

CHAPTER 4

GAINING TRACTION—THE IEO’S INFLUENCE AT THE IMF

YASEMIN BAL GÜNDÜZ

INTRODUCTION

As called for in its Terms of Reference (TOR), the Independent Evaluation Office (IEO)’s objectives are to enhance the learning culture within the International Monetary Fund (IMF), strengthen the IMF’s external credibility, and support the Board’s institutional governance and oversight responsibilities. The IEO seeks to achieve these objectives through its contributions to changes in work practices, shifts in priorities in use of budgetary resources, adjustments in policies or development of new policy frameworks, and longer-term efforts to reshape the culture of the institution. While changes may not be solely, or even primarily, motivated by IEO evaluations and recommendations, the IEO complements the self-review and self-evaluation work by the IMF to enhance the institution’s ability to draw lessons from its experience and integrate improvements into its future work.

Taking stock of experience with IEO recommendations from 16 IEO evaluations over 2011–21, this chapter discusses the IEO’s influence on the IMF, focusing on the formal process for following up on recommendations made by IEO evaluations. It explores the extent to which IEO recommendations were endorsed by the Board, the alignment of the recommendations and the Management Implementation Plan (MIP) actions, the implementation of MIP actions, and the factors affecting the time to implement.

A key objective of this chapter is to take stock of the strengths and weaknesses of the current follow-up process with a view to identifying how to further improve the traction and impact of IEO evaluations. In that regard, this chapter complements the discussion of the impact of IEO evaluations in Chapter 2. It provides an overview of the follow-up process embedded in the IEO’s theory of change and how this process has evolved over time, discusses the traction of recommendations, and presents concluding remarks.

THE FOLLOW-UP PROCESS

As discussed in Chapter 2, the formal follow-up process to implement evaluation recommendations endorsed by the Executive Board has evolved and strengthened over time, primarily driven by the feedback from the periodic external evaluations of the IEO. In 2007, in response to the first external evaluation of the IEO (Lissakers, Husain, and Woods 2006), known as the “Lissakers” report, the IMF’s Executive Board approved a more systematic approach for following up and monitoring the implementation of IEO recommendations by establishing MIPs and Periodic Monitoring Reports (PMRs). Further changes to this process were approved following the second and third external evaluations.

Current Process

The key components of the follow-up process—the summing up of the Board discussion, the MIP, and the PMR—can be understood by considering a theory of change (ToC) for how IEO evaluations impact the institution. A ToC is used by evaluators to depict:

- ▶ The causal pathway from activities to outputs to a sequence of outcomes to final desired impacts; and
- ▶ The causal assumptions showing why, and under what conditions, the various links in the causal pathway are expected to work; that is, what has to happen for the causal linkages to be realized (Mayne 2015).

The ToC shown in Figure 4.1 depicts how IEO interventions (evaluations) are meant to work to help the IMF be more effective in supporting its membership to achieve high-level goals of sustainable economic growth and macroeconomic and financial stability.

The results chain, or the causal pathway, of the IEO evaluation work is as follows (Figure 4.1, textboxes in green):

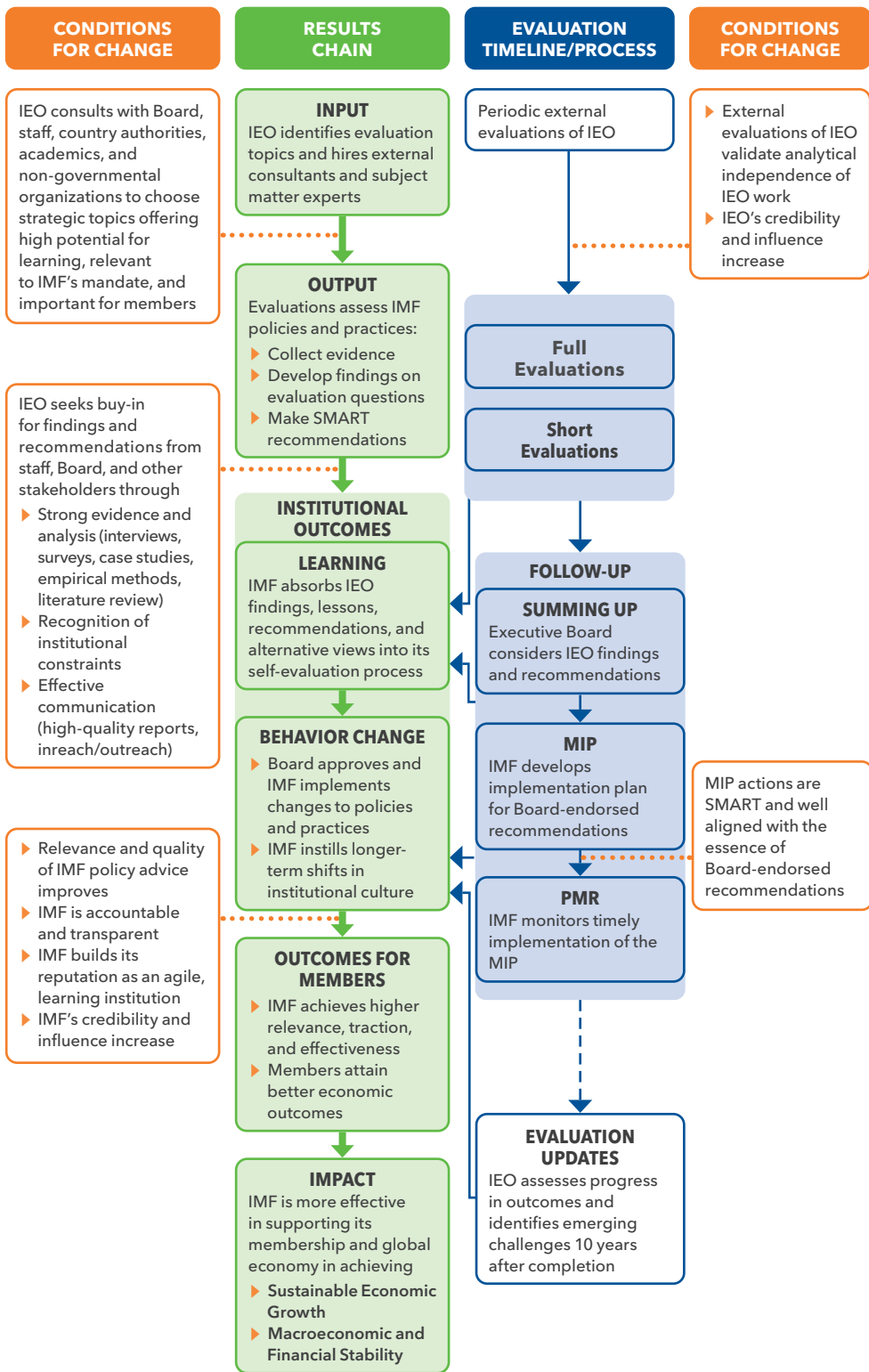
- ▶ Inputs and outputs are fully under the IEO’s control. In consultation with stakeholders, the IEO Director selects a strategic evaluation topic with high learning potential and decides on the evaluation format—a full or a short evaluation. The IEO sets up an evaluation team, which collectively possesses evaluation and subject matter expertise, experience in macroeconomic policy analysis, and institutional knowledge. The evaluation team collects evidence through various evaluation methods (interviews, surveys, case studies, empirical methods,

and literature review), analyzes/triangulates the evidence, and produces an evaluation report. Recommendations are typically included in full evaluations, while short evaluations may present recommendations if judged necessary by the IEO. Ten years after the completion of a full or a short evaluation, the IEO may undertake an evaluation update. Updates take stock of progress in areas identified by the original evaluation, assess whether the lessons and recommendations still remain relevant, and identify emerging issues and new challenges. Depending on the findings of an evaluation update or a short evaluation, the IEO may decide to launch a full evaluation.

- ▶ Outcomes are expected to take place sequentially at two levels: institutional outcomes for the IMF and outcomes for members.
- ▶ Institutional outcomes are the direct contributions of the IEO evaluation work to the IMF's learning and behavior change. All full evaluations and short evaluations with recommendations are reinforced by a rigorous follow-up process: the IMF develops MIPs for the Board-endorsed IEO recommendations and annually prepares monitoring reports on the implementation of the MIP to the Board. Nevertheless, IMF learning and behavior change are expected to ensue not only from the specific recommendations and the follow-up process, but also from the broader range of IEO findings in all evaluations and evaluation updates.
- ▶ Outcomes for members are the higher relevance, traction, and effectiveness of IMF advice to members as a result of improved institutional outcomes and ensuing better immediate economic outcomes for members. Economic outcomes include variables such as tax-to-Gross Domestic Product (GDP) ratio, government spending, exchange rate policy, and macro-critical structural reforms.
- ▶ Impact represents the highest level of results in the results chain, corresponding to the contributions of IEO evaluations to the IMF's mandate of providing support for members and the global economy aimed at achieving sustainable economic growth and macroeconomic and financial stability. Attribution at the impact level is more difficult to assess than at the outcome level given the wide range of factors affecting a country's economic performance.

The orange textboxes in Figure 4.1 present the causal assumptions, that is, the conditions under which the outcomes and the desired impact will be realized. A successful transition through the results chain, thereby an impactful evaluation,

FIGURE 4.1. IEO'S THEORY OF CHANGE



depends on whether the following conditions for change are accomplished at each stage:

- ▶ **Input.** Fully informed by the diverse views of stakeholders, the IEO chooses a strategic evaluation topic offering a high potential for learning, relevant to the IMF’s mandate, and important for members.
- ▶ **Output.** The IEO prepares the conditions for effective buy-in for its findings and recommendations from staff, the Board, and other stakeholders, which requires the following: the evaluation report is of high quality in substance and presents strong evidence and analysis, well triangulated through a range of evaluation methods; takes into account binding institutional and resource constraints; has SMART (Specific, Measurable, Attainable, Relevant, and Timely) recommendations; and clearly conveys its findings and the rationale for recommendations. Furthermore, the IEO interacts closely with staff during the evaluation to identify constraints to implementation and communicates its findings and recommendations effectively through in-reach/outreach activities.
- ▶ **Institutional outcomes.** Improved learning and behavior change at the institutional level would lead to better outcomes for members provided such changes are well aligned with the essence of the recommendations and improve the relevance and quality of IMF advice to members. IEO evaluations and observable changes in institutional outcomes help improve accountability and transparency of the IMF and build its reputation as an agile and learning institution. As a result, the IMF’s credibility and influence increase.
- ▶ **Outcomes for members and impact.** Higher relevance and quality of IMF advice combined with stronger accountability, transparency, and credibility help the IMF achieve higher traction with members, better economic outcomes, and higher impact.

Notably, the IEO and the IMF jointly learn from and implement recommendations of periodic external evaluations of the IEO. Such independent “evaluations of the evaluator” validate the independence of the IEO work and assess the traction of Board-endorsed recommendations, helping improve the IEO’s credibility and influence.

The current roles and responsibilities of the various participants in the evaluation and the follow-up process are as follows:

IEO

- ▶ **Topic selection.** The IEO prepares a menu of possible evaluation topics, which is discussed with the Board in an informal seminar. The IEO Director decides on the evaluation topics, in consultation with Executive Directors and management, as well as with informed and interested stakeholders outside the IMF. Topics are focused on issues of importance to the IMF's membership and of relevance to the IMF's mandate, in line with current institutional priorities but without interfering with the IMF's operational activities. The IEO conducts the evaluation and formulates its recommendations, drawing on input from management, staff, the Executive Board, country officials, subject matter experts, and other stakeholders. The IEO typically also includes more detailed sub-recommendations or suggestions on how the top-line recommendations could be achieved.
- ▶ **Board discussion.** The IEO responds to directors' comments and questions and reviews the summing up (SU) to help ensure that it accurately reflects the Board discussion.
- ▶ **MIP.** As input to the Evaluation Committee discussion of the MIP, IEO assesses how well the MIP actions align with the corresponding IEO recommendations, whether the actions are SMART, and whether the actions are adequate for achieving the goals sought by the recommendations.
- ▶ **PMR.** IEO provides input to staff preparing the PMR and issues a statement ahead of the Board discussion to raise any remaining concerns regarding the information provided about implementation status. During the internal review process of the PMR, IEO also reviews and comments on staff proposals to retire open MIP actions.

Management and Staff

- ▶ **Topic selection.** Management and staff can suggest evaluation topics and comment on the IEO's menu of options.
- ▶ **Evaluation.** Management approves staff comments on the draft IEO report and recommendations for IEO's consideration during the internal review process.

- ▶ **Board discussion.** In preparation for Board discussion of the IEO evaluation report, the Managing Director provides a written statement of her position on each IEO recommendation, which is circulated to the Board along with any written comments on the more detailed specific suggestions made by the IEO.
- ▶ **MIP.** Within six months of the Board discussion, management and staff are required to set out a forward-looking implementation plan identifying actions to follow up on Board-endorsed IEO recommendations, together with a timeline and cost estimates.
- ▶ **PMR.** Once a year, staff prepares a PMR on the state of implementation of actions contained in approved MIPs and not judged completed in the previous PMR. Since 2014, the Office of Internal Audit (OIA) has been responsible for preparing PMRs and reporting to the Board. The OIA reports directly to the Managing Director and has no direct operational responsibility or authority over any of the activities audited or reviewed.

The Executive Board

- ▶ **Topic selection.** Executive Directors can suggest evaluation topics and comment on the IEO's menu of options.
- ▶ **Board discussion.** The IEO's evaluation report is discussed by the Executive Board. Executive Directors typically issue written statements laying out their reaction to evaluation findings and recommendations. Under the IMF's "rule of silence," if they do not comment on one or more recommendation(s), that is interpreted as support for the Managing Director's view. Directors may also comment on the more specific suggestions made by the IEO evaluation. Directors may then further refine or even alter their views during interactions at the Board meeting. A summing up (SU) of this meeting conveys the response of the Board to the evaluation and its support for the recommendations made by the evaluation. Directors are given the opportunity to ensure that SUs accurately reflect their views.
- ▶ **MIP.** The Board endorses the MIPs, usually on a lapse of time basis, after review by the Evaluation Committee to ensure that IMF management and staff are adequately pursuing the Board-endorsed recommendations. The Evaluation Committee may request changes to the draft MIP to provide stronger or timelier follow up to Board-endorsed recommendations.

- ▶ PMR. Since 2019, the PMR is discussed and approved by the Board (prior to 2019, the PMR discussion was delegated to the Evaluation Committee). Decisions to retire action items can be made at each PMR discussion. A triage exercise to identify actions to be reformulated took place on a one-off basis in 2019.

EVOLUTION: HOW DID WE GET HERE?

Concerns raised by external evaluations of the IEO helped improve the key components of the follow-up process.

Recommendations

Articulation of IEO recommendations provides the jumping-off point between the evaluation and follow up on its findings and conclusions. Formulation of concise and well-targeted recommendations helps lay the basis for clear decisions by the Executive Board about whether and how the IMF should respond. Over time, the IEO has received varying advice and feedback on its recommendations—including, on the one hand, calls to provide more specific proposals to address issues it identifies and, on the other, to keep its advice at a high level and refrain from micro-managing the institution.

The second external evaluation of the IEO (Ocampo, Pickford, and Rustomjee 2013) henceforth referred to as the “Ocampo report,” expressed concern that IEO recommendations had become “too process-oriented” and called for recommendations to “be aimed at the outcomes to be achieved by the IMF, leaving it to Management and the Board to design the appropriate actions to deliver those outcomes.” (Ocampo 2013, p. 3).

In 2018, by contrast, the third external evaluation found that some recommendations had been “too general and hard to implement,” with “too little attention given in IEO reports to explaining the usefulness and effect of the recommendations made,” and too much attention to laying out evidence and “pointing out gaps and shortcomings” (Kaberuka, Jiun, and Meyersson 2018, pp. 10–11). The report, henceforth referred to as the “Kaberuka report,” thus called for the IEO to design “impactful” recommendations that are practical, as well as SMART. The report essentially asked the IEO to achieve a balance between competing priorities—calling for the IEO both to provide directions that were “broad enough ... to give management room to find the best way to implement,” and to articulate proposals that were “fully appropriate in the operating context” and provided “adequate guidance” for development of follow-up plans.

Summings Up

The initial draft summing up is prepared in advance of the Board discussion by IMF staff in the Secretary's Department, in consultation with the IEO, the Strategy, Policy, and Review Department (SPR), and other IMF departments as relevant, drawing on Executive Directors' written statements (known as "grays"). The explicit role of the IEO in the summing-up process was strengthened following the Kaberuka report in 2018. Most substantively, the IEO has the opportunity to comment on Secretary's Department's initial drafts of the summing up ahead of, as well as after, the Board meeting.

MIPs

MIPs were introduced in 2007 following the first external evaluation of the IEO, the "Lissakers report." The report identified the lack of a formal mechanism for the Board to follow up specific recommendations made by the IEO after observing that "the implementation of IEO recommendations appears to remain firmly under the control of Management. Unless Management initiates follow-up, nothing happens." (p. 25) The Lissakers report suggested several ways to strengthen monitoring and follow-up that envisaged a more active role for the Board and the Evaluation Committee: "One would be for the IEO periodically to issue, for Board discussion, an evaluation of follow-up and implementation of its recommendations and findings on specific issues that appeared to enjoy support from the Board. Another measure would be for the Evaluation Committee to play a more active role. Soon after an IEO report has been presented and discussed by the Board, the Evaluation Committee could schedule a meeting with Management or relevant department heads, together with the team leader for the IEO, to discuss what follow-up is being planned or considered. The Evaluation Committee could continue to monitor and report back to the Board" (p. 26).

Subsequent external evaluations have led to substantial reinforcement of the MIP process. The Ocampo report cautioned about the significant delays from the finalization of the IEO evaluations to the approval of the MIP and recommended that "Management should present the MIP within one to three months following the Board discussion. The Evaluation Committee should review the MIP on behalf of the Board and ensure that it is approved no later than two months after its release" (p. 34). Subsequently, it was decided to introduce a requirement to present the draft MIP within six months of the Board discussion, to provide adequate time for staff to develop a comprehensive plan to meet objectives while being consistent with the IMF's broader agenda and work processes. The Kaberuka report recommended that the Evaluation Committee should enforce clear expectations for Management

Implementation Plans to meet the SMART test. Subsequently, management and staff have been committed to be tightly guided by the SMART principles to accountability.

Periodic Monitoring Reports

With the introduction of a formal follow-up process in 2007 came a PMR, initially prepared by SPR, on the status of implementation of actions included in the MIPs. Beginning with early reports, directors raised questions about the extent to which recommendations were being addressed without the impact always being seen. Some issues, such as staff tenure on country teams, were raised repeatedly during PMR discussions and kept “on the table” for monitoring. The Ocampo report pointed to the conflict of interest in management’s dual responsibilities in preparing the MIP and monitoring its application. The report recommended that “Periodic Monitoring Reports should continue to be prepared annually, but by the Internal Audit Office (OIA) ... given its mandates, its experience in evaluation, its arm’s-length relationship with management and the capacity and skills base of its staff” (pp. 26 and 34). Beginning in 2014, preparation of the PMR—and thus the assessment of whether actions had been completed—was moved to the OIA. Relatedly, the Ocampo report did not support a more central role for the IEO in the follow-up process, noting that “the comparative advantage of the IEO is to identify policy issues and desirable outcomes ... [Monitoring of specific recommendations] is not its primary responsibility and would detract from its core objectives and mandates” (pp 25–26). The Kaberuka report found that the PMRs prepared by the OIA were detailed and professional, and that “candor and usefulness” had continued to improve with each report (p. 16).

At the Board discussion of the first PMR in 2008, directors agreed to refer future PMRs to the Evaluation Committee for initial consideration and review in order to make more efficient use of the Executive Board’s time. According to this framework, if the Evaluation Committee discussion did not raise any issues, the committee would recommend endorsement of the PMR by the full Board on a lapse of time basis. The committee could also recommend a Board discussion of the whole PMR or just a few issues that would merit further consideration. The Kaberuka report recommended that discussion of the PMR be restored to a formal Board meeting with management and staff present. The Kaberuka report elaborated that “The current practice is for the EVC to discuss the PMRs, but ... neither management nor the staff responsible for the implementation of open items have been required to explain the delays in person to the EVC; instead it has been the OIA, charged with monitoring implementation, and SPR, that have been in attendance... the current

BOX 4.1. CATEGORIZATION OF OPEN ACTIONS IN MIPS IN 2019 TRIAGE FRAMEWORK

Category 1. Actions that are open with no obvious challenges; routine follow up in the PMR. These actions are clearly defined, measurable, and monitorable; progress is being made and they are expected to be implemented in reasonable time.

Category 2. Actions for which desired outcomes are not being achieved despite implementation guidelines being in place; revisions are needed to embed greater incentives or stronger enforcement. Management will propose revisions that seek to achieve the same goal as the original action, while enhancing monitorability, embedding compliance incentives, and/or strengthening accountability mechanisms.

Category 3. Actions that are insufficiently specific and have no clear measures of success; revisions are needed to make them SMART (Specific, Measurable, Attainable, Relevant, and Timely). Management will propose revisions to the actions to make them more specific, measurable, attainable, relevant, and time bound.

Category 4. Actions that involve long-term technical or culture change and are not suitable for a binary open-closed classification; better addressed through other mechanisms. Actions in this category would be followed up in five reporting cycles of the PMR, following their original approval, to review progress. After this period, management will propose (with an explanation of the progress to date) that the action should be retired from the PMR monitoring process, unless a longer monitoring period is required—in which case the period of PMR monitoring would be extended. If needed, at the time an action is proposed to be retired, management, jointly with the accountable departments, would identify other mechanisms of the Board where future updates could be available (for example, a Board policy or administrative review that looks into the broad topic of concern).

Category 5. Actions for which full implementation is no longer seen as valuable, have duplicates in other MIPS, or have become redundant; retire from PMR. Management will propose that such actions should be retired from the PMR because it has encountered significant challenges or other factors that make full implementation unlikely (with no obvious remedy), or because the actions have become redundant, having been superseded by subsequent events, priorities, or MIPS.

In 2020, the Board placed eight actions into categories 2 and 3 to retain them under PMR monitoring, as these actions were not fully covered under other workstreams, emphasizing particularly the reformulation of actions on staff tenure and handover of country assignments. The Board asked staff to take a holistic approach to ensure mission teams' deep understanding of country circumstances, and their strong relationships and traction with member countries. These eight actions were reformulated in 2021. The Board retired 15 actions under categories 4 and 5 while stressing that most of these actions remained highly relevant. Five out of six actions in category 4 had met the five-year PMR reporting requirement and had been substituted by other implementation workstreams and monitoring mechanisms. For one action in category 4 on regional and gender diversity, annual monitoring through the PMR was extended for five additional PMR cycles to help provide additional progress on regional and gender diversity objectives. Ten actions under category 5 were retired as they were assessed to be superseded by new workstreams and further PMR monitoring was considered redundant.

Source: IMF (2019, 2020, and 2021).

process gives insufficient attention and importance to the implementation progress” (pp 16–17). Since 2020, PMRs have been considered in formal Board meetings.

The 2019 Triage Framework

Despite improvements in the PMR process, the number of actions to be monitored continued to grow. Actions in new MIPs outpaced the implementation of outstanding actions as some items, particularly from earlier MIPs (before the greater emphasis on SMARTer commitments), continued to lag. This motivated the most recent innovation in the process: to triage open actions by determining whether some were no longer relevant and could be retired, and identifying those that were still important and needed new impetus. The Kaberuka report stated clearly that “too many items have remained open for too long,” noting that action on some was difficult to assess because they were “too broadly phrased” or were focused on general calls to strengthen culture in a continuing way; others remained open “simply due to poor enforcement and insufficient accountability,” and some may have been “overtaken by events and therefore are no longer relevant in the current form” (p. 16). The report recommended that “the Board should comprehensively address the backlog of open management actions on the basis of an enhanced Periodic Monitoring Report containing recommendations to deal with each category of open action items” (p. 22).

In 2019, the Board approved a framework to address open management actions in response to Board-endorsed IEO recommendations (IMF 2019). The framework was implemented in “Categorization of Open Actions in Management Implementation Plans” (IMF 2020), which undertook a triage of the backlog of open actions into five broad categories (Box 4.1). As a result, 15 open actions were retired, 1 item was retained, and 8 items were to be reformulated. In January 2022, the Board endorsed the MIP for eight reformulated actions, which included a number of outcome indicators for monitoring and reporting progress (IMF 2022b). These two reports were prepared by the OIA, in consultation with the IEO, SPR, and other IMF departments.

Recurring Issues

Over the years, some observers have pointed to a weakness in the follow-up process, namely that it tracks only specific actions arising from formal recommendations, without considering whether sufficient attention is brought to bear on broader issues raised by IEO evaluations that also represent important areas for learning and change. The Ocampo report articulated that “the more generic and substantive issues raised by IEO reports, which are often not encapsulated in specific recommendations, also need monitoring... this type of review should be carried out biennially by the IEO itself...[and] should be selective and concentrate on the ‘bigger issues’ that are particularly relevant... It would be for the Board to decide whether further actions are needed in light of these IEO reviews” (p. 26). The IEO subsequently prepared a review of recurring issues raised by IEO evaluations (IEO 2014). (Progress on these issues is discussed in Chapter 2.) The Board supported the recommendation that the IEO conduct similar evaluations every five years and that staff prepare a separate high-level report on the status of initiatives, addressing the recurring issues identified by the IEO. Staff prepared a High Level Status Report in 2016 on progress made in addressing these issues. However, in light of the need to streamline the work program and the existence of other monitoring processes, the formal process for tracking recurring issues was discontinued.

Evaluation Updates

IEO evaluation updates, introduced in 2013, provide a vehicle to assess how issues raised in earlier evaluations may have evolved and point to areas that may need further attention. However, updates do not include recommendations and are only discussed informally by the Board, rather than for action, so they do not have a formal follow-up process.

TRACTION OF IEO RECOMMENDATIONS

This section covers the traction of recommendations from IEO evaluations completed during 2011–21. Two phases of traction will be covered sequentially: first, the Board consideration of recommendations; and second, the follow-up of MIP actions related to the Board-endorsed recommendations. This analysis covers 16 evaluations for which monitoring data is available in the Twelfth PMR (IMF 2022a).

Discussion of the first phase covers the extent to which each recommendation in IEO evaluations was endorsed by the Board. It specifically takes stock of recommendations that were rejected or only partially accepted, and why. The discussion of the second phase presents some descriptive analysis on the nature of IEO recommendations versus MIP actions, using two metrics: proximity to higher-level results (result score); and the depth of institutional change (IC-depth score). Finally, the traction of the MIP actions and the factors affecting the time to implement the actions will be explored, using descriptive statistics as well as some empirical methods.

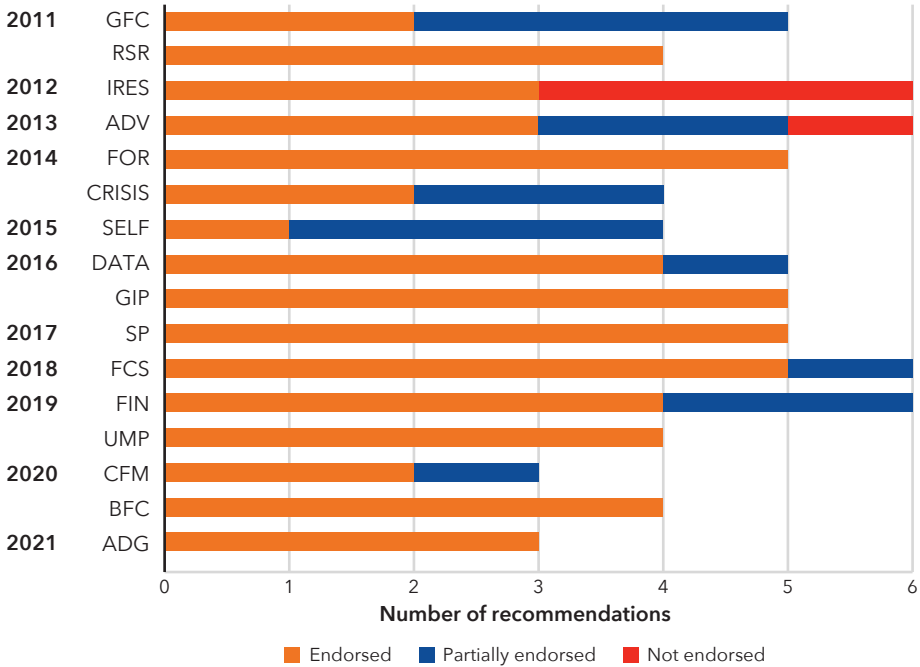
The Board Response to IEO Recommendations

Overall, IEO recommendations appear to have been selective and focused on key issues. Across 16 evaluations completed over 2011–21, the IEO made 75 recommendations with 3 to 6 recommendations for each evaluation (Figure 4.2). On average, the Board fully endorsed almost three-quarters of the 75 recommendations. Only 4 recommendations were not endorsed while 15 recommendations were endorsed with qualifications.¹ Annex 4.2 presents the recommendations not endorsed or only partially endorsed, along with the Board response in the summings up.

The Board rarely rejected IEO recommendations outright. For evaluations completed in the earlier part of the decade, the Board did not endorse three recommendations in the evaluation of *International Reserves: IMF Concerns and Country Perspectives*, and one recommendation in the evaluation of *The Role of the IMF as Trusted Advisor*. In addition, some sub-recommendations in the evaluation of *IMF Performance in the Run-Up to the Financial and Economic Crisis* were not endorsed. The key reasons stipulated in the summings up for not endorsing some

¹ In two evaluations, *IMF Performance in the Run-Up to the Financial and Economic Crisis* and *The Role of the IMF as Trusted Advisor*, some top-line recommendations were not sufficiently specific. Therefore, a bottom-up approach based on sub-recommendations is used for these evaluations. Accordingly, a recommendation is considered endorsed if all relevant sub-recommendations were endorsed by the Board, and not endorsed if all relevant sub-recommendations were rejected by the Board. A recommendation is considered partially accepted if it is not wholly accepted or rejected.

FIGURE 4.2. BOARD RESPONSE TO IEO RECOMMENDATIONS, 2011-21



Source: IEO evaluation reports and the author’s calculations.

Note: Covers all recommendations for 16 evaluations completed over 2011-21.

The acronyms used for the evaluations are presented in Annex 4.1.

recommendations, and thus not having any corresponding MIP actions, were as follow:

- ▶ Already addressed by existing policies, products, or analytical tools. Three recommendations contained in the evaluation of *International Reserves: IMF Concerns and Country Perspectives* fell under this category: first, “targeting policy distortions directly,” second, “embedding the discussion of reserve accumulation in the multilateral context in a more comprehensive treatment of threats to global financial stability that was informed by developments in global liquidity and financial markets,” and third, “Policy initiatives that were meant to deal with systemic externalities must take into account the relative size of countries’ contributions to the externality” (pp. 17–18). Directors agreed with these approaches but considered that these were already addressed by existing IMF policies, tools, and multilateral surveillance products, such as the External Sector Report, spillover reports, and reserve adequacy assessments.

- ▶ Limited value. With regard to *The Role of the IMF as Trusted Advisor*, directors did not see the merit of the recommendation to “Incorporate early and openly the views of all countries—particularly those that stand to be most affected by changes in the Fund’s policy stance—during the preparation of major policy papers on which analytical debate is still ongoing” (p. 29). Directors supported the need for extensive consultation with the IMF membership, but there was limited support for setting out all views in policy papers. Similarly, the directors were wary of introducing a more bureaucratic process they saw resulting in the context of this sub-recommendation: “in close consultation with country authorities, develop a medium-term strategic plan which is expected to be actively used as a guidance for subsequent mission teams and would allow Management to monitor continuity as well as progress with key medium-term objectives” (p. 28).
- ▶ Disagreement with substance. The Board did not endorse the substance of two sub-recommendations in *IMF Performance in the Run-Up to the Financial and Economic Crisis*. The first stated that “In order to promote more effective bilateral surveillance, consideration must be given to the possibility of issuing staff reports without the need for Board endorsement. This could be followed by a peer review process structured to give surveillance greater traction” (p. 22). The second stated that “On issues of systemic importance, the Fund should be ready to err more often in the direction of emphasizing risks and vulnerabilities, rather than focusing on possible benign scenarios. This change in approach would need to be discussed and agreed by the membership at large” (p. 23). On the latter, a number of directors cautioned that the recommendation to err more often emphasizing risks and vulnerabilities could lead to more false alarms and thereby reduce the credibility and traction of surveillance. In *The Role of the IMF as Trusted Advisor*, directors did not support the sub-recommendation of “developing incentives for mission chiefs and resident representatives that make their role as trusted advisors an important part of their performance,” as they were concerned about how such incentives might affect the independence of staff advice (p. 29).
- ▶ Weak or vague formulation. The Board accepted management’s view that this sub-recommendation in *IMF Performance in the Run-Up to the Financial and Economic Crisis* was vague: “Encourage the staff to be more candid about the ‘known unknowns,’ to be more ready to challenge their own preconceptions, and to frankly disclose the limitations of data and technical tools underlying its analysis” (p. 22).

Comment and Suggestions

It is worth emphasizing that the extent to which the Board endorses IEO recommendations should not be considered as a performance metric for the success of an IEO evaluation. Rather, the IEO's key role of “speaking truth to power” with an independent and well-informed voice is critical for helping the IMF achieve its mandate. Delivering timely and, at times, uncomfortable messages is a crucial part of the job.

Nevertheless, a recommendation that is not endorsed does not lead to any explicit follow-up process and is thus likely to have limited traction, unless it somehow contributes to changing mindsets. Some of the factors leading to a negative response from the Board were due to unforced errors that the IEO has been addressing in its second decade. The main lessons are as follows:

- ▶ Improve the clarity and precision of recommendations. Clarity and precision should be sought in two dimensions: The first is complementarity. When making recommendations on IMF policies, the IEO should recognize whether they overlap with or complement the existing policy and practices, while highlighting what else is needed and gaps in implementation. The evaluation of *IMF Advice on Capital Flows* provides a good example of an evaluation clearly identifying the gaps in the IMF's policy and providing timely, clear, and specific policy recommendations for updating the IMF's Institutional View (IV) on capital flow measures, although not all of IEO's detailed suggestions were accepted in a subsequent review of the IV. Moreover, an explicit recognition of how recommendations overlap with existing policies and recent initiatives may strengthen staff buy-in by giving credit to staff work. A good practice would be to report the baseline of existing policies and practices, both at the beginning and end of the evaluation, reflecting any evolution over the course of the evaluation. A recent example was the introduction of the Resilience and Sustainability Trust as the evaluation of *IMF Engagement with Small Developing States* was being prepared (IEO 2022). The second dimension regards SMART formulation. As discussed extensively in the previous section, following the Kaberuka report, the IEO has been paying particular attention to formulating SMART recommendations, focusing on providing detailed, actionable suggestions for how top-line recommendations can be implemented.
- ▶ Engage more closely with stakeholders to increase buy-in. Some cases of “limited value” or “disagreement in substance” could be avoided by closer engagement with staff and the Board during the evaluation, especially when formulating the recommendations, to fully understand their concerns and

potential constraints to implementation. Such close engagement does not necessarily, and should not, come at the expense of the IEO's independent voice. The IEO has increased efforts to reach out to stakeholders to get their views on recommendations and explain IEO's rationale for them. At times, IEO recommendations were met with defensive responses from staff, especially when perceived as requiring additional budgetary resources in an environment of "no real increase" in the institution's budget and staff already overstretched. In that regard, the IEO has made recommendations on how the IMF can economize in some areas to create space for changes that will require more staff and budgetary resources. Examples include its recommendations in the evaluation of *IMF Financial Surveillance* for more flexible, dynamic, and risk-based allocation of Financial Sector Assessment Program resources; its recommendations in *Behind the Scenes with Data at the IMF: An IEO Evaluation* on prioritizing the IMF's data requirements and weighing the benefits and the costs for any additional requests; and its recommendation in *IMF Engagement with Small Developing States* for further streamlining the application of data-demanding diagnostic tools.

- ▶ Avoid too many recommendations. It is important to focus recommendations on a few prominent issues with high potential impact, well supported by the evidence, and actionable by the IMF. This allows the Board discussion to focus on the most important high-level recommendations while allowing management and staff to develop SMART actions that take institutional constraints into account

Recommendations vs. MIP Actions: Were They Aligned?

The MIPs for 16 evaluations completed over 2011–21 had 185 actions to address the IEO recommendations. An important question is to what extent these MIP actions responded to the essence of the corresponding IEO recommendations. This question is multifaceted and hard to tackle with high confidence. The alignment of recommendations and the MIP actions can be examined at three levels: policy content at the output level; results achieved at the outcome level; and the depth of institutional change.

Policy Content

There are three crucial points in time for assessing the alignment of MIP actions and recommendations in terms of policy content. The first is at the time of the MIP endorsement by the Board; the second is at the time of the Board approval of changes to the related policies and practices committed by the MIP; and the third

is at the time of the PMR approval. One important point to emphasize is that many MIP actions relate to processes rather than end results: in many cases, the MIP actions involve preparing policy papers to establish or update IMF policies and practices. Such MIP actions do not prejudice the substance and conclusions of the corresponding papers but note that the IEO recommendations will be considered. Therefore, at the time of the MIP approval, it is not possible to fully assess whether the final policy changes will be aligned with the essence of the IEO's recommendations. In addition, at the time of the changes to policies, the IEO is not involved in the internal review process for the policy papers, nor does it make a statement to the Board conveying its views on the alignment of policy changes with its recommendations. At the time of the PMR approval, the OIA reports on implementation on the basis of whether such a policy paper was discussed and approved by the Board. However, the OIA and the IEO do not formally assess or validate the extent to which the policy paper and final approved policy changes align with the essence of the IEO's analysis behind its corresponding recommendations.

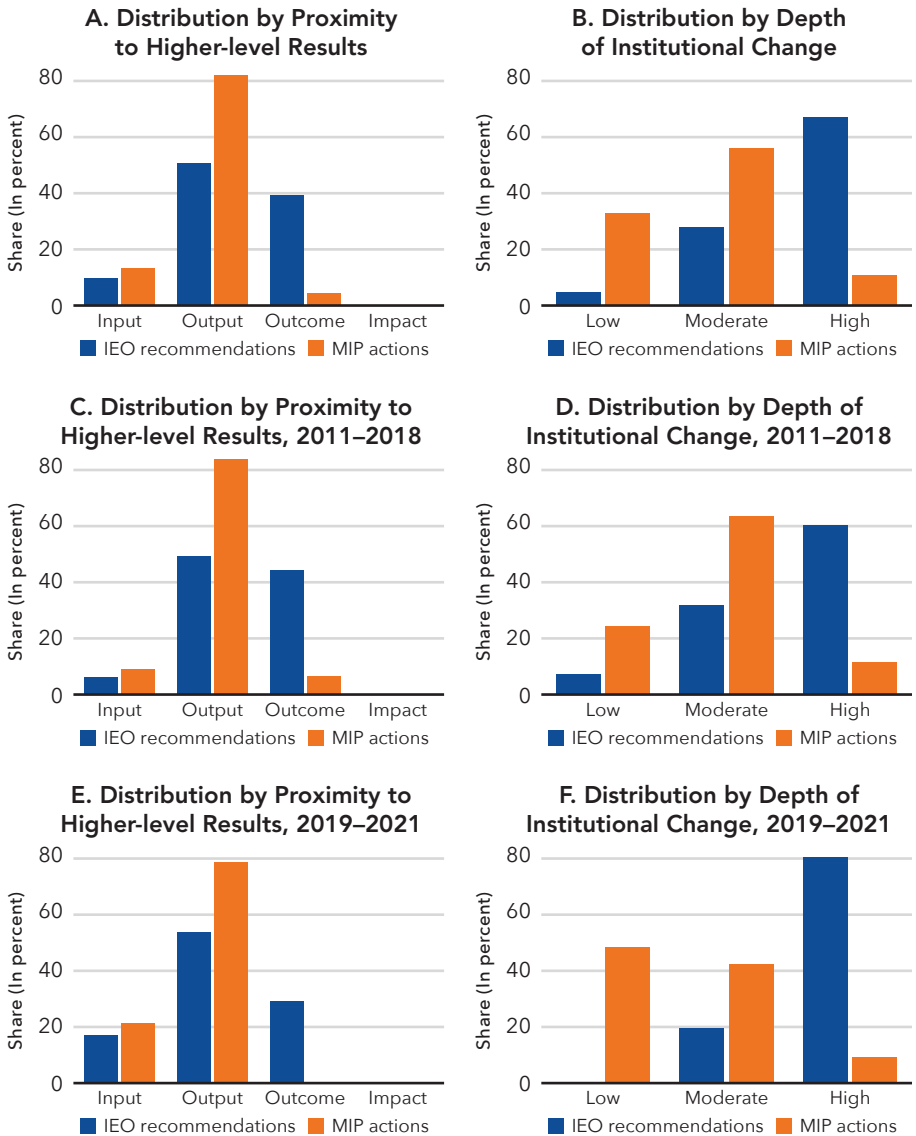
Results and Institutional Change

We analyze the alignment of recommendations and MIP actions based on two metrics introduced in this chapter²: proximity to higher-level results and depth of institutional change. The results metric refers to the (implicit) results chain or the theory of change for the IMF intervention evaluated by the IEO and assigns a rating corresponding to whether the recommendation is pitched at the input, output, outcome, or impact level respectively, on an escalating order of one to four. Ideally, such ratings would be assigned based on an explicit theory of change for the IMF intervention, prepared as part of the evaluation by the IEO in consultation with the stakeholders. The metric showing depth of institutional change assesses the significance of the institutional change on a rating scale of one to three, from low or no depth, moderate depth, to high depth, in escalating order. Each metric is applied to 75 IEO recommendations and 185 MIP actions covering all evaluations completed between 2011 and 2021.

The distribution of ratings by proximity to higher-level results is quite distinct for IEO recommendations versus MIP actions (Figure 4.3, panel A). While the IEO recommendations cover a range of actions primarily tilted towards output and outcome levels, the MIP actions addressing the Board-endorsed recommendations are specified overwhelmingly at the output level. The follow-up process by the PMR monitors whether the outputs are delivered but does not assess or monitor whether

² The details of these metrics along with their respective rating scales and examples are presented in 4A.3.

FIGURE 4.3. DISTRIBUTION OF RATINGS FOR RECOMMENDATIONS AND MIP ACTIONS BY PROXIMITY TO HIGHER-LEVEL RESULTS AND DEPTH OF INSTITUTIONAL CHANGE, 2011-21



Sources: IEO evaluations, MIPs, and the author's calculations.

Note: Covers all recommendations and MIP actions for 16 evaluations completed over 2011-21.

the intended outcomes for underlying recommendations are being achieved by the MIP actions.

Notably, none of the recommendations or the MIP actions were formulated at the impact level over 2011–21. While it's hard to assess the contribution of IMF interventions at the impact level, the recent evaluation of *Growth and Adjustment in IMF-Supported Programs* (IEO 2021) employed a range of empirical methods to estimate the impact of IMF-supported programs on economic growth. The evaluation found that IMF-supported programs over 2008–19 yielded growth benefits relative to a counterfactual of no IMF engagement, and that stabilization and reforms implemented in the program context boosted post-program growth performance. *Growth and Adjustment in IMF-Supported Programs* made three recommendations to further enhance program countries' capacity to sustain activity while undertaking needed adjustment during the program period, and to enhance growth prospects beyond the program.

While the IEO recommendations involved institutional changes predominantly at high depth, MIP actions were tilted more heavily towards moderate to low-depth actions (Figure 4.3, panel B). As explained earlier, at the time of the MIP approval, the degree of policy changes to be contemplated in the final Board paper is not known with certainty, and such an assessment can be made accurately only when the new or revised policy is approved by the Board. The assessment here is based on the formulation of actions at the time of the MIP approval. While a thorough retrospective assessment of approved Board papers is beyond the scope of this chapter, anecdotal evidence points to instances of substantial differences between the approved policy paper and the IEO recommendations. An example is referenced earlier in this chapter on *IMF Advice on Capital Flows* (IMF 2002a) regarding the updated Institutional View on capital flows.

To assess the effect of the shift towards SMARTer recommendations and MIP actions in 2018, Figure 4.3 panels C through F presents the same histograms for evaluations completed before and after 2018. The proximity of actions to higher-level results and their institutional depth both seem to have been affected, as no MIP action was formulated at the outcome level and the metric on depth of institutional change was heavily tilted towards low-depth activities. The composition of IEO recommendations appears to have shifted somewhat to inputs at the expense of outcomes. On the other hand, the depth dimension of IEO recommendations improved significantly, suggesting that the post-2018 recommendations envisaged greater transformational changes in the policy content.

Comments and Suggestions

Key lessons from experience have already started to influence the IEO's and the IMF's approach to formulating and monitoring the implementation of MIP actions and the Board-endorsed recommendations. Recent progress and some ideas for the path forward in the follow-up process are as follows:

- ▶ **Monitor outcomes.** Recommendations and MIP actions could be linked more explicitly to the intended higher-level results, where feasible and cost effective. As highlighted in the results-based management literature, when setting performance expectations for outcomes, it is crucial to avoid distorting behavior by focusing on higher-level outcomes that are closely related to the true objectives of the activities, and by regularly reviewing and updating the outcome indicators to ensure they remain relevant and useful and are not causing perverse behavior that could undermine the achievement of high-level objectives (Mayne 2007). In most cases, the MIP commitments aim at changing IMF policies and practices, which are outputs and interim steps necessary, but not necessarily sufficient, to achieve the desired outcomes.

To address this issue, the recent MIP reformulating eight long-standing open actions in previous MIPs (IMF 2022b) more systematically incorporated some outcome indicators to stipulate how to assess success for MIP actions at the outcome level (for example, regular surveys with specific and time-bound targets for satisfaction metrics are being used to monitor progress in improving handover of country assignments and to assess satisfaction of country authorities with IMF country teams' handling of confidential information).

It should also be recognized that in some cases, monitoring outcomes may not be feasible or cost effective. For example, a MIP action for the evaluation of research at the IMF (IEO 2011) was to “adopt new procedures to ensure the quality of working papers” to improve the technical quality of analytical work. Although new guidelines for research were introduced in 2012, this action had remained open owing to no process for measuring compliance with the guidelines and difficulty in assessing the quality of working papers. The action was reformulated in 2022 but noted that measuring and monitoring the quality of individual working papers

was neither feasible nor meaningful from a cost-benefit analysis and in light of limited resources.³

The Kaberuka report also weighed in on outcome monitoring for recommendations involving culture change. The report classified recommendations into two groups: first, recommendations that are more specific and actionable with a shorter time frame; and second, qualitative recommendations aimed at changing culture, which are likely more long-term in nature and where success in implementation is less directly measurable. The report suggested separating follow-up processes for these two groups. It noted that while simply monitoring the status of implementation as open or closed would suffice for the former, the latter would require identifying proxy or qualitative measures of impact, and tracking continued progress over several years.

- ▶ Assess impact where feasible. It is notoriously difficult to monitor the achievement of higher-level results at the impact level. This is especially so at the IMF, whose mandate is to provide support for members and the global economy aimed at achieving sustainable economic growth and macroeconomic and financial stability. The methods used in *Growth and Adjustment in IMF-Supported Programs* offer a promising avenue to explore the impact of various IMF interventions through one particular lens—that of sustaining and fostering growth in the program context—but also illustrate the empirical challenges involved.
- ▶ Validate substance at the output level. To help bridge the gap between outputs and outcomes, the IEO could provide comments, in a focused manner on drafts for the selected, key policy papers listed among the MIP actions and convey its views to staff on whether the policy paper—hence, the suggested revisions to IMF policies and practices—was aligned with the essence of the IEO recommendations and whether any divergence was justified from the IEO’s perspective. The IEO could then issue a written assessment to the Board ahead of the Board discussion, which would support the Board’s oversight of the implementation of Board-endorsed recommendations. A similar process does occur with the PMR prepared by OIA. The IEO has an opportunity to comment on the draft report circulated to departments for review, and then provides a

³ The reformulated MIP action includes the following output-level commitments: “Identify and disseminate good practices on the departmental review of WPs; update Fund guidelines on publication of WPs and departmental procedures on the review of WPs; increase transparency and accountability for the interdepartmental review of WPs; and improve outlets for disseminating country analytical work.” (IMF 2022a, p. 13)

written comment circulated to the Board for discussion. The IEO could also use public channels to convey its views more broadly. As discussed in Chapter 2, a recent example is an opinion piece by Korinek, Ostry, and Loungani (2022) on the update to the Institutional View on capital flows (IMF 2022c). The authors, two of whom took part in the evaluation of *IMF Advice on Capital Flows*, conclude that the updated framework includes welcome fixes, but major rethinking is still needed.

Traction of MIP Actions

Traction, as measured by the implementation of the MIP commitments, has been quite strong, but with uneven progress across evaluation topics. The OIA, in its PMRs, assessed that 128 out of 185 actions included in the MIPs for 16 evaluations completed during 2010–21 were implemented: thus, the overall completion rate is already around 70 percent, and this rate will go up as actions in the more recent implementation plans are achieved.⁴

Each evaluation had 5–18 MIP actions, with the lower and the upper end of the range recorded by *International Reserves: IMF Concerns and Country Perspectives* and *Behind the Scenes with Data at the IMF: An IEO Evaluation*, respectively (Figure 4.4, panel A). With one exception, all of the actions for the evaluations completed during 2011–15 were fully implemented, retired, or reformulated: for one action, the Board extended annual monitoring for five additional PMR cycles to help provide additional progress on regional and gender diversity objectives.⁵

The average implementation rate was 84 percent for the evaluations completed during 2011–15, although only *International Reserves: IMF Concerns and Country Perspectives* and the evaluation of *Self-Evaluation at the IMF: An IEO Assessment* were fully implemented. The implementation rates for others ranged between 65 percent for *IMF Performance in the Run-Up to the Financial and Economic Crisis* to 91 percent for the evaluation of *IMF Response to the Financial and Economic Crisis*. (Figure 4.4, panel B).⁶ Notably, fully implemented evaluations completed

⁴ The implemented actions include three retired and two reformulated actions deemed completed in the categorization exercise and the Twelfth PMR, respectively.

⁵ The action formulated in response to a recommendation in *The Role of the IMF as Trusted Advisor* and proposed to be retired in 2019 was to “implement measures to raise the share of nationals from underrepresented regions among senior staff and set targets for representation of women at senior levels.” (IMF 2021, p. 5)

⁶ The time to implement each MIP action is calculated in years from the date of the MIP approval to the date of the PMR confirming the action as completed.

during 2011–15 had fewer than 10 MIP actions while those with the lowest implementation rates had 11 to 17 MIP actions.

For the evaluations completed over 2016–21, the average implementation rate was 71 percent. The implementation rates across evaluations ranged from full implementation for *IMF Advice on Unconventional Monetary Policies* to 10 percent for *Growth and Adjustment in IMF-Supported Programs*, the most recent evaluation covered in the Twelfth PMR and completed only in 2021. According to the Twelfth PMR, which examined 88 open actions from 10 MIPs over 2015–22, 39 were completed, while 49 remained open, of which 12 were overdue by more than one year.⁷ While the completion rate for the more recent six years seems a bit lower, this is largely because less time has gone by. In fact, the rate at which actions are completed has increased for recent evaluations.

As for the pace of implementation, *IMF Advice on Unconventional Monetary Policies* was the front-runner, with all of its actions fully implemented in three years after the MIP approval, followed by *IMF Performance in the Run-Up to the Financial and Economic Crisis*, with a 91 percent implementation rate in three years. Notably, several evaluations achieved implementation rates around 80 percent in two years, including *The IMF and the Crises in Greece, Ireland, and Portugal*, *IMF Collaboration with the World Bank on Macro-Structural Issues*, *IMF Advice on Unconventional Monetary Policies*, and *Self-Evaluation at the IMF: An IEO Assessment*. This may suggest that the IMF attached a high priority to learning from the global financial crisis and the euro-area crisis, including the use of unconventional monetary policies, both through its own reviews already in the pipeline and through the IEO recommendations. At the other end, the lowest implementation rate after four years was for *Behind the Scenes with Data at the IMF: An IEO Evaluation*. Though its implementation rate caught up sharply in the Twelfth PMR, it remained the lowest in more than five years after the MIP approval. As highlighted in Chapter 2, the slow implementation for this evaluation is partly explained by broader delays in the review timetable due to the COVID-19 pandemic, as well as the complexity of the tasks involved.

⁷ The number of open actions reported in the Twelfth PMR differs from the actions classified as open by the methodology adopted in this chapter. The PMR dropped the retired and reformulated actions and added 22 new actions replacing 8 reformulated actions (IMF 2022b). The methodology described in Annex 4 classified some retired actions as open as they were superseded by other workstreams and/or more recent IEO evaluations. Furthermore, the reformulated actions were retained without adding the 22 new actions; however, the status of their implementations was tied to the full implementation of all new actions that replaced them. Finally, only the unique MIP actions were included to avoid double counting the same action listed as addressing several IEO recommendations.

Overall, the pandemic shock appeared to have affected some evaluations more than others: the IMF has made good progress in implementing the MIP actions in response to evaluations of *IMF Financial Surveillance*, *IMF Advice on Unconventional Monetary Policies*, *IMF Advice on Capital Flows*, and *IMF Collaboration with the World Bank on Macro-Structural Issues*, while progress on implementation for *Behind the Scenes with Data at the IMF: An IEO Evaluation*, *The IMF and Social Protection*, and *The IMF and Fragile States* have been more adversely affected. Amid the pandemic, progress on recommendations on *IMF Advice on Capital Flows* within the first year of the MIP approval is particularly noteworthy, a feat achieved only for two other evaluations, *Self-Evaluation at the IMF: An IEO Assessment*, and *Growth and Adjustment in IMF-Supported Programs*. Faced with pressing decisions to reallocate staff resources to pandemic-related work, the IMF nevertheless appears to have protected progress in its policy agenda on core areas of its mandate.

What Determines the Time to Implement?

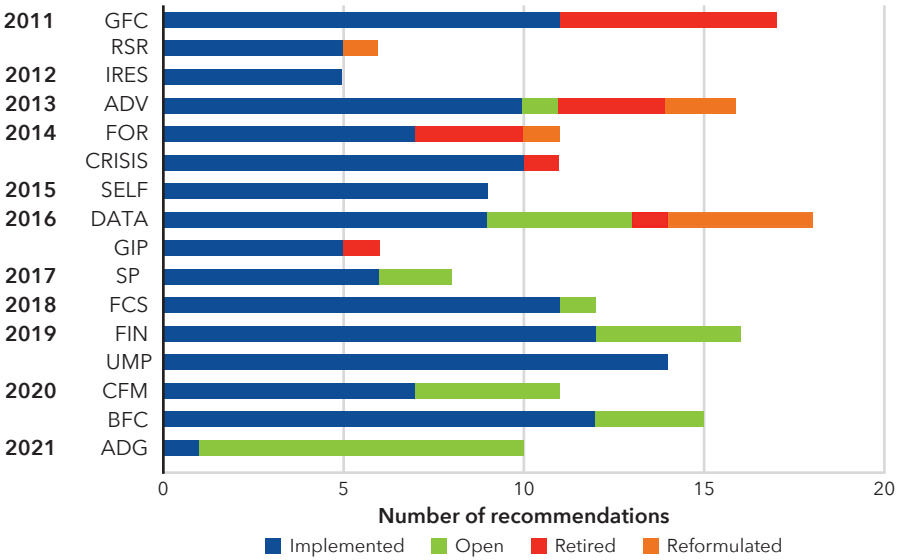
It is plausible to expect that the MIP actions that are closer to higher-level results and entail deeper institutional change may take longer to implement. In fact, the distributions of time to implement are remarkably similar for actions at the input and output level (Figure 4.5, panel A). In most cases, these actions were implemented during the second year after the MIP approval. However, at the outcome level, implementation has taken considerably longer—or was not achieved. Of eight MIP actions at this level, only five were completed, while one was retired and two were reformulated. All but one of the completed actions took at least six years to implement while the one implemented in one year was achieved by the long-awaited approval of the IMF's 2010 quota and governance reforms, an artifact of lucky timing.⁸ Time to implement tends to rise by the depth of institutional change (Figure 4.5, panel B).

For the 41 open MIP actions, the average time since approval, as of the Twelfth PMR, was 3.6 years. There were long-standing open actions at all three levels of results (Figure 4.5, panel C) and mostly at moderate and high institutional depth (Figure 4.5, panel D). Almost half of open actions were from the more recent IEO evaluations (*IMF Advice on Capital Flows*, *IMF Advice on Unconventional Monetary Policies*, *IMF Financial Surveillance*, *IMF Collaboration with the World Bank on Macro-Structural Issues*, and *Growth and Adjustment in IMF-Supported Programs*)

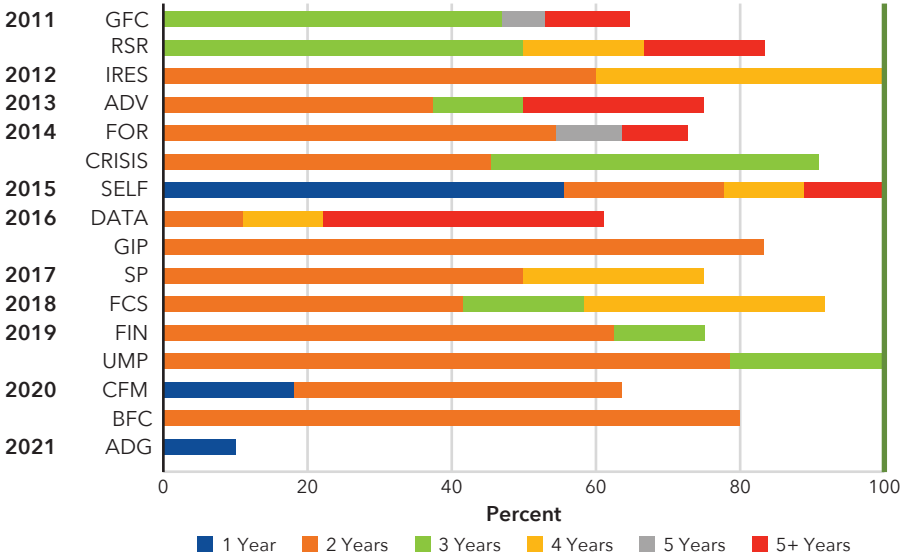
⁸ In survival analysis of factors affecting time to implement the MIP actions, this data point was treated as an outlier.

FIGURE 4.4. STATUS AND IMPLEMENTATION RATE FOR THE MIP ACTIONS

A. Number of MIP Actions by Implementation Status



B. Percentage Implemented Over Time



Source: Periodic Monitoring Reports and the author's calculations.

Note: Covers all MIP actions for 16 evaluations completed over 2011-21. Panel B presents the percentage of MIP actions implemented over one to five plus years from the date of the MIP approval. The implementation rate for five plus years is calculated as of the date of the Twelfth Periodic Monitoring Report and includes three retired and two reformulated actions deemed completed (4A.4).

and remained open for less than three years. Slightly more than a third of open actions were retired or reformulated, as explained in 4A.4, with an average time since approval of 5.8 years (ranging from 2.9 years for an action resulting from *The IMF and the Crises in Greece, Ireland, and Portugal*, to 7.7 years for 5 actions related to *IMF Performance in the Run-Up to the Financial and Economic Crisis*). The remaining open actions were 23 for evaluations completed after 2017, 4 for *Behind the Scenes with Data at the IMF: An IEO Evaluation*, 2 for *The IMF and Social Protection*, and 1 for *The Role of the IMF as Trusted Advisor*. The latter has been open since 2013 and is the longest-standing open action: it seeks to raise the share of nationals from underrepresented regions among senior staff and set targets for representation of women at senior levels.

A complementary perspective is provided by survival analysis, a frequently used empirical method well suited to analyze time to occurrence of an event, which allows using full information on both implemented and unimplemented actions, as well as exploring the effects of different factors in a multivariate analysis.

The results from survival analysis are presented in 4A.4. Compared to the base category in regressions represented by low depth output-level actions, the MIP actions at the input and outcome levels take about 10 percent and 215 percent longer to implement, respectively, though only the latter is significant. As for institutional depth, on average, it takes significantly longer to implement actions that involve institutional change at moderate and high depth compared to the base category, by about 60 percent and 72 percent, respectively. Finally, the MIP actions that involve long-term technical or culture change take 319 percent longer to implement than the base category.

Notably, a dummy variable for three input-level actions that involved structural shifts in the IMF's human resource (HR) structure and practices turns out to be highly significant, indicating the prolonged challenges in making progress in this area. These actions in MIPs for *IMF Performance in the Run-Up to the Financial and Economic Crisis* and *The Role of the IMF as Trusted Advisor* included the following:

- ▶ Continue efforts to broaden the professional diversity of staff, including their educational background and skill mix, with more staff being hired with financial sector experience;
- ▶ Target three years average tenure for country assignments, to be monitored by Human Resources Department (HRD) and reported to the Board on a regular basis; and

- ▶ Implement measures to raise the share of nationals from underrepresented regions among senior staff and set targets for representation of women at senior levels.

Acknowledging the limited progress on these actions, in 2020 the Board extended annual monitoring for five additional PMR cycles for the last item and in 2021, reformulated the action on average tenure. The reformulated action strengthened the monitoring and accountability framework towards three-year average tenure for country assignments and introduced interim targets to be monitored at a more granular level across country groups. Moreover, by end-2023, staff will prepare an analysis of issues surrounding mission chief and country team turnover, including recent trends and drivers and their effect on tenure in country assignments, and propose policy options and targets, if necessary, to ensure adequate tenure in country assignments.

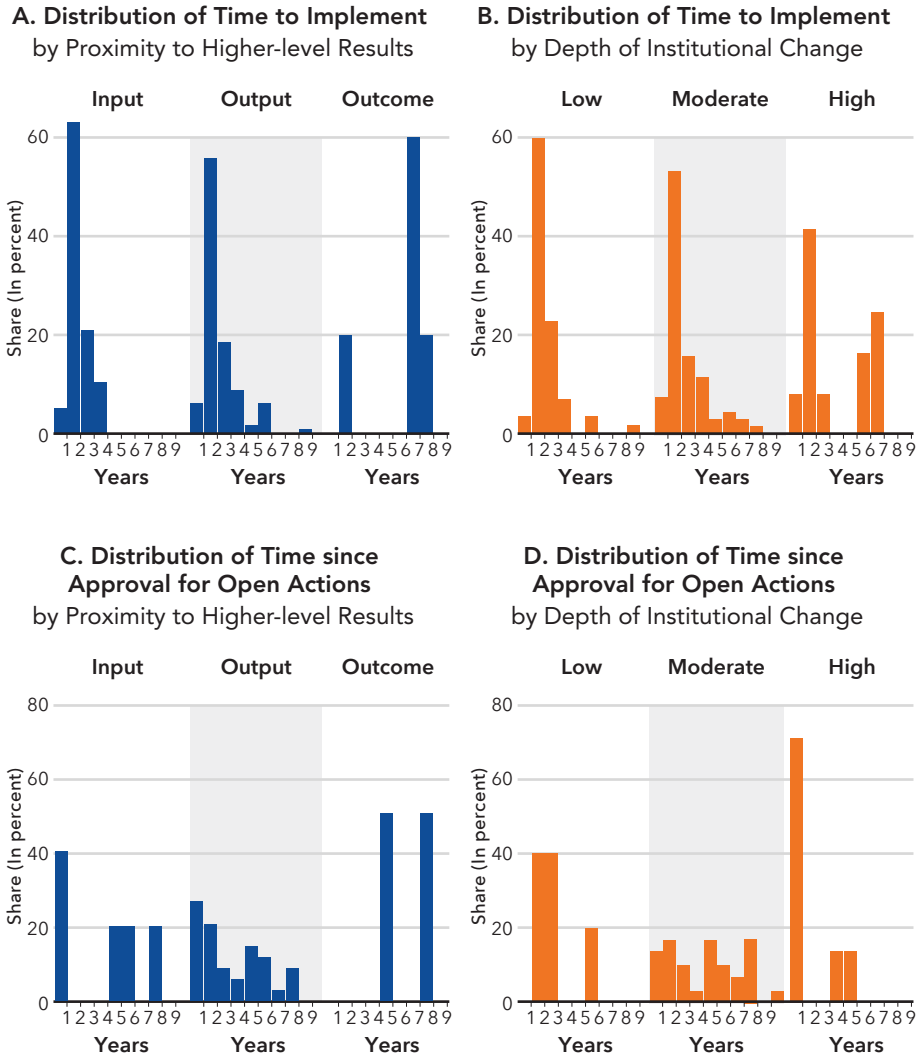
To control for the correlation within groups, dummy variables are included in the regression for each evaluation. After controlling for the effects for other explanatory factors, the results across evaluations highlighted significant differences in time to complete the MIP actions. At one end of the spectrum, the actions in response to *IMF Advice on Capital Flows* are estimated to have been completed significantly faster than for other evaluations, while at the other end, it appears that actions in response to *Behind the Scenes with Data at the IMF: An IEO Evaluation*, are taking the longest to implement.

According to the sensitivity analysis, input/output level actions that are low depth have a high predicted probability of being implemented in less than two years, while outcome level actions that also involve changing institutional culture still have a low probability of implementation even after eight years (sensitivity analysis, Figure 4.6). As for the mean estimated time to implement in years, when culture change is set to zero, the predictions range on average from 1.2 years for input/output level actions at low depth to 6.6 years for outcome level actions at high depth (4A.3). For actions that involve culture change, the predicted average time to implement increases to 5.1 to 9.6 years for input/output level actions while outcome level actions take significantly longer, at 16 to 27.5 years.

Comments and Suggestions

- ▶ Enhance monitoring of outcomes/culture change. The results suggest that for MIP actions that are specified at the outcome level, especially when they also involve culture change, the monitoring framework should be significantly extended, or other longer-term continuous monitoring and reporting mechanisms should be considered. Targets or milestones should be set where feasible, with due regard to what constitutes realistic but sufficiently ambitious progress over time. Furthermore, the IEO's evaluation updates that return to topics after around 10 years can be a useful tool to provide an assessment of progress with a longer horizon and at a deeper level than in the MIP and PMR processes.
- ▶ Sustain momentum on structural HR actions. Making progress on actions that involved structural shifts in the IMF's HR structure and practices has proven to be very challenging. It is critical to understand the root causes of the limited progress on such actions. A recently strengthened monitoring and accountability framework, as well as a forthcoming analysis of the drivers of staff turnover and policy options to address them, are welcome steps in the right direction, but staff and the IEO need to sustain the momentum on HR actions.

FIGURE 4.5. DISTRIBUTIONS OF TIME TO IMPLEMENT AND TIME SINCE APPROVAL FOR COMPLETED AND OPEN MIP ACTIONS BY PROXIMITY TO HIGHER-LEVEL RESULTS AND DEPTH OF INSTITUTIONAL CHANGE



Source: Periodic Monitoring Reports and the author's calculations.
 Note: Covers all MIP actions for 16 evaluations completed over 2011-21 and followed up in the PMRs. Reformulated actions, except for those under category 3, were included as still open as of the date of the categorization exercise. All but 4 of 15 retired actions were classified as open, with 3 completed and 1 dropped (4A.4).

CHALLENGES AND OPPORTUNITIES

The IMF has substantially strengthened the follow-up framework for the implementation of Board-endorsed IEO recommendations over time, driven primarily by recommendations in the external evaluations of the IEO's work. In-reach to staff to promote awareness of evaluation findings and lessons is also expected to complement the follow-up framework to support institutional learning but may not be as effective in overcoming institutional inertia and driving change. Attention to formulating SMARTer actions seems to have helped implementation but also tilted the actions more towards inputs and outputs rather than outcomes. Overall, the current framework is quite robust, with some state-of-the-art good practices and clear roles and responsibilities for management, the Board, and the IEO.

Nevertheless, this chapter identifies some challenges as well as opportunities to build on the progress already made.

Linking Outputs to Intended Outcomes

While the IEO recommendations cover a range of actions primarily tilted towards output and outcome levels, the vast majority of MIP actions addressing the Board-endorsed recommendations are specified at the output level. The follow-up process assesses whether the output has been delivered but has not systematically monitored whether the intended outcomes in the recommendations are being achieved by the MIP actions. Recommendations and MIP actions could be linked more explicitly to the intended higher-level results, where feasible and cost effective. It is encouraging that staff has started to incorporate some outcome indicators in the recent MIPs, which will help reverse the shift in the composition of actions towards input and output levels after the introduction of SMARTer actions.

Assessing Substance at the Output Level

MIP actions often provide commitments to Board reviews of IMF policies and practices (that is, specified at the output level), some of which are new and others already planned. Such MIP actions do not prejudge the substance and conclusions of the corresponding review papers but note that the IEO recommendations will be considered. At the time of the MIP approval, it is not possible to prejudge whether the final policy changes will be aligned with the essence of the recommendations. To address this gap, the IEO could provide input, in a focused manner, specifically providing comments to staff on drafts of selected policy papers listed among the MIP actions, conveying its views on whether the approved paper is

aligned with the IEO recommendations and whether any divergence was justified from the IEO's perspective. Ahead of the Board discussion, the IEO could issue a statement to the Board that assesses how well staff proposals meet the objectives of the recommendations, thus supporting Board oversight of the implementation of its recommendations.

Monitoring Impact

It is notoriously difficult to monitor the achievement of higher-level results at the impact level. This is especially so at the IMF, whose mandate is to provide support for members and the global economy aimed at achieving sustainable economic growth and macroeconomic and financial stability. IEO's 2021 evaluation, *Growth and Adjustment in IMF-Supported Programs*, employed a range of methods to estimate the impact of IMF-supported programs through one particular lens—sustaining growth in the program context. The use of such techniques offers a promising avenue for IEO's evaluations to explore the impact of various IMF interventions.

Cultural Change

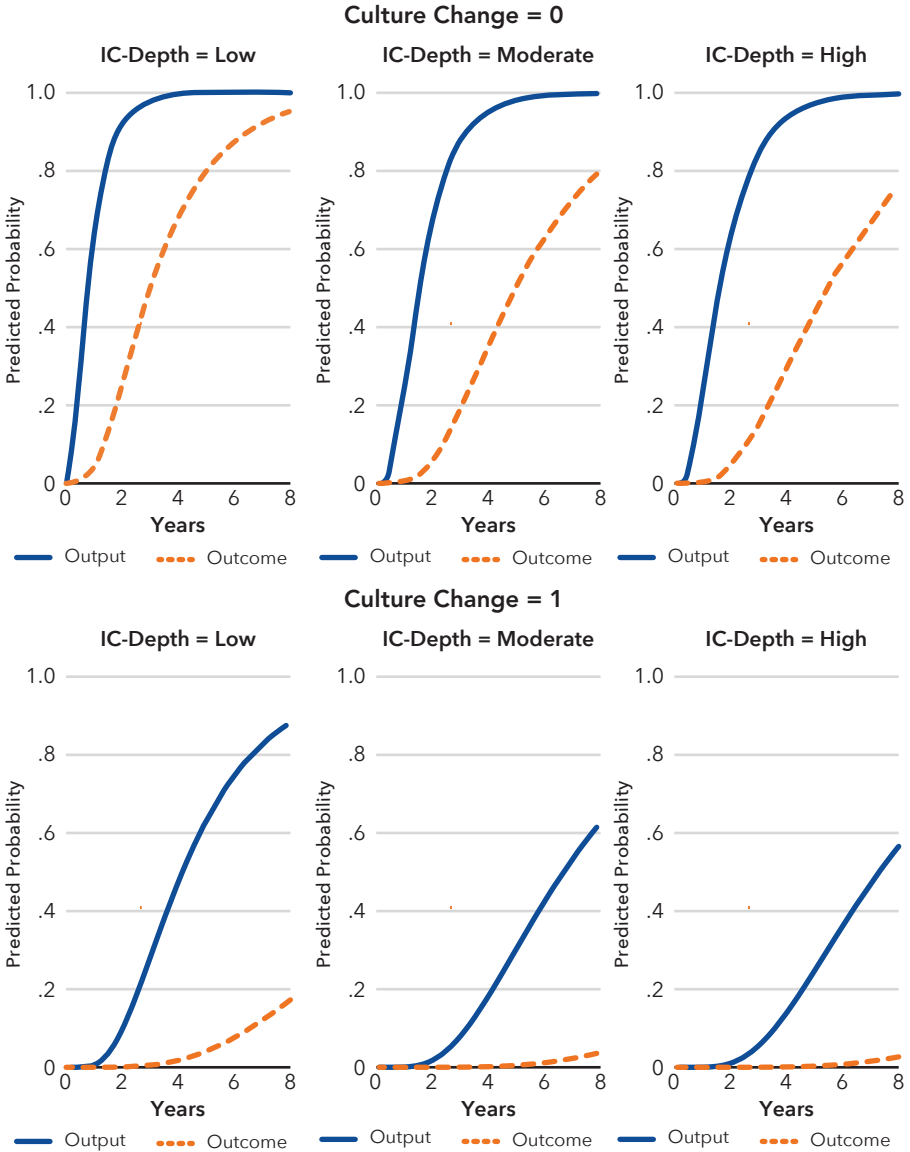
Recommendations that involve a change in the institutional culture may take a very long time to have an impact, and hard to monitor in substance. This suggests value in evaluation updates that return to topics after around 10 years, which can provide an assessment of progress with a longer horizon and at a deeper level than in the MIP and PMR processes.

Budgetary Implications

At times, recommendations are met with defensive responses from staff, especially when perceived as requiring additional budgetary resources in an environment of “no real increase” in the institution's budget and staff already overstretched. Similarly, recommendations that would require substantial reallocation of budgetary resources tend to take much longer to implement. The IEO could look more systematically into operational efficiency and cost-benefit analysis in its evaluations, presenting management and the Board with some options on what the IMF could do less of, or differently, in the area under evaluation, fully informed by the stakeholders' views. While the IEO may conclude that some increased resources may be justified to raise value added, it could also present prioritized recommendations intended to be budget neutral.

FIGURE 4.6. SENSITIVITY ANALYSIS: PREDICTED PROBABILITY OF IMPLEMENTING MIP ACTIONS

(Lognormal Regression Estimates)



Source: The author's calculations.

Note: Based on the benchmark lognormal regression in Table 4.2. Dummies for evaluations, structural HR actions, and two outliers are set to zero.

ANNEX 4.1

ACRONYMS FOR IEO EVALUATIONS

ADG	<i>Growth and Adjustment in IMF-Supported Programs (2021)</i>
ADV	<i>The Role of the IMF as Trusted Advisor (2013)</i>
BFC	<i>IMF Collaboration with the World Bank on Macro-Structural Issues (2020)</i>
CFM	<i>IMF Advice on Capital Flows (2020)</i>
CRISIS	<i>IMF Response to the Financial and Economic Crisis (2014)</i>
DATA	<i>Behind the Scenes with Data at the IMF: An IEO Evaluation (2016)</i>
FCS	<i>The IMF and Fragile States (2018)</i>
FIN	<i>IMF Financial Surveillance (2019)</i>
FOR	<i>IMF Forecasts: Process, Quality, and Country Perspectives (2014)</i>
GFC	<i>IMF Performance in the Run-Up to the Financial and Economic Crisis (2011)</i>
GIP	<i>The IMF and the Crises in Greece, Ireland, and Portugal (2016)</i>
IRES	<i>International Reserves: IMF Concerns and Country Perspectives (2012)</i>
RSR	<i>Research at the IMF: Relevance and Utilization (2011)</i>
SDS	<i>IMF Engagement with Small Developing States (2022)</i>
SELF	<i>Self-Evaluation at the IMF: An IEO Assessment (2015)</i>
SP	<i>The IMF and Social Protection (2017)</i>
UMP	<i>IMF Advice on Unconventional Monetary Policies (2019)</i>

ANNEX 4.2

IEO RECOMMENDATIONS NOT ENDORSED OR PARTIALLY ENDORSED BY THE BOARD

EVALUATION	IEO RECOMMENDATIONS NOT ENDORSED BY THE BOARD
GFC	<p>Recommendation (Rec). Actively seek alternative or dissenting views by regularly involving eminent outside analysts in Board and/or management discussions.</p> <p>Executive Board (EB). At the time the MIP was discussed, the Board did not support making this a more regular aspect of Board and/or management discussions.</p>
GFC	<p>Rec. Encourage the staff to be more candid about the “known unknowns,” to be more ready to challenge their own preconceptions, and to frankly disclose the limitations of data and technical tools underlying its analysis.</p> <p>Staff. Staff considered this sub-recommendation to be vague.</p>
GFC	<p>Rec. In order to promote more effective bilateral surveillance, consideration must be given to the possibility of issuing staff reports without the need for Board endorsement. This could be followed by a peer review process structured to give surveillance greater traction.</p> <p>EB. Not endorsed by the Board.</p>
GFC	<p>Rec. On issues of systemic importance, the IMF should be ready to err more often in the direction of emphasizing risks and vulnerabilities, rather than focusing on possible benign scenarios. This change in approach would need to be discussed and agreed by the membership at large.</p> <p>EB. No follow-up action identified. A number of directors cautioned that the recommendation to err more often towards emphasizing risks and vulnerabilities could lead to more false alarms and thereby reduce the credibility and traction of surveillance.</p>

EVALUATION	IEO RECOMMENDATIONS NOT ENDORSED BY THE BOARD
IRES	<p>Rec. Target perceived policy distortions directly.</p> <p>Managing Director (MD). "I believe the evaluation errs when it considers the rationale of the Fund and its membership, through the Board and IMFC, in undertaking work on reserves. This work was cast in the broader workstream on the international monetary system (IMS), and the approach to reserve issues was broad and multipronged, with emphasis placed on the development of a strong global safety net, the articulation of adequate policies to manage volatile capital flows, and the integration of the discussion of reserves with other external sector policies. Within the IMS, the demand for reserve assets matters in the context of their limited supply. Hence, the motivation was to make progress in assessing the right level of reserves for precautionary purposes given the need to balance their insurance benefits (demonstrated in this and past crises) with the fact that holding excessive reserves is subject to diminishing returns and can be costly both to the domestic and global economy." (IEO 2012 p. 27)</p> <p>EB. Directors noted with satisfaction the broad congruence between the IEO's recommendations and current IMF policies and practices. In particular, the latest triennial surveillance review calls for widening the scope of external stability assessments in bilateral surveillance, and the recently launched External Sector Report lays out, in a multilaterally consistent manner, the evolution of external imbalances across the largest economies, integrating advice on reserve adequacy with advice on related policy areas.</p>

EVALUATION	IEO RECOMMENDATIONS NOT ENDORSED BY THE BOARD
IRES	<p>Rec. Embed the discussion of reserve accumulation in the multi-lateral context in a more comprehensive treatment of threats to global financial stability that was informed by developments in global liquidity and financial markets.</p> <p>MD. “I find myself in agreement with most of the IEO’s formal recommendations, which are remarkably congruent with a number of recent Fund policy positions and new surveillance initiatives.” (IEO 2012 p. 27)</p> <p>Staff. “The Fund has already developed products to address this recommendation. This is particularly the case for the new External Sector Report. Further, the metric developed in the Assessing Reserve Adequacy paper also goes towards addressing this concern.” (IEO 2012 p. 30)</p> <p>EB. Executive Directors broadly supported the general recommendations presented in the evaluation, while recognizing that the IMF had already made progress in many of the areas the evaluation highlighted. Directors considered that there is a need to better understand the risks to the IMS arising from a buildup of reserves beyond levels driven by precautionary motives. Many directors shared the IEO’s view that excessive reserve accumulation is only a symptom of underlying distortions and stressed the importance of other factors like developments in global liquidity and capital flows. A number of other directors agreed with staff that it is justified to highlight that, in addition to being a symptom, reserve accumulation can be a potential cause of instability for the IMS, given the limited availability of reserve assets. Directors agreed that IMF surveillance, particularly multilateral surveillance, is well positioned to identify such risks and propose remedial action. A number of directors noted that the IEO’s report pointed to the need for more clarity in the IMF’s view on precautionary reserve holdings. In particular, they saw room for the IMF to better explain its work on reserve adequacy and external sector risks, both in its external communication and in its policy dialog with country authorities.</p>

EVALUATION	IEO RECOMMENDATIONS NOT ENDORSED BY THE BOARD
IRES	<p>Rec. Policy initiatives that were meant to deal with systemic externalities must take into account the relative size of countries' contributions to the externality.</p> <p>EB. Directors noted that IMF surveillance of systemically important countries has struck the right balance in deploying those tools, taking into account other factors that affect external conditions for these countries. As regards global spillovers from excessive reserve accumulation, a number of directors supported the IEO's recommendation that the relative contribution of a country to such externality be taken into account, without prejudice to the principle that all IMF members should be treated evenhandedly. This approach is already being applied in the IMF's multilateral surveillance products, including the spillover reports.</p>
ADV	<p>Rec. In close consultation with country authorities, develop a medium-term strategic plan which is expected to be actively used as guidance for subsequent mission teams and would allow management to monitor continuity as well as progress with key medium-term objectives.</p> <p>EB. Many directors did not support the proposal for medium-term strategic plans, on grounds that it could introduce more bureaucratic processes.</p>
ADV	<p>Rec. Develop incentives for mission chiefs and resident representatives that make their role as trusted advisor an important part of their performance.</p> <p>EB. A number of directors agreed on the need to develop incentives for staff to better act as a trusted advisor, while a few others were concerned about how such incentives might affect the independence of staff advice.</p>
ADV	<p>Rec. Incorporate early and openly the views of all countries—particularly those that stand to be most affected by changes in the IMF's policy stance—during the preparation of major policy papers on which analytical debate is still ongoing. Consult early with Executive Director offices to exchange views and keep them abreast of emerging policy frameworks, so as to facilitate their communication with their authorities.</p> <p>EB. On the recommendation to incorporate the views of all countries during the preparation of major policy papers, directors supported the need for extensive consultation with the membership, but there was limited support for setting out all views in the papers.</p>

EVALUATION	IEO RECOMMENDATIONS PARTIALLY ACCEPTED BY THE BOARD
GFC	<p>Rec. Create a risk-assessment unit that reports directly to management with the purpose of developing risk scenarios for the systemically important countries and analyzing tail risks for the global economy. This unit should organize periodic Board seminars on the risk scenarios and provide an assessment on whether its analysis was appropriately incorporated into multilateral and bilateral surveillance.</p> <p>EB. MIP indicated that the Board did not agree with the specific recommendation to create a risk-assessment unit.</p>
ADV	<p>Rec. Share the major policy issues, the preliminary macroeconomic framework, medium-term projections, and preliminary policy lines with the authorities well before the mission to help facilitate more in-depth discussions.</p> <p>EB. Directors supported early informal consultations with country authorities on key areas of interest and stressed that their offices could play a key role in this process and, more broadly, in promoting dialogue between mission teams and country authorities. Directors, nonetheless, had different views on whether sharing the macroeconomic framework and key policy recommendations with the authorities ahead of missions would add value to Article IV consultations.</p>
CRISIS	<p>Rec. Management needs to consolidate and simplify the current framework to identify and assess risks and vulnerabilities. In particular, the Early Warning Exercise (EWE) needs to be made more user-friendly; it should foster greater debate and input by participants, and outreach on its results should aim to reach authorities.</p> <p>MD. Qualified support. It is not proposed to enhance the outreach of the EWE at this time.</p> <p>EB. Directors expressed a range of views on the appropriateness of disseminating the EWE findings to a wider audience, including by debriefing the Board. There was agreement that any such effort should not compromise candor or access to confidential information.</p>
CRISIS	<p>Rec. Financial System Stability Assessments (FSSAs) for the five to seven largest systemic financial centers should be updated annually in conjunction with IMF's bilateral surveillance.</p> <p>EB. Most directors saw limited merit in the IEO recommendation and instead look forward to forthcoming proposals to mainstream macro financial surveillance.</p>

EVALUATION	IEO RECOMMENDATIONS PARTIALLY ACCEPTED BY THE BOARD
SELF	<p>Rec. Adopt a broad policy or general principles to establish an explicit, institution-wide framework for self-evaluation in the IMF (including its goals, scope, outputs, utilization, and follow up). The IMF should then conduct periodic review of this function as a basis to adapt the policy to changing circumstances.</p> <p>MD. Qualified support. I see limited value in adopting a broad policy to establish an explicit, institution-wide framework for self-evaluation in the IMF that is subject to review every few years.</p> <p>EB. [EDs] took note of the report’s finding that the IMF does not have an institution-wide framework for self-evaluation, instead using a variety of tools and mechanisms that contain an explicit or implicit self-evaluation element. Directors were reassured by the report’s finding that this flexible approach for the most part has served the IMF relatively well. Therefore, while a number of directors saw merit in establishing a new, explicit, institution-wide framework for self-evaluation, many directors considered it more useful to build on existing processes to deliver the necessary strategic approach. Directors agreed that self-evaluation must evolve with the policy and operational environment and that a strategic approach will avoid introducing excessive rigidity as to when and how the IMF conducts self-evaluation. Directors called for integrating efforts to strengthen self-evaluation into the IMF’s institution-wide strategic planning framework.</p>
SELF	<p>Rec. Conduct self-assessments for every IMF-supported program. The scope and format to these assessments could vary across programs, but it is critical that country authorities be given the opportunity to express their views on design and results of each program, as well as on IMF performance.</p> <p>MD. Rejected</p> <p>EB. While some directors preferred self-assessing every IMF-supported program, most noted the likely limited value this would generate relative to costs, in light of the significant amount of self-assessment of programs already existing. These directors favored a more selective, risk-based approach. Directors underscored the importance of better integrating country authorities’ views, whether utilizing existing mechanisms or other new approaches.</p>

EVALUATION	IEO RECOMMENDATIONS PARTIALLY ACCEPTED BY THE BOARD
SELF	<p>Rec. Each policy and thematic review should explicitly set out a plan for how the policies and operations it covers will be self-evaluated going forward.</p> <p>MD. Setting out a plan ex ante for how the self-assessment of every policy and thematic review should be conducted risks turning self-assessment into a routine, box-ticking exercise divorced from new and rapidly evolving challenges that the institution should adapt to instead.</p> <p>EB. Directors broadly agreed that at the outset, policy and thematic reviews should define the objectives of the review and what would constitute policy success. However, they did not support spelling out ex ante how the self-assessment of such reviews should be conducted, noting that this would allow plans for such reviews to be integrated within the IMF’s overall planning framework and adapted to take account of the changing needs facing the institution.</p>
DATA	<p>Rec. Reconsider the role and mandate of the Statistics Department (STA). The work of STA could be refocused towards what is needed to support the IMF’s core operations, making the provision of services to the IMF the nucleus of the department’s activity. This would entail a change in the department’s culture and organization—including increased attention to the timeliness and operational relevance of the data it manages, reallocation of resources towards activities that more directly support the IMF’s main mandate, and inclusion of more staff with IMF operational experience.</p> <p>EB. Directors generally considered that a decision on whether to move the new data management structure into STA should be taken in the context of the long-term strategy.</p>
FCS	<p>Rec. The IMF should take steps to incentivize high-quality and experienced staff to work on individual fragile states, ensure that adequate resources are allocated to support their work, and find pragmatic ways of increasing field presence in high-risk locations while taking necessary security arrangements, even at high cost.</p> <p>EB. Directors noted that an increase in field staff presence in high-risk locations should be weighed against the paramount objective of protecting staff safety.</p>
FIN	<p>Rec. The IMF should revisit the current approach to allocating Financial Sector Assessment Program resources to achieve a more flexible, dynamic, and risk-based allocation across countries and issues.</p> <p>EB. Qualified support. A number of directors were opposed to limiting mandatory assessments to five jurisdictions with the most systemically important financial sectors (S5).</p>

EVALUATION	IEO RECOMMENDATIONS PARTIALLY ACCEPTED BY THE BOARD
FIN	<p>Rec. The IMF should continue to work to enhance the impact of IMF multilateral surveillance by increasing rigor and transparency, and by deepening collaboration with international partners.</p> <p>EB. Qualified support. Some directors supported wider dissemination of the Early Warning Exercise to senior officials, while others cautioned that wider dissemination could weaken its effectiveness.</p>
CFM	<p>Rec. Revisit the Institutional View (IV) in the light of recent experience and recent research. In particular, the following changes to the IV should be carefully considered: allow for preemptive and more long-lasting use of capital flow measures (CFM) in some circumstances; consider distributional implications as part of the strategy for capital account liberalization within the IV; and rethink the concept of the CFM.</p> <p>EB. The Board supported preemptive and more lasting CFMs but expressed diverse views on outflow CFMs outside of crisis or near- crisis circumstances and reconsideration of CFM Macprudential Measure (MPM) classifications.</p>

ANNEX 4.3

METRICS FOR PROXIMITY TO HIGHER-LEVEL RESULTS AND INSTITUTIONAL CHANGE

PROXIMITY TO HIGHER-LEVEL RESULTS RATINGS

Unit of analysis. Recommendations and Management Implementation Plan (MIP) actions.

Reference metric. The (implicit) results chain or the theory of change (ToC) for the International Monetary Fund (IMF) intervention evaluated by the Independent Evaluation Office (IEO). Ideally, such ratings would be assigned based on an explicit ToC for the IMF intervention prepared by the IEO in consultation with the stakeholders.

Rating Scale

4 Impact level. If fully implemented, the recommendation or the MIP action will achieve a significant improvement in sustainable economic growth and macroeconomic and financial stability for the member(s), the global economy, or both; that is, the highest-level results envisaged in the IMF's mandate.

3 Outcome level. If fully implemented, the recommendation or the MIP action will achieve an outcome that may be necessary, but not sufficient, to secure the results at the impact level. In most cases, attribution to the IEO recommendations and the MIP actions will be feasible at most at this level. Examples include the outcomes of changes in IMF policies and practices envisaged in the IEO recommendations, such as “creating an environment that encourages candor and diverse/dissenting views,” or “better coordinating the work on fragile and conflict-affected states by the Fund and other stakeholders,” or “enhancing the technical quality of analytical work.”

2 Output level. If fully implemented, the recommendation or the MIP action will achieve an output that may be necessary, but not sufficient, to achieve the results at the outcome level. Examples include the issuance of a new or revised guidance note on a policy issue aligned with the substance of the IEO recommendation, or the delivery of a working group report advising the management on a specific policy.

1 Input level. If fully implemented, the recommendation or MIP action will secure an input that may be necessary, but not sufficient, to achieve the results at the

output level. Examples include setting up a working group or a task force to advise management on a specific policy issue, or a change in HR practices to hire a priority group of specialists the IEO deemed necessary to achieve the IMF's mandate.

INSTITUTIONAL CHANGE—DEPTH RATINGS

Unit of analysis. Recommendations and MIP actions.

Reference metric. Significance of the institutional change to improve the effectiveness of the IMF in achieving its mandate.

Rating Scale

3 High depth. The recommendations or MIP actions that, by themselves, would bring about durable and highly significant institutional changes. In principle, adoption of any major or transformational revisions to the IMF's existing policy framework or strategic direction will be included in this category. Examples include the adoption of "The Liberalization and Management of Capital Flows: An Institutional View," and the incorporation in Article IV consultations of emerging issues, such as climate change mitigation, inequality, and digital technology. Major innovations by the IMF to introduce transformative analytical tools to support global economic and financial stability, such as the Financial Sector Assessment Program, or a major shift in the incentives towards achieving the intended high-level results could also be included in this category. Such shifts in incentives could be within the institution or with respect to effective collaboration with relevant partners based on the respective comparative advantages.

2 Moderate depth. The recommendations or MIP actions will achieve institutional changes of moderate significance. In some cases, they could include substantial actions necessary, but not sufficient, to deliver durable and highly significant institutional changes, such as a substantial shift in resource allocation to better support the IMF's mandate (for example, a change in HR practices to hire a priority group of specialists deemed necessary to deliver the mandate). In other cases, they could include substantial, but not major, revisions to the existing policies that would bring about more gradual improvements in policies. This category could also include the introduction of new analytical tools, or upgrades to the existing analytical tools, which could significantly improve the quality of policy advice.

1 Low or no depth. The recommendations or MIP actions will achieve institutional changes of low significance or no change, usually as stepping stones for more significant changes. Examples would include the delivery of a working group report advising management on a specific policy issue or setting up a working group or a task force to advise management on a specific policy.

ANNEX 4.4

AN EMPIRICAL ANALYSIS OF TIME TO IMPLEMENT THE MIP ACTIONS

A PRIMER ON SURVIVAL ANALYSIS

A class of models known as survival analysis, or duration analysis in social sciences, analyzes the time to occurrence of an event, that is, the duration $T \geq 0$ (for an overview, see Wooldridge 2010, and Cleves, Gould, and Marchenko 2016). Linear regression is not suitable to estimate such models as time to event does not follow a normal distribution. Parametric estimation of survival models substitutes the normality assumption with a more reasonable distributional assumption for T . Semi-parametric and non-parametric models make no assumptions about the distribution of T . Cox regression is a multivariate semi-parametric method that fits a conditional logistic model to predict the occurrence of successive events conditional on a set of explanatory variables, that is, covariates, and Kaplan-Meier is a univariate non-parametric method frequently used in survival analysis.

In the empirical model estimated in this section, T is a continuous random variable representing the length of time it takes to implement a MIP action. For a particular value t of T , the cumulative distribution function (cdf) of T is the probability of a MIP action being completed by time t .

$$F(t) = P(T \leq t) \quad t \geq 0$$

The hazard function plays a central role in survival analysis. It allows approximating the probability of exiting the initial state within a short interval and defined as:

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{P(t \leq T < t + h | T \geq t)}{h}$$

For each t , $\lambda(t)$ is the instantaneous rate of completion of MIP actions per unit of time. For a differentiable cdf,

$$\lambda(t) = \frac{f(t)}{S(t)}$$

where $S(t) = 1 - F(t)$, also called the survivor function, is the probability of a MIP action not being completed by t . The density of T is then derived as a function of $\lambda(t)$

$$f(t) = \lambda(t) \exp \left[- \int_0^t \lambda(s) ds \right]$$

The shape of the hazard function is of primary interest in empirical applications and depends on the process driving T and the event examined. The simplest case is the constant hazard function, derived under the exponential distribution, which means the process driving T is memoryless: the probability of exit in the next interval does not depend on how much time has been spent in the initial state. In this application, a constant hazard function does not make much sense as the time for completion would depend on how long an action has remained open. Instead, a Weibull distribution is a relatively simple way to capture duration dependence as it accommodates a wide range of shapes for the hazard function over time with a density function given by:

$$f(t) = \gamma \alpha t^{\alpha-1} \exp(-\gamma t^\alpha)$$

Its hazard function is then derived as

$$\lambda(t) = f(t) / S(t) = \gamma \alpha t^{\alpha-1}$$

When $\alpha=1$, the Weibull distribution reduces to the exponential distribution if $\lambda=\gamma$. If $\alpha>1$, the hazard is monotonically increasing, so the hazard everywhere exhibits positive duration dependence; if $\alpha<1$, the hazard is monotonically decreasing. While the shape of the hazard function is less restricted compared to the exponential distribution, monotonicity may not be a realistic assumption depending on the specific empirical case. Both the exponential and the Weibull distributions belong to a class of models known as proportional hazard models. A proportional hazard can be written as

$$\lambda(t|x) = K(x)\lambda_0(t)$$

where x is a vector of explanatory variables, $K(x)>0$ is a positive function of x and $\lambda_0(t)>0$ is called the baseline hazard. The baseline hazard is common to all units in the population; individual hazard functions differ proportionately based on a function $K(x)$ of observed covariates. Typically, $K(x)=\exp(x\beta)$, where β is a vector of parameters. Then

$$\log \lambda(t|x) = x\beta + \log \lambda_0(t)$$

Where β_j measures the semielasticity of the hazard with respect to x_j . Cox (1972) suggested a partial likelihood method for estimating the parameters in a proportional hazard model without specifying the baseline hazard. The strength of the

Cox model is that the model makes no assumptions about the shape of the hazard function. Among the parametric models, lognormal, Gompertz, and loglogistic distributions relax the proportional hazard assumption. Which distribution fits better is chosen according to the model selection criteria of Akaike's information criterion (AIC) and Bayesian information criterion (BIC).

EMPIRICAL MODEL

Time and event variables. The survival analysis dataset requires a time and an event variable. The time variable, T is the length of time it takes in years to implement a MIP action. It is calculated from the approval of the MIP to the date of the PMR staff report that reports the action as completed. The event variable, that is, the implementation of the MIP action, takes the value of 1 if completed, or 0 if still open at the end of the observation period. Survival analysis makes use of both completed and open actions. The data used in the estimation is as of the date of the Eleventh PMR.

Explanatory variables include the two metrics on proximity to higher-level results and the depth of institutional change, as well as dummy variables for each evaluation. For both metrics, each category enters the regression as a 0 to 1 indicator variable. The output level actions with institutional change at low depth are chosen as the omitted base category for the indicator variables. Which categories are chosen to omit makes no difference but affects how results are reported. Another indicator variable is included for actions that involve long-term technical or culture change (that is, actions classified under category 4 in the triage framework). While remaining open for 6.5 years or longer, three input-level actions involving structural shifts in the IMF's human resource structure and practices have not yet been completed.¹ A dummy variable is constructed to capture the impact of these actions on results.

Treatment of retired and reformulated actions. Reformulated actions except for those under category 3 (that is, six actions reformulated to better align with SMART principles) and all but four of the retired actions were included in the estimation sample and recorded as still open as of the date of the categorization exercise. As

¹ These actions in GFC and ADV MIPs included the following: efforts would continue to broaden the professional diversity of staff, including their educational background and skill mix, with more staff being hired with financial sector experience; target three years average tenure for country assignments, to be monitored by the Human Resources Department (HRD) and reported to the Board on a regular basis and; implement measures to raise the share of nationals from underrepresented regions among senior staff and set targets for representation of women at senior levels. As noted earlier, in 2020, the Board extended annual monitoring for five additional PMR cycles for the latter and in 2021, reformulated the action on average tenure.

category 3 actions were too broadly worded with no clear measures of success and not time-bound, they remained open until the categorization exercise. As such, their inclusion might introduce undue noise in estimation.

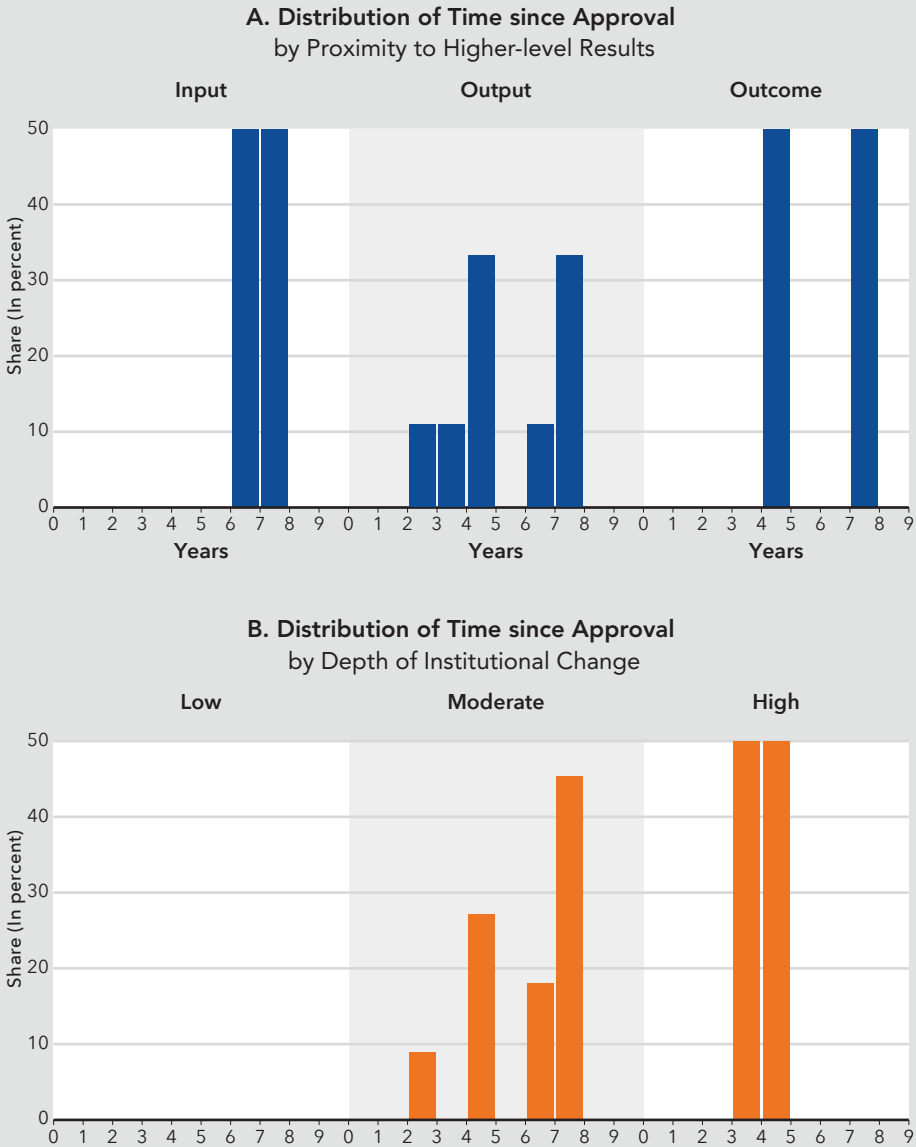
Most of the retired actions in category 5 were superseded by other workstreams and/or the MIP actions addressing recommendations in more recent IEO evaluations and marked as open. For instance, five actions, four arising from GFC and one from CRISIS, were superseded by the UMP actions to improve the IMF's financial spillover analyses and strengthen financial and macro-financial analysis in surveillance, and by the FIN actions to further develop financial expertise at the IMF. A DATA action was superseded by the FCS actions related to capacity development in fragile states, with one specifically on data-related issues. A GIP action on improving the analytical underpinnings of both surveillance and program design, especially in the areas of economic forecast, external sector assessments, and integrated surveillance, was also recorded as open as it is being addressed through various IMF workstreams. Similarly, a GFC action on reporting on the traction of past advice and an ADV action on enhancing the use of cross-country examples in Article IV reports were classified as open and they continue to be monitored and addressed through other IMF workstreams.² Two ADV actions and one action in the evaluation of *IMF Forecasts: Process, Quality, and Country Perspectives* (FOR) were marked as completed at the time of the categorization paper.³ On the other hand, two FOR actions were retired without being implemented as they were replaced by new priorities owing to the IMF's budget constraints.⁴ These two actions were

² IMF (2019) noted that while reporting of traction in staff reports improved sharply, increasing from 30 percent in 2011 to 80 percent in 2018, this GFC action would continue to be monitored in the Comprehensive Surveillance Review (CSR) based on a more comprehensive concept of traction. For the ADV action, sharing cross-country knowledge was identified as a priority by the Knowledge Management Unit (KMU) and is being addressed by the CSR and the KMU.

³ Two completed ADV actions were: "Early consultations with country authorities will now be expected of all teams" and "Mission teams will be expected to share major policy questions and global assumptions underlying projections at least one week ahead of the mission." These actions are marked as completed as the 2019 Office of Executive Directors (OED) survey results confirmed that early and substantive interactions were the norm. IMF (2019) noted that OEDs particularly appreciated the staff's candid and open exchange of views, suggesting important progress by mission chiefs in explaining the rationale of the Fund's policy advice. The completed FOR action was: "Policy consultation meetings for Article IV consultations and program reviews will discuss the underlying causes of significant forecast errors (not pertaining to data revisions) for key macroeconomic variables." IMF (2019) noted that major work had been completed to enhance learning from forecast errors.

⁴ Two FOR actions retired without being implemented were: first, start a pilot process (to be broadened gradually for most major economies and country groups) with top-down guidance for medium-term forecasts in the spirit of the approach used for short-term forecasts; and second, develop a multicountry macro econometric forecasting model to ensure medium-term consistency of forecasts.

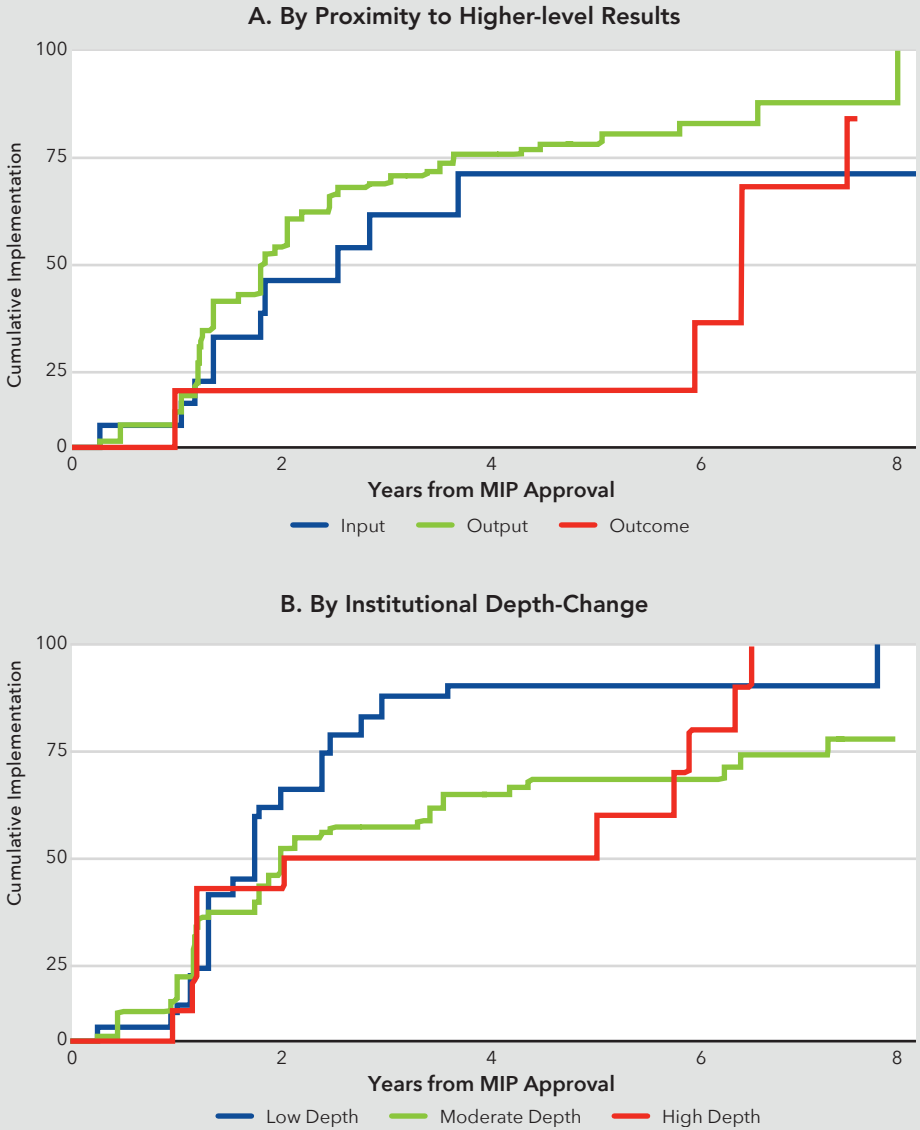
FIGURE 4A.1. DISTRIBUTION OF TIME SINCE APPROVAL OF RETIRED OR REFORMULATED MIP ACTIONS MARKED AS OPEN



Sources: PMRs, IMF (2019), and the author’s calculations.

Note: Includes only the actions that are classified as open, thereby, excludes six actions under category 3 (reformulated to better align with SMART principles) as well as one duplicate and three completed actions.

FIGURE 4A.2. KAPLAN-MEIER CUMULATIVE PROBABILITY ESTIMATES FOR THE IMPLEMENTATION OF MIP ACTIONS



Source: The author's calculations.

marked as open as of the date that they were retired (that is, recorded as right-censored data). Finally, a retired GFC action was dropped as it was essentially a duplicate of another retired GFC action, as recognized in IMF (2019).

Figure 4A.1 presents the distribution of time since approval of the retired or reformulated MIP actions that are classified as open. Except for one DATA and one GIP action that were open for less than three years, all others remained open four to eight years prior to being formulated. These actions are still being addressed by other IMF workstreams or the MIP actions addressing recommendations in more recent IEO evaluations, or both.

RESULTS

The non-parametric Kaplan-Meier estimates for the cumulative probability of implementation by two metrics are presented in Figure 4A.2. This univariate analysis reveals differences in the cumulative probabilities across categories. The log-rank test rejects the equality of cumulative probabilities of implementation across input-output-outcome categories and low-moderate-high depth categories at 10 percent and 5 percent significance, respectively. A close overlap of functions at the input and the output levels is noteworthy and will be picked up later in the parametric estimation results. Nevertheless, the distribution at the outcome level looks quite distinct from those at the input and output levels. For all three categories of institutional depth, distributions are discernably different.

According to AIC and BIC model selection criteria, the parametric lognormal distribution fits best to the data generating process (Table 4A.1). The performance of loglogistic model is equally good as both distributions are suited well to events whose probability of occurrence increases initially and decreases later. Both the superior performance of the lognormal model compared to the exponential and Weibull distributions and the univariate estimates of the smoothed hazard function by categories suggest that the proportional hazard and monotonicity assumptions do not hold.

TABLE 4A.1. MODEL SELECTION: AKAIKE'S INFORMATION CRITERION (AIC) AND BAYESIAN INFORMATION CRITERION (BIC)

DISTRIBUTIONS	AIC	BIC
Exponential	361.7	419.2
Weibull	311.3	371.9
Gompertz	343.3	403.9
Lognormal	290.0	350.6
Loglogistic	290.1	350.7

Source: The author's calculations.

The results from lognormal regressions are presented in Table 4A.2. The base category in regressions is represented by low-depth output-level actions (that is, the omitted categories). Based on the benchmark regression (Table 4A.2, column 1), at the input and outcome levels, the MIP actions take about 10 percent and 215 percent longer to implement respectively than the base category, though only the latter is significant. In other words, the time to implement at the input and the output levels is not significantly different. This result could be explained by some MIP actions at the input level that required a substantial reallocation of budget resources and took very long to implement. Actions that involve institutional change at moderate and high depth take significantly longer to implement compared to the base category, by about 60 and 72 percent, respectively. Notably, actions that involve long-term technical or culture change take 319 percent longer to implement than the base category. Two outliers turn out to be highly significant and influential in regression results (Table 4A.2, columns (2)-(3), column (4) excludes these observations).⁵ Finally, the dummy variable for four input-level actions that involved structural shifts in the IMF's human resource (HR) structure and practices is highly significant, which flags the prolonged challenges in making progress in this area.

The results for evaluation dummies highlight significant differences in the completion of MIP actions across evaluations. At one end of the spectrum, the actions in response to CFM are estimated to be completed significantly faster than other evaluations; it appears that DATA actions take the longest to implement. GIP

⁵ The first outlier is a low-depth output-level RSR action that took very long to implement. The second outlier is a high-depth, outcome-level CRISIS action that was completed very quickly owing to fortunate timing of long awaited 2010 quota reforms that became effective in 2016.

TABLE 4A.2. LOGNORMAL REGRESSION RESULTS FOR TIME TO IMPLEMENT THE MIP ACTIONS

EXPLANATORY VARIABLES	(1) BENCHMARK	(2)	(3)	(4)
Proximity to Higher-level Results				
Input	1.103	1.217	1.303	1.104
	(0.139)	(0.211)	(0.222)	(0.140)
Outcome	3.149***	2.259**	3.029***	3.153***
	(0.715)	(0.717)	(0.680)	(0.717)
Institutional Change-Depth				
Moderate Depth	1.601***	1.591***	1.671***	1.605***
	(0.194)	(0.205)	(0.210)	(0.195)
High Depth	1.716*	1.471	1.742*	1.718*
	(0.531)	(0.465)	(0.539)	(0.532)
Culture Change	4.191***	4.523***	4.499***	4.199***
	(0.910)	(1.209)	(1.147)	(0.915)
Structural HR Actions	47.48***			48.66***
	(12.51)			(12.87)
Outlier 1	3.341***		3.672***	
	(0.506)		(0.572)	
Outlier 2	0.119***		0.124***	
	(0.0396)		(0.0409)	
CFM	0.412***	0.347***	0.371***	0.414***
	(0.121)	(0.109)	(0.115)	(0.121)
CRISIS	1.493	1.072	1.377	1.496
	(0.422)	(0.344)	(0.404)	(0.423)
DATA	4.317***	3.761***	4.041***	4.343***
	(1.137)	(1.075)	(1.102)	(1.148)

EXPLANATORY VARIABLES	(1) BENCHMARK	(2)	(3)	(4)
FCS	1.838** (0.485)	1.545 (0.442)	1.642* (0.459)	1.844** (0.488)
FIN	1.853*** (0.392)	1.541* (0.367)	1.674** (0.382)	1.860*** (0.395)
FOR	1.338 (0.393)	1.175 (0.370)	1.211 (0.370)	1.339 (0.395)
GFC	1.864*** (0.421)	1.720** (0.414)	1.713** (0.399)	1.866*** (0.422)
GIP	0.861 (0.268)	0.783 (0.254)	0.785 (0.252)	0.862 (0.269)
IRES	1.272 (0.370)	1.064 (0.327)	1.152 (0.347)	1.273 (0.371)
RSR	2.305*** (0.560)	2.349** (0.811)	1.964** (0.520)	2.308*** (0.562)
SELF	0.626 (0.227)	0.533* (0.203)	0.563 (0.210)	0.626 (0.227)
SP	1.787** (0.440)	1.491 (0.408)	1.583* (0.419)	1.790** (0.442)
UMP	1.354 (0.283)	1.114 (0.264)	1.223 (0.277)	1.358 (0.285)
Constant	1.045 (0.212)	1.257 (0.282)	1.116 (0.235)	1.043 (0.212)
No. of Observations	153	153	153	151

Source: The author's calculations.

Note: Dependent variable: Time since the MIP approval at which the MIP action is observed either as completed or right-censored. Coefficients are reported in multiplicative form for lognormal distribution, t where . Significant at 10 percent:*; 5 percent:**; and 1 percent:***, robust standard errors are in parentheses.

and SELF actions take a relatively shorter time to implement and actions related to FCS, FIN, GFC, the evaluation of *Research at the IMF: Relevance and Utilization* (RSR), and SP relatively longer to implement.

SENSITIVITY ANALYSIS

Based on the benchmark regression, the estimated average time to complete the MIP actions for preset values of covariates are presented in Table 4A.3. This analysis is essentially counterfactual in nature as some combinations of covariates may be unlikely (such as input level low-depth actions also involving culture change). When culture change is set to zero, the predictions range on average from 1.2 to 6.6 years for MIP actions with input/output level actions at the low end and outcome level actions at the high end of the spectrum. For actions that involve culture change, the predicted average time to implement increases to 5.1 to 9.6 years for input/output level actions while outcome level actions take significantly longer, at 16.0 to 27.5 years.

TABLE 4A.3. SENSITIVITY ANALYSIS: PREDICTED AVERAGE TIME TO IMPLEMENT MIP ACTIONS

(In Years)

Proximity to Higher-level Results	Institutional Change-Depth	CULTURE CHANGE (CC)			
		CC=0	95% Confidence Interval	CC=1	95% Confidence Interval
Input	Low	1.3	[0.7-2.0]	5.6	[3.2-8.0]
Input	Moderate	2.1	[1.1-3.1]	9.0	[5.5-12.5]
Input	High	2.3	[0.7-3.9]	9.6	[5.2-14.0]
Output	Low	1.2	[0.7-1.7]	5.1	[3.2-7.0]
Output	Moderate	1.9	[1.2-2.7]	8.1	[5.7-10.6]
Output	High	2.1	[0.7-3.5]	8.7	[5.4-12.1]
Outcome	Low	3.8	[1.9-5.7]	16.0	[5.2-26.8]
Outcome	Moderate	6.1	[3.4-8.8]	25.7	[10.2-41.1]
Outcome	High	6.6	[3.4-9.7]	27.5	[15.9-39.1]

Source: The author's calculations.

Note: Based on the benchmark lognormal regression in Table 4A.2. Dummies for evaluations, structural HR actions, and two outliers are set to zero.

REFERENCES

Independent Evaluation Office of the International Monetary Fund (IEO), 2011, *IMF Performance in the Run-Up to the Financial and Economic Crisis: IMF Surveillance in 2004–07* (Washington: IMF).

_____, 2011, *Research at the IMF: Relevance and Utilization* (Washington: IMF).

_____, 2012, *International Reserves: IMF Concerns and Country Perspectives* (Washington: IMF).

_____, 2013, *The Role of the IMF as Trusted Advisor* (Washington: IMF).

_____, 2014a, *IMF Response to the Financial and Economic Crisis* (Washington: IMF).

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

_____, 2014c, *Recurring Issues from a Decade of Evaluation: Lessons for the IMF* (Washington: IMF).

_____, 2015, *Self-Evaluation at the IMF: An IEO Assessment* (Washington: IMF).

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

_____, 2016b, *Behind the Scenes with Data at the IMF: An IEO Evaluation* (Washington: IMF).

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

_____, 2018, *The IMF and Fragile States* (Washington: IMF).

_____, 2019a, *IMF Financial Surveillance* (Washington: IMF).

_____, 2019b, *IMF Advice on Unconventional Monetary Policies* (Washington: IMF).

_____, 2020a, *IMF Advice on Capital Flows* (Washington: IMF).

_____, 2020b, *IMF Collaboration with the World Bank on Macro-Structural Issues* (Washington: IMF).

_____, 2021, *Growth and Adjustment in IMF-Supported Programs* (Washington: IMF).

_____, 2022, *IMF Engagement with Small Developing States* (Washington: IMF).

International Monetary Fund, 2019, “A Framework to Address Open Management Actions in Response to Board-endorsed IEO Recommendations,” March (Washington).

_____, 2020, “Categorization of Open Actions in Management Implementation Plans,” March (Washington).

_____, 2021, “Eleventh Periodic Monitoring Report on the Status of Management Implementation Plans in Response to Board-endorsed IEO Recommendations,” November (Washington).

_____, 2022a, “Twelfth Periodic Monitoring Report on the Status of Management Implementation Plans in Response to Board-endorsed IEO Recommendations,” November (Washington).

_____, 2022b, Implementation Plan in Response to the Board-endorsed Recommendations Categorization of Open Actions in Management Implementation Plans,” March (Washington).

_____, 2022c, “Review of The Institutional View on The Liberalization and Management of Capital Flows,” March (Washington).

Kaberuka, Donald, Chia Der Jiun, and Pernilla Meyersson, 2018, “Time for a reboot at a critical time for multilateralism: The Third External Evaluation of the IEO (Washington: IMF). <https://ieo.IMF.org/en/our-mandate/External-Evaluations>.

Korinek, Anton, Prakash Loungani, and Jonathan Ostry, 2022, “The IMF’s updated view on capital controls: Welcome fixes but major rethinking is still needed.” <https://www.brookings.edu/blog/up-front/2022/04/18/the-IMFs-updated-view-on-capital-controls-welcome-fixes-but-major-rethinking-is-still-needed/>

Lissakers, Karin, Ishrat Husain, and Ngaire Woods, 2006, “Report of the External Evaluation of the Independent Evaluation Office” (Washington: IMF). <https://ieo.IMF.org/en/our-mandate/External-Evaluations>.

Mayne, John, 2007, “Challenges and Lessons in Implementing Results-Based Management,” *Evaluation*, 13(1) pp. 87–109.

Ocampo, J.A., Stephen Pickford, and Cyrus Rustomjee, 2013, “External Evaluation of the Independent Evaluation Office: Report of the Panel Convened by the Executive Board” (Washington: IMF). <https://ieo.IMF.org/en/our-mandate/External-Evaluations>.