current debt management practices to identify key constraints to balancing the cost of financing with prudent risk management, including refinancing or rollover risks.

The absence of consistent guidance for EAC3 effectively loosens EAC2. Since 2016, EA has been possible even if debt is in the gray zone provided that other forms of financing are available. One such form of financing is market access. Favorable market access assumptions, therefore, support debt sustainability. The absence of an analytical basis or consistent guidance for how to assess market access prospects is a gap in the framework. In practice, EAC3 judgments have often been backward looking, based on current or recent access (Ecuador 2020, Egypt 2020), rather than on prospects. Separately, some stakeholders noted a tension in how restructurable debt is considered in staff's assessments of EAC2 and EAC3. The assessments tended to view a sufficiently high share of restructurable debt as strengthening assessments of EAC2. But the signal sent by treating private debt as restructurable could risk undermining the prospects for regaining private market access.

EAC4

EAC4 has often been interpreted very narrowly in terms of political and institutional capacity to implement the program, rather than on the overall soundness of program design. This criterion requires reasonably strong prospects of success, including not only the member's adjustment plans but also its institutional and political capacity to deliver that adjustment. That is, EAC4 requires assurances about prospects of program success, taking into account national ownership and implementation capacity. In practice, however, assessments of this criterion have often focused on the member's political and institutional capacity, rather than on establishing that program design is solid enough to provide reasonably strong prospects of success in resolving the member's BOP problem. From this angle, EAC4 would be the most important of the EACs, requiring that the program as a whole, including aspects related to the other three criteria, will address the member's problems in a timely and effective manner.

Relatedly, EAC4 does not establish a substantively higher expectation for program success relative to NA programs. Any Fund program should have at least "reasonably strong" prospects of success in order to be approved, and it should be realistic in terms of the authorities' political and institutional capacity to implement the adjustment. In line with the conditionality guidelines, which assert that higher access should be accompanied by stronger safeguards, it seems prudent to require stronger ex ante justification of program adequacy for EA programs,[46] taking into account the tensions that may exist between the requirement of a reasonable prospect for program success and the elevated program risks and challenges in EA cases.

The Fund does not have a framework or consistent guidance for how to assess prospects for program success, including institutional and political capacity. Prospects are assessed more or less on an ad hoc basis, justified by country circumstances. But that makes it hard to assess the soundness of assessments, the standards, and comparisons. Suggestions for such guidance raised during the evaluation included taking into account the member's history of Fund engagement and program experience with certain measures (which may help inform assessments of ownership and prospects for success relating to similar measures), the full suite of DSA realism tools, and the adequacy of social protection measures, as well as tailoring program design, duration, and instruments to the nature of BOP pressures.[47]

EAC4 assessments were particularly difficult in cases with potential changes in government during the program period. The Fund's conditionality guidance has clear advice on gaining political assurances, including if elections are upcoming. Nonetheless, meaningful assurances can be hard to obtain ahead of elections. This difficulty can be exacerbated if the program's phasing involves frontloaded disbursements but backloaded adjustment. The staff teams nonetheless sought and received political assurances in a variety of forms—from private assurances by stakeholders across the political spectrum to presidential or parliamentary letters supporting the program—and in many cases set out the associated risks clearly in the staff reports—including, for example, for Argentina (2018) and Ecuador (2020).

[46] Higher relative safeguards already apply in other parts of the Fund's lending framework; for example (although in a different context), the FCL has more demanding qualification criteria than the PLL.

[47] The conditionality guidance (IMF, 2024a) requires staff to discuss two subsets of the SRDSF realism tools on (i) fiscal adjustment and growth; and (ii) growth and the output gap. To strengthen the higher evidentiary standard for EA programs, more robust realism checks could include substantiating EAC4 by discussing the full set of SRDSF realism tools and integrating them with the macro framework (Bal Gündüz, 2024; Erce, 2024; and Giugale and Bal Gündüz, 2024 discuss this point further).