



INTERNATIONAL MONETARY FUND

EXTERNAL INDEPENDENT EVALUATION OF THE JAPAN
TECHNICAL ASSISTANCE SUBACCOUNT

FINAL REPORT - VOLUME I (MAIN REPORT)

MARCH 2024

CONTENTS

EXECUTIVE SUMMARY.....	II
1. INTRODUCTION	1
1.1 Background to IMF CD and the JSA	1
1.2 About the current evaluation.....	3
1.3 Structure of this report.....	5
2. HOW WELL DID JSA-FUNDED ACTIVITIES PERFORM?.....	6
2.1 Relevance	6
2.2 Coherence.....	8
2.3 Effectiveness.....	10
2.4 Efficiency.....	14
2.5 Impact	16
2.6 Sustainability	17
2.7 Lessons learned.....	20
3. CONCLUSIONS AND RECOMMENDATIONS	21
3.1 Conclusions.....	21
3.2 Recommendations.....	24

GLOSSARY OF ABBREVIATIONS

AFR	African Department
APD	Asia Pacific Department
CD	Capacity Development
CDD	Capacity Development Department
CDOT	IMF Capacity Development Office in Thailand
CEF	Common Evaluation Framework
DAC	Development Assistance Committee
ECF	Extended Credit Facility
EFF	Extended Fund Facility
EME	Emerging Market Economy
FAD	Fiscal Affairs Department
GFS	Government Finance Statistics
ICD	Institute for Capacity Development Department
IGF	Infrastructure Governance Facility
IMF	International Monetary Fund
JMOF	Ministry of Finance, Japan
JSA	Japan Sub-Account for Selected Fund Activities
LEG	Legal Department
LIC	Low-Income Country
LTX	Long-Term Expert
MC	Mission Chief
MCM	Monetary and Capital Markets Department
MOOC	Massive Open Online Course
OECD	Organisation for Economic Cooperation and Development
RBM	Results-Based Management
RCDC	Regional Capacity Development Centre
RR	Resident Representative
SOE	State-Owned Enterprise
STA	Statistics Department
STI	Singapore Regional Training Institute
STX	Short-Term Expert
TA	Technical Assistance
VFM	Value for Money
WB	World Bank

EXECUTIVE SUMMARY

BACKGROUND

In January 2023, the International Monetary Fund (IMF) appointed BDO LLP to carry out an independent mid-term evaluation of capacity development (CD) activities funded by the Government of Japan through the Japan Sub-Account for Selected Fund Activities (JSA). The evaluation was requested and sponsored by the Ministry of Finance, Japan (JMOF) in its capacity as agency responsible for the JSA, and covered 32 JSA-funded programs delivered between 1 May 2017 and 30 April 2022. It examined performance against the six Organisation for Economic Cooperation and Development - Development Assistance Committee (OECD-DAC) evaluation criteria of relevance, coherence, effectiveness, efficiency, impact and sustainability.

The evaluation fieldwork was carried out between June and October 2023, with the key methodologies being document review, semi-structured stakeholder interviews, e-surveys of recipient authorities and IMF staff, and detailed analysis and assessment of a sample of 15 JSA-funded programs. In-person field visits were carried out to two countries where multiple JSA-funded programs had been taking place, namely Cambodia and Nigeria.

FINDINGS AND CONCLUSIONS

Our key findings and conclusions from the evaluation are set out below.

- JSA-funded programs constitute a key part of overall IMF CD, and contribute greatly to developing capacity and strengthening institutions in Asia and some African countries, particularly fragile states. JSA-funded training is effective in building knowledge and policy-oriented skills across the IMF membership on issues critical to macroeconomic and financial stability, fiscal and monetary frameworks, and emerging global challenges.
- The period covered by this evaluation was dominated by the pandemic, and the Fund and JSA deserve credit for adapting to the circumstances. The pandemic restrictions halted all CD activities overnight, and they resumed only gradually. The Fund was able to respond by replacing in-person activities with virtual delivery of technical assistance (TA) and training, and an expansion of online learning programs. These new modalities contributed to integration of TA and training with surveillance - one of the Fund's strategic goals.
- JSA-funded programs score highly for relevance and coherence, i.e., the extent to which they address identified needs, and complement other activities being carried out in target countries by development partners or the recipient authority. JSA-funded programs were also well aligned with priorities identified in IMF surveillance and lending programs during the evaluation period.
- Most JSA-funded programs have been relatively effective, in that they largely or at least partially achieved their milestones, outcomes, and objectives - despite significant disruption from the COVID pandemic. There were some significant differences in effectiveness across sectors - although caution must be exercised in using the Results-Based Management (RBM) framework for such comparisons. The evaluation also found significant differences in the achievements of programs across types of countries - e.g., fragile versus non-fragile countries, emerging market economies (EMEs) versus low-income countries (LICs). However, these differences do not translate into widely different effectiveness ratings, since projects in fragile and LIC countries generally have less ambitious objectives than those in less fragile or higher-income settings. This means the likelihood of projects achieving their objectives is broadly similar between the two groups.

In addition to issues with comparability between the level of expected achievement in different countries or settings, the RBM database remains limited as a tool for assessing each program's progress against milestones, outputs, outcomes and objectives. There has been significant progress during the evaluation period in entering the required information and in rating compliance. But there are still gaps and sometimes confusion between inputs, outputs, and outcomes.

- Multi-country programs have been a flexible and often efficient way to deliver JSA-funded CD at scale. However, they can also complicate coordination within the Fund, and can pose challenges to transparency and accountability. One advantage of such programs is their flexibility to shift resources between recipient countries in response to changing priorities and circumstances. They can also offer opportunities for peer
-

learning between countries, and possible cost savings through sharing a single Long-Term Expert (LTX) between multiple countries. Possible challenges of the model include identifying and deploying appropriate experts if the recipient countries are very heterogeneous, monitoring and managing resources in the absence of granular information by individual country, and evaluating program performance when some countries have achieved results while others have not.

- Qualitative evidence suggests that the Value for Money (VFM) of JSA-funded programs has improved since the pandemic - as new delivery modalities are less costly and appear equally effective. That said, assessing the VFM of JSA-funded activities is challenging due to the lack of complete cost data, and difficulty in monetizing outcomes.
- In our view, it would not be practical to try to optimize VFM by identifying programs for early cancellation and reallocation of funds - particularly given the already short duration of most JSA-funded interventions. Even at the mid-point of program implementation, it is usually difficult for program managers to make a robust assessment of the likely ultimate achievements and costs - as confirmed by the limited information available in interim assessment reports. That said, managers of multi-country programs sometimes reallocate resources from poorly performing or stalled country projects to countries where the resources could be used more effectively.
- With a few notable exceptions, we were not able to identify wider impacts from JSA-funded programs (i.e., higher-level effects beyond their initial outputs and outcomes). However, this is not because JSA-funded programs lack impact *per se*. Instead, it is due to short timeframes, difficulty in identifying and attributing possible wider effects, and lack of routine data collection beyond the end of the programs.
- For the same reasons, there is little information on JSA programs' sustainability (i.e., the extent to which their benefits are likely to continue later). Hence we based our sustainability assessments largely on information available at or before completion - e.g., whether there was strong recipient authority 'buy-in', whether knowledge, tools and systems had been absorbed by the recipient organization, whether the requisite financial and human resources were in place for reforms to continue, and whether staff who participated in the TA were likely to stay in post and transmit the experience to colleagues. Based on these metrics, the sustainability of JSA programs appears relatively modest, particularly in fragile states.

SUMMARY RATINGS OF SAMPLE PROGRAMS AGAINST OECD-DAC CRITERIA

Our ratings of JSA programs against the OECD-DAC evaluation criteria of relevance, coherence, effectiveness, efficiency, impact and sustainability are summarized below, including sub-analysis by geographical and thematic criteria. The ratings are based on a four-point scale, where 1 is 'poor' and 4 is 'excellent'. The underlying evidence and rationale for the ratings awarded are presented in **Section 2** of our report.

	Relevance	Coherence	Effectiveness	Efficiency ¹	Impact	Sustainability	Overall rating ²	Number of projects ³ in sample
Overall Average	2.9	3.0	2.5	2.9	2.1	2.4	2.6	58
By workstream								
FAD	2.9	3.0	2.3	2.8	2.2	2.2	2.5	31
ICD	3.2	3.3	3.3	3.2	2.1	2.8	3.0	8
MCM	2.8	2.9	2.5	2.7	1.9	2.5	2.5	13
STA	2.9	2.9	2.7	3.1	2.3	3.0	2.8	6

¹ Due to the lack of robust and complete information on project expenditure and progress (see **Section 2.4**), the figures for efficiency are indicative only, and should be treated with caution.

² The overall ratings for each program are a weighted average of the scores under the individual OECD-DAC criteria, using the following weightings: Relevance 15%; Coherence 15%; Effectiveness 30%; Efficiency 10%; Impact 15%; Sustainability 15%.

³ Our sample consisted of 15 JSA-funded programs. Most of these were multi-country programs comprising several country-level projects, which we assessed and scored individually. The figures in this column represent the total number of individual projects assessed within each category / sub-group.

	Relevance	Coherence	Effectiveness	Efficiency ¹	Impact	Sustainability	Overall rating ²	Number of projects ³ in sample
By region								
Asia-Pacific	3.0	3.1	2.4	2.7	2.1	2.6	2.6	19
Africa	2.9	2.9	2.5	2.9	2.2	2.3	2.6	33
Global / Mixed ⁴	3.0	3.0	3.2	3.0	2.2	2.5	2.9	6
By fragility								
Fragile States	2.9	3.0	2.4	3.0	2.2	2.4	2.6	19
Non-fragile States	2.9	3.0	2.5	2.8	2.1	2.5	2.6	33
Global / Mixed ⁵	3.0	3.0	3.2	3.0	2.2	2.5	2.9	6

RECOMMENDATIONS

Our recommendations to the IMF and/or the JSA are set out below. More detail of their rationale and suggested implementation approach is set out in **Section 3**.

1. The Fund should maintain or adopt important innovations and adaptations catalyzed by the pandemic - notably those relating to technology and remote delivery - across all relevant activities. *(Priority: High)*
2. Regarding training, the Fund should seek to maintain the progress made in using online and virtual courses rather than reverting fully to former modes of delivery. *(Priority: High)*
3. The Fund should seek more and earlier inputs from recipient authorities in deciding on a program, its design, and on the selection of an LTX. *(Priority: High)*
4. JSA should consider reallocating some of the funding from Asian emerging market economies to other regions, including African countries, which have greater needs for capacity building. *(Priority: High)*
5. The Fund and JSA should explore ways to structure multi-country programs that maintain their flexibility, while also supporting transparency and accountability, and reducing managerial challenges. *(Priority: Medium)*
6. The Fund should make greater efforts to make recipient authorities and IMF Mission Chiefs / Resident Representatives aware that JSA CD is funded by the Government of Japan. This should include increasing the JSA 'branding' on project documents, and requiring IMF project leads to inform local IMF staff and recipient authority counterparts as to the source of the funding. *(Priority: Medium)*
7. The Fund should consider ways to gather evidence routinely on the *de facto* sustainability and impacts achieved by JSA-funded programs after implementation. *(Priority: Low)*

We would like to thank all those involved in the evaluation for their time and assistance in providing documents and information to support our work.



Caroline Hagmann
Partner
BDO LLP
25 March 2024

55 Baker Street
London
W1U 8EW

⁴ Global /Mixed refers to programs such as online training, which could include participants from more than one region.

⁵ Global /Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

1. INTRODUCTION

1.1 BACKGROUND TO IMF CD AND THE JSA

1.1.1 IMF CAPACITY DEVELOPMENT

The International Monetary Fund (IMF) is a financial agency of the United Nations made up of 190 member countries, that works to achieve sustainable growth and prosperity for all its members. It does so by supporting economic policies that promote financial stability and monetary cooperation, which are essential to increased productivity, job creation, and economic well-being. The IMF aims to promote international monetary cooperation, exchange stability, and orderly exchange arrangements; to provide temporary financial assistance to countries with balance of payments difficulties; and to foster sustainable economic growth. To achieve these objectives, the IMF undertakes three main functions:

- **Surveillance** is the process by which the IMF maintains a policy dialogue with each of its members and appraises country and global macroeconomic conditions.
- **Financial assistance** includes providing credits and loans to member countries with balance of payments problems so that they can restore conditions for financial and macroeconomic stability and sustainable economic growth.
- **Capacity development (CD)** consists of expert advice and training provided to government institutions of member countries to help them strengthen their human and institutional capacity and design and implement effective macroeconomic and structural policies.

The IMF offers CD in its areas of core expertise, including revenue administration, public financial management, macroeconomic statistics, financial supervision and regulation, macroeconomic frameworks, central bank operations, tax policy and financial integrity.⁶

1.1.2 JAPANESE SUPPORT THROUGH THE JSA

Japan is the largest contributor to IMF CD, including technical assistance (TA) and training. Since 1990, Japan-financed IMF CD has helped country authorities build capacity for formulating and implementing sound macroeconomic policies in the fiscal, monetary, financial, and related statistical fields, and legal and administrative capacity.

Contributions to IMF CD are provided through the Japan Subaccount for Selected Fund Activities (JSA). Since 1990, Japan has contributed about \$807 million to JSA. As at April 2023, some \$661 million had been used to support IMF CD. This enabled the IMF to assist more than 100 member countries with formulating and implementing sound macroeconomic policies (in the fiscal, monetary, financial, and statistical fields) as well as strengthening capacity by conducting training in these areas. JSA funds are used to cover program and project costs, including the salaries and travel costs of staff and experts and the costs associated with developing training and organizing seminars and workshops.⁷ JSA-funded CD mainly supports low- and lower-middle-income countries as they implement more growth-enhancing macroeconomic policies. Since April 2010, many of the JSA-funded TA activities are being delivered under a multi-country programmatic approach to provide region-wide and medium-term assistance, encourage synergies, and enhance effectiveness and sustainability.

JSA-funded programs constitute a key part of overall IMF CD, and they are the main component in activities such as online training and support for JSA-priority Southeast Asian countries, namely Cambodia, Lao PDR, Myanmar, and Vietnam. Myanmar was important during the first half of the review period, but IMF CD support ceased in early 2021 following a military coup. The IMF redirected JSA resources initially allocated for Myanmar to other countries, mainly in the APD region - e.g., PNG, and the Philippines. JSA also supports CD for small Pacific Island states and several countries in Sub-Saharan Africa, particularly Fragile States (also a priority for Japan).

⁶ *IMF Policies and Practices on Capacity Development* (<https://www.imf.org/en/Publications/Policy-Papers/Issues/2019/11/14/IMF-Policies-and-Practices-on-Capacity-Development-48811>).

⁷ JSA-funded activities are complemented by Fund-financed staff backstopping and project management, diagnostic missions, and installation and inspection missions.

During the evaluation period (FY2018 to FY2022⁸), JSA contributed around US\$174 million to support IMF CD activities, making it the major donor and covering 30 percent of external financing for IMF CD (or about 20 percent of the Fund’s total expenditure on CD).⁹ JSA has been a major funding source for IMF training, with most of the IMF’s online courses and the IMF Institute Learning channel financed by JSA. Along with the Singaporean government, JSA is also the major supporter of the Singapore Regional Training Institute (STI). In addition, Japan provides support for other important initiatives such as research on Central Bank Digital Currencies, and multi-partner thematic initiatives such as the Anti-Money Laundering/ Combating the Financing of Terrorism Thematic Fund (AML/CFT), Data for Decision Fund (D4D), Tax Administration Diagnostic Assessment Tool (TADAT), and bilateral Infrastructure Governance Facility (an IMF-Japan CD Flagship Program).¹⁰

As shown in **Figure 1** below, around 40% of budgeted JSA expenditure in the evaluation period (1 May 2017 to 30 April 2022) was on programs delivered by the Fiscal Affairs Department (FAD), mainly covering public financial management and revenue administration. The next highest expenditure (27%) was by the Institute of Capacity Development (ICD), followed by the Monetary and Capital Markets (MCM) and Statistics (STA) Departments (22% and 9% respectively). The Legal Department (LEG) accounted for just 1% of the total.

Figure 1: Budgeted JSA expenditure FY18-22, by IMF CD department

DEPARTMENT	BUDGETED EXPENDITURE (US\$)	% OF TOTAL
Fiscal Affairs Department (FAD)	44,663,853	40%
Institute for Capacity Development (ICD)	24,991,795	27%
Monetary and Capital Markets (MCM)	30,373,586	22%
Statistics Department (STA)	9,574,744	9%
Legal Department (LEG)	1,507,170	1%
TOTAL	111,111,148	100%

Source: BDO analysis of IMF datasheet ‘Japan Projects and Execution FY18-FY22’.

Note: Figures are for budgeted JSA expenditure during the evaluation period. They may not sum precisely due to rounding.

1.1.3 PREVIOUS EVALUATIONS OF JSA

Since 2003, there have been four independent external evaluations of JSA-funded CD - the most recent in 2017. The 2017 evaluation found that the quality of JSA outputs was consistently very high, although the evaluators rated the effectiveness and impact of many programs lower than the IMF’s own self-assessments. The evaluation noted that several programs did not have specific defined goals, and that some had confused outputs with outcomes, or had outcomes and objectives that were not well defined.

Regarding delivery modalities, the evaluation found that long-term experts (LTX) were generally preferred over short-term experts (STX) or HQ support in terms of effectiveness. The evaluation found donor coordination to be good in the sense of avoiding duplication and ensuring complementarity, but that “true synergies” were attained only in a few cases. The visibility of Japan as JSA donor was found to be high, with nearly every beneficiary aware that the funding came from Japan.

The evaluation highlighted the importance of institutional and political factors as ultimate determinants of CD success. Recipient countries’ limited institutional capacity was found to be a key factor limiting the sustainability of IMF CD. Moreover, the IMF’s lack of expertise with institutional development (despite its expertise in technical areas) limited its ability to promote institutional change, as opposed to improving competence at the individual level. The evaluation found that IMF experts’ efforts to incorporate “institutional constraints” or “political economy challenges” into project design were *ad hoc* or not systematic. The evaluation therefore recommended:

1. More ‘design thinking’, to better understand the political environment of recipient countries concerned rather than focusing just on transfer of technical knowledge.
2. Operationalizing institutional development, potentially by hiring institutional development experts with knowledge of how to assess an institution, formulate realistic institutional development plans, and manage institutional change.

⁸ IMF fiscal years run from 1 May to 30 April. The period covered by the evaluation is 1 May 2017 to 30 April 2022.

⁹ See *Japan-IMF Partnership on Capacity Development, Annual Report 2023*, Table 1, p. 7.

¹⁰ *Ibid.*

3. ‘Political economy thinking’, making use of ‘political economy analysis’ to inform decision-making on whether to engage in an intervention, or to change its goals.

Our current evaluation found many positive developments in the design and performance of JSA-funded CD, although some of the same challenges remain. CD outputs are of high technical quality, and the Fund has generally avoided duplication with other CD providers. Improved use of Results-Based Management (RBM) and log frames has led to better defined milestones that help monitor implementation, and there has been progress in defining objectives, outcomes and outputs. The strengthened strategic goal of aligning CD with surveillance and lending has *de facto* led to more political economy considerations in the design and implementation of programs. That said, the visibility of the Government of Japan as funding source for JSA appears to have declined since the previous evaluation. While all IMF implementing staff are aware that their JSA project is funded by the Government of Japan, this is often not the case among recipient authorities and Area Department staff.¹¹

1.2 ABOUT THE CURRENT EVALUATION

1.2.1 BACKGROUND AND AIMS

This independent evaluation was requested and sponsored by the Ministry of Finance, Japan (JMOF) in its capacity as agency responsible for the JSA.

The aim of the evaluation was to assess the degree to which the JSA-funded projects achieved their respective objectives, and to inform the IMF and Japan’s MOF regarding the performance of JSA-funded programs. It was also intended to identify good practices that could be replicated in other relevant CD contexts, as well as clear and actionable recommendations to further improve CD delivery and entity-level operations - including factors which either enhanced or detracted from the ability of JSA-funded programs to achieve their objectives.

The evaluation also paid attention to certain issues of interest to Japan - namely coordination in the field between JSA-funded IMF CD and other donor partners; integration of JSA-funded programs with IMF surveillance and lending activities; maximizing and accurately capturing the outcomes achieved by JSA-funded programs, and increasing visibility for Japan on JSA-funded CD activities.

The evaluation assessed CD programs (most of which consist of several individual country-level projects¹²) taking place between 1 May 2017 and 30 April 2022. This included JSA-funded online training and the STI, but - by agreement with the IMF and the JMOF - excluded the IMF-Japan Tax seminars (since these are not evaluable in the same way as other JSA-funded activities).

1.2.2 EVALUATION CRITERIA AND QUESTIONS

The evaluation examined JSA’s performance against CD objectives and related outcomes (defined in the IMF’s Results-Based Management (RBM) framework), according to the Organisation for Economic Cooperation and Development - Development Assistance Committee (OECD-DAC) criteria of relevance, efficiency, coherence, effectiveness, impact and sustainability.¹³ The overarching framework for assessing CD at the project level was provided by the IMF’s Updated Common Evaluation Framework (CEF) and Guidance Note.¹⁴

The main basis of the evaluation of individual projects were the log frames in the IMF’s RBM Framework.¹⁵ Log frames and monitoring data for JSA-funded programs were made available to us during the inception phase.¹⁶ We used the RBM log frames to understand project goals and design, as well as to inform our judgements when assessing selected projects against the six OECD-DAC criteria, namely:

¹¹ See Section 2.2.2.

¹² NB - under the IMF’s previous terminology, ‘project’ referred to overall initiatives rather than individual country-level activities. However, the term ‘program’ is now used to refer to overall activities, with ‘project’ describing an individual country-level element of a program. For the current evaluation, we use the latter approach - i.e., using the term ‘project’ to describe individual country-level elements of overall JSA ‘programs’.

¹³ <https://www.oecd.org/dac/evaluation/revise-evaluation-criteria-dec-2019.pdf>

¹⁴ <https://www.imf.org/en/Publications/Policy-Papers/Issues/2020/09/24/Updated-Common-Evaluation-Framework-For-IMF-Capacity-Development-And-Guidance-Note-49779>

¹⁵ The RBM Framework is applicable to single and multi-country TA and training.

¹⁶ IMF CD providers are required to specify standardized log frames that draw on the Fund’s RBM catalogue. However, while CD projects are increasingly linked to logical frameworks, full coverage has not been accomplished. The evaluation therefore drew on the RBM Governance Framework where possible.

- Relevance - the extent to which the intervention objectives and design respond to beneficiaries, global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.
- Coherence - the compatibility of the intervention with other interventions in a country, sector or institution.
- Efficiency - the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.
- Effectiveness - the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.
- Impact - the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.
- Sustainability - the extent to which the net benefits of the intervention will continue.

These criteria were applied throughout the evaluation, including to frame and structure our methodologies and the presentation of our findings in the current report.

1.2.3 METHODOLOGY

The evaluation aimed to gather and analyse data on all 32 JSA-funded programs taking place in the period subject to evaluation. A sample of selected programs was subject to more in-depth analysis, including document review and scoring against the OECD-DAC evaluation criteria. Our key methodologies are summarized below.

Document review

We examined a range of documents relevant to the JSA and its wider implementation environment, including overall JSA annual reports, individual project and program documentation, and third-party reports and analysis.

Semi-structured interviews

We carried out semi-structured interviews with a range of stakeholders at central/global, regional and country levels. The majority of these interviews were carried out remotely via MS Teams. In-person interviews were held with some IMF staff based in Washington DC, and with staff and stakeholders in Cambodia and Nigeria (where we carried out on-site field visits).

E-survey of JSA recipient authorities and IMF Staff

We ran three e-surveys using the MS Forms platform, covering the following groups:

- JSA recipient authorities - survey in English or French
- IMF implementing staff - survey in English
- IMF Country Mission Chiefs (MCs) and Resident Representatives (RRs) - survey in English

These surveys were carried out between 22 September and 29 October 2023. The survey of recipient authorities received 39 responses (33 in English and six in French). This equated to 17% of possible individual respondents, covering at least one project from 64%¹⁷ of JSA-funded programs.¹⁸ The survey of IMF implementing staff received 54 responses (a response rate of 59%), and the survey of MCs and RRs 25 responses (a response rate of 20%).

¹⁷ The numerator for this figure excludes four training programs, since these were not covered by our survey.

¹⁸ NB - these are relatively high response rates for this type of survey, where factors such as poor connectivity, staff turnover and limited administrative capacity generally limit recipients' ability to respond. In this case, we note that IMF also struggled to provide us with a complete list of contact names and email addresses for all JSA-funded projects, leading to implementation challenges on the recipient authority survey in particular. These issues would be addressed by our Recommendation 7, i.e., that the Fund maintain centrally an up-to-date list of contacts for each project, both during implementation and for at least one year after completion.

In-depth review of a sample of JSA programs

We carried out in-depth reviews of a sample of 15¹⁹ JSA programs. These covered all the individual countries/projects within the selected programs, including a detailed review of all program/project documents and data, triangulated against interviews with IMF staff and other stakeholders, and review of other contextual documents such as IMF Article IV Surveillance / Use of Fund Resources reports, and other country publications.²⁰

The selection of the 15 programs by lead department was based on the proportion of the JSA budget²¹ allocated to each department during the evaluation period, as shown in **Figure 2** below.

Figure 2: Proportion of JSA funds allocated to IMF departments, and corresponding sample profile

DEPARTMENT	% OF BUDGETED JSA EXPENDITURE ²²	IN PROPORTION FOR SAMPLE OF 15	ACTUAL PROFILE ACHIEVED IN SAMPLE
Fiscal Affairs Department (FAD)	40%	6	6
Institute for Capacity Development (ICD)	27%	4	4
Monetary and Capital Markets (MCM)	22%	4	4
Statistics Department (STA)	9%	1	1
Legal Department (LEG)	1%	0	0
TOTAL	100%	15	15

Source: BDO analysis of IMF datasheet ‘Japan Projects and Execution FY18-FY22’

We assessed every program in our sample against OECD-DAC criteria, using a 1-4 scale.²³ Our reviewers used a standard scoring template, based on the detailed definitions and sub-criteria in the IMF Updated Common Evaluation Framework (CEF) and Guidance Note.²⁴ For the effectiveness criterion, we started with the RBM ratings on milestones and outcomes, which we triangulated against the views of recipient authorities and IMF staff as gathered through surveys and interviews. We also reviewed program assessments prepared by IMF staff, and TA and surveillance reports. In assessing other OECD-DAC criteria, we used evidence from surveys, interviews and document reviews to produce a composite picture of performance.

For five of these programs, we conducted ‘deep dives’, which supplemented the document review through semi-structured interviews with a range of stakeholders (including recipient authority officials, IMF implementation staff and other development partners). Some ‘deep dives’ included in-person fieldwork in two countries where the programs had taken place, Cambodia and Nigeria. We examined several projects from different programs delivered in these countries. In other countries, ‘deep dive’ programs were covered through virtual interviews.

1.3 STRUCTURE OF THIS REPORT

This main evaluation report provides a summary of the work carried out by the BDO team, and the resultant findings, conclusions and recommendations. It is structured as follows:

- **Section 1** sets out the background and context to the evaluation.
- **Section 2** presents our findings on how well JSA-funded activities have performed against the six OECD-DAC criteria, including key lessons learned for design and implementation of future CD.
- **Section 3** sets out our conclusions and recommendations.

¹⁹ NB - one program we covered (ICD_APD_2022_01) was effectively a continuation of an earlier program (ICD_STI_2019_01), delivered by the STI. We reviewed the documentation for both programs, but have counted them as one program for the purposes of reporting.

²⁰ <https://www.imf.org/en/Publications/SPROLLs/Article-iv-staff-reports#sort=%40imfdate%20descending>

²¹ Source: BDO analysis of IMF datasheet ‘Japan Projects and Execution FY18-FY22’.

²² Figures are for budgeted JSA expenditure during the evaluation period. They may not sum precisely due to rounding.

²³ The ratings had the following interpretation: 1=poor, 2=modest, 3=good, and 4=excellent.

²⁴ <https://www.imf.org/en/Publications/Policy-Papers/Issues/2020/09/24/Updated-Common-Evaluation-Framework-For-IMF-Capacity-Development-And-Guidance-Note-49779>

2. HOW WELL DID JSA-FUNDED ACTIVITIES PERFORM?

This section sets out our assessment of how well JSA-funded activities performed against the six OECD-DAC criteria (relevance, coherence, effectiveness, efficiency, impact and sustainability). It presents aggregate scores for each of these criteria, as well as commentary on particular points of interest and examples.

Where relevant, the scoring of sample programs is disaggregated by lead department (FAD, ICD, MCM or STA²⁵) and by the region where the programs (or, in some cases, their individual projects) took place. We also provide sub-analysis of scores between fragile/conflict-affected states²⁶ and others, since this was an area of interest expressed by the Government of Japan at the inception stage.

2.1 RELEVANCE

2.1.1 OVERVIEW

Relevance is the extent to which the intervention objectives and design respond to beneficiaries, global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.

Our detailed review of a sample of JSA-funded programs found that these are generally relevant to their context and the needs of the recipient authorities (Figure 3). We found no significant difference between the relevance of projects delivered in African vs Asia Pacific countries (Figure 4), or fragile/conflict-affected states versus other (chiefly emerging economies) (Figure 5).

Figure 3: BDO reviewers' scores for the relevance of JSA-funded projects, by lead department

Lead Department	FAD	ICD	MCM	STA	JSA overall
Relevance (mean assessment score)	2.9	3.2	2.8	2.9	2.9

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 4: BDO reviewers' scores for the relevance of JSA-funded projects, by region

Region	African Countries	Asia Pacific Countries	Global / Mixed ²⁷	JSA overall
Relevance (mean assessment score)	2.9	3.0	3.0	2.9

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 5: BDO reviewers' scores for the relevance of JSA-funded projects, fragile vs non-fragile states

Classification	Fragile states	Non-fragile states	Global / Mixed ²⁸	JSA overall
Relevance (mean assessment score)	2.9	2.9	3.0	2.9

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

²⁵ The sample of 15 programs selected for in-depth document review was based on the proportion of JSA expenditure by each lead department during the evaluation period (see Section 1.2.3). As the IMF Legal Department (LEG) TA funded by JSA made up just 1% of total JSA expenditure, it did not feature in our sample of 15 programs. It was, however, included in our surveys of IMF staff and recipient authorities.

²⁶ For this analysis, we defined countries as fragile or conflict-affected if they appeared on the World Bank's current FCS list: <https://thedocs.worldbank.org/en/doc/608a53dd83f21ef6712b5dfef050b00b-0090082023/original/FCListFY24-final.pdf>. Although a few countries may have been added to or removed from this list during the evaluation period, this is unlikely to significantly affect the analysis presented here.

²⁷ Global / Mixed refers to programs such as online training, which could include participants from more than one region.

²⁸ Global / Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

In most cases we were able to identify appropriate linkages to government priorities / activities, and to findings/recommendations from IMF surveillance and strategy documents. We also identified evidence that many projects were addressing a need expressed or acknowledged by recipient authorities, which in turn led to higher scores for these projects on effectiveness and sustainability. This suggests that the IMF is successfully implementing the 2018 CD Review strategic goals²⁹ of aligning CD with surveillance and lending activities, and giving a greater role to Area Department country teams in setting priorities and selecting CD activities. Greater involvement of Area Departments and further integration with surveillance and lending activities has led to better alignment of CD activities with country needs and willingness to address capacity gaps, since Area Department country teams are better placed than CD departments to conduct needs analysis and capacity assessments.³⁰

There were a small number of projects where recipient authority 'buy-in' was clearly lacking from an early stage. This was often in the context of a multi-country program, where the Fund had expected that buy-in in certain countries would develop after the project was launched, and that problems could be rectified during implementation. Although this happened in some cases, in others the lack of ownership persisted throughout - and ultimately led to poor scores for effectiveness and sustainability.³¹ In some of the poorest performers, resources were eventually reassigned to other countries.

The related issue of low absorption capacity at recipient authorities remains a problem, particularly in low-income and fragile states. The Risk Assessment and Mitigation module of project proposals requires teams to consider questions about political support, resource adequacy, and commitment by management and technical staff - all aspects of absorption capacity. However, the risk management and mitigating measures are often generic rather than country-specific, or left blank / pending until the program has been launched, and authorities can be involved.

Interim project assessments are generally more specific about the extent of risks and intended mitigations, and final assessments often discuss whether risks materialized, or are likely to upon completion. While absorption capacity will remain a serious risk in low-income and fragile states, there may be scope for proposal and planning documentation to include more explicit evidence (e.g., an official request from the recipient authority) that the intervention meets a need identified or acknowledged by the authorities, and that the authority has the commitment to support its delivery. Our work suggests that such commitment can be easier to obtain where authorities are already engaged with the IMF through surveillance or lending arrangements - as was the case with several fragile states in our sample that were accessing the IMF Extended Credit Facility (ECF) whilst also receiving TA.³²

Notwithstanding the above-mentioned cases of limited 'buy-in', our e-survey of IMF staff and recipient authorities found that JSA-funded TA is generally seen as relevant to its context and recipients' needs. While recipient authorities noted that the impetus for initiating projects frequently came from the IMF rather than themselves,³³ this does not appear to have prevented projects from targeting the areas of greatest need, and (in most cases) taking on board the views of recipient authorities when designing scope and approach. For example, 88%³⁴ of recipient authorities responding to our survey agreed that their JSA-funded project was

²⁹ See *2018 Review of the Fund's Capacity Development Strategy - Overview Paper* (<https://www.imf.org/en/Publications/Policy-Papers/Issues/2018/11/20/2018-review-of-the-funds-capacity-development-strategy>), p. 22.

³⁰ These processes did not go as far as the full systematic 'political economy analysis' and 'design thinking' recommended by the previous JSA evaluation (Section 1.1.3), but they clearly enhanced the relevance of CD activities. In fact, given the size of each CD country project, it would not be realistic to expect such a comprehensive political economy analysis in advance of each CD intervention.

³¹ For example one country-project within the program FAD_AFR_2019_01 ('Strengthening fiscal sustainability through core budget functions in Fragile States in Sub-Saharan Africa') struggled initially with authority 'buy-in'. But intensive engagement by the LTX appears to have improved working relationships and fostered greater ownership by the authority as the project progressed. Conversely, one country project under the program FAD_APD_2020_02 (Strengthening PFM in selected South-East Asian countries) set out to implement Cash Basis IPSAS and (in collaboration with the World Bank) modernize the chart of accounts (CoA). However, due to authority capacity constraints (qualified managers already being overloaded), lack of understanding among counterparts and poor collaboration between various departments within the Ministry of Finance, the project struggled to achieve traction throughout - scoring poorly not just for relevance, but also for effectiveness, impact and sustainability.

³² E.g., in the program MCM_AFR_2021_02 ('Strengthening debt management operational frameworks in African countries').

³³ For example, 41% (±14%) of recipient authorities said that IMF staff played the leading role in initiating the project. 33% (±13%) said it was an equal joint decision between IMF staff and national authorities, and 26% (±13%) said it was the Ministry of Finance or national authorities which played the leading role. Source: BDO e-survey of recipient authorities (n=39).

³⁴ Margin of error at 95% confidence level = ±9%.

‘correctly targeted at our country’s development needs’, and the same proportion that it was ‘correctly targeted to fill gaps in our national capacity in the relevant sector.’ 77%³⁵ agreed that their views ‘were taken on board in the design of the JSA-funded project.’

IMF staff are also positive about the relevance of JSA-funded projects, with both Implementing Managers³⁶ and Mission Chiefs / Resident Representatives³⁷ reporting relevance scores similar to (or higher than) recipient authorities.

2.2 COHERENCE

2.2.1 OVERVIEW

Coherence is the compatibility of the intervention with other interventions in a country, sector or institution.

Our reviewers found that the coherence of most JSA-funded interventions was generally strong (Figure 6). In most cases we were able to identify appropriate linkages and complementarities with government priorities / activities and other donor-funded interventions (both IMF and others) in the relevant country. This was equally true of all the departments covered by our review, and with no major differences between geographical regions or types of country (Figure 7 and Figure 8).

Figure 6: BDO reviewers’ scores for the coherence of JSA-funded projects, by lead department

Lead Department	FAD	ICD	MCM	STA	JSA overall
Coherence (mean assessment score)	3.0	3.3	2.9	2.9	3.0

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 7: BDO reviewers’ scores for the coherence of JSA-funded projects, by region

Region	African Countries	Asia Pacific Countries	Global / Mixed ³⁸	JSA overall
Coherence (mean assessment score)	2.9	3.1	3.0	3.0

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 8: BDO reviewers’ scores for the coherence of JSA-funded projects, fragile vs non-fragile states

Classification	Fragile states	Non-fragile states	Global / Mixed ³⁹	JSA overall
Coherence (mean assessment score)	3.0	3.0	3.0	3.0

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

In line with the 2018 strategic objective to better integrate CD with surveillance and lending, we found that JSA-funded TA and training were almost always directed towards priorities identified in Article IV consultations and/or contributing to implementation of Fund-supported lending programs. This integration appears to have gone beyond teams simply listing known nearby CD activities in program documentation, and to have enhanced the relevance and coherence of CD, as well as its likely effectiveness. Examples of such integration included the Revenue Administration COVID-19 (FAD_IMF_2021_09) program, which supported two African countries in implementing their respective ECF-supported programs. Similarly, MCM’s CD support to Mongolia (MCM_MNG_2018_04) contributed to implementation of the financial sector reforms included under its Fund-supported Extended Funding Facility (EFF). Another example was the program Macroeconomic Frameworks

³⁵ Margin of error at 95% confidence level = $\pm 12\%$.

³⁶ Between 85% and 100% of IMF Implementing Managers agreed with the equivalent statements regarding relevance (Margin of error at 95% confidence level = $\pm 6\%$ and $\pm 2\%$ respectively). Source: BDO e-survey of IMF implementing staff (n=54).

³⁷ Between 76% and 84% of IMF MCs and RRs agreed with the equivalent statements regarding relevance (Margin of error at 95% confidence level = $\pm 15\%$ and $\pm 13\%$ respectively). Source: BDO e-survey of Mission Chiefs and Resident Representatives and (n=25).

³⁸ Global /Mixed refers to programs such as online training, which could include participants from more than one region.

³⁹ Global /Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

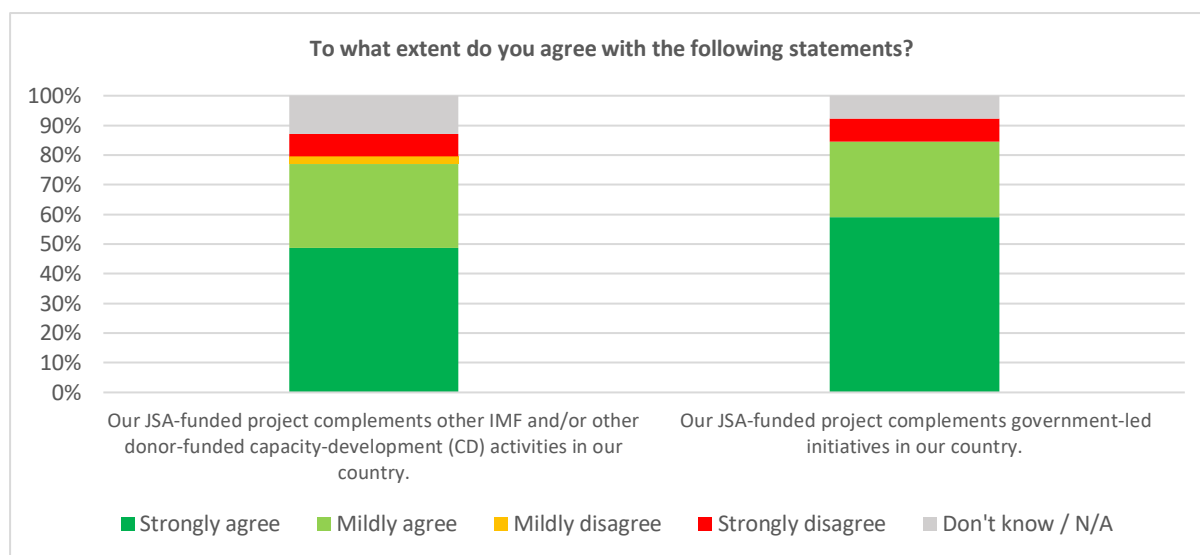
(ICD_IMF_2021-05). This program assisted several central banks to improve their forecasting capacities, which IMF surveillance had identified as a priority in the participating countries.⁴⁰

Like the previous JSA evaluation team⁴¹ we identified few examples of actual active synergies between JSA-funded projects and other donor-funded interventions - although this was mainly due to lack of overlap with the subject matter rather than actual missed opportunities for joint working. Notable exceptions included a three-way collaboration on revenue administration on one African country, whereby the JSA-funded IMF project explicitly built on and incorporated work by the African Development Bank (who prepared a reform plan) and the Japan International Cooperation Agency (who developed customs classification rules).

Several interviewees noted the key role of recipient authorities in coordinating the different TA interventions to avoid duplication between different CD providers and projects. Also, several stakeholders noted that a coherent approach between the IMF and the World Bank was facilitated by good relations and communication at HQ level in Washington DC, as well as contact in-country between IMF experts and WB resident missions.

Respondents to our e-surveys were also very positive about the coherence of JSA-funded projects, both with other donor-funded interventions and with domestic government policies. Around 80%⁴² of recipient authorities agreed that the JSA-funded intervention complemented other relevant activities in their country (Figure 9).

Figure 9: Recipient authorities' views on the coherence of their JSA-funded project



Source: BDO e-survey of recipient authorities (n=39)

2.2.2 AWARENESS OF THE GOVERNMENT OF JAPAN AS THE SOURCE OF JSA FUNDING

One specific area we were asked to consider as part of our evaluation was 'visibility', i.e., the extent to which recipient authorities and other stakeholders were aware that the source of funding for IMF JSA TA was the Government of Japan.

We explored this question chiefly through our e-surveys, triangulated with evidence from stakeholder interviews and case studies. In general, we found that awareness of the funding source for JSA TA amongst recipient authorities was fairly good, albeit with scope for further improvement. For example, more than a third⁴³ of recipient authorities responding to our survey were unaware that their JSA project was funded by the Government of Japan.

⁴⁰ The goals and design of this program were highly relevant, which (as mentioned below) contributed to achieving outstanding results in most of the participating countries. However, in one participating fragile state there was very limited progress. This was because the authorities were distracted by elections and the pandemic, despite agreeing with the Fund that the activities under this program were high priority.

⁴¹ ENCLUDE and SEO Amsterdam Economics, *Final Report - Evaluation of the Japan Technical Assistance Subaccount (JSA) Capacity Development Programs*, 2017.

⁴² Margin of error of at 95% confidence level = \pm c.12%.

⁴³ 36%. Margin of error at 95% confidence level = \pm 14%.

Notably, almost two thirds of the recipient authorities who were aware that the funding was from the Government of Japan had learned this only through discussions with IMF staff rather than ‘branding’ or other sources. This suggests there may be scope to improve visibility by increasing the JSA ‘branding’ on project documents and - equally importantly - ensuring that IMF staff make clear to their counterparts at recipient authorities who is funding the TA.

IMF implementation staff are almost always aware that JSA is the source of funding. However, more than half⁴⁴ of IMF Mission Chiefs and Resident Representatives responding to our survey were not. Indeed, a number of potential respondents from this group informed us that they could not complete a survey specifically about JSA-funded TA, since they were not aware of which projects and programs in their country were funded from that source. This is not surprising (although this information is available), since the key concern for Mission Chiefs and Resident Representatives is likely to be the focus and design of the CD, not its funding source. However, given the importance which donor partners (including the Government of Japan) place on visibility, it may be helpful to improve branding and communication between TA managers and permanent in-country IMF staff.

2.3 EFFECTIVENESS

2.3.1 OVERVIEW

Effectiveness is the extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.

Our reviewers assessed the effectiveness of JSA-funded projects based on both RBM data / assessment reports and other sources - notably interviews, wider document review and our e-surveys. Overall, we found that JSA-funded TA was fairly effective, including some outstanding examples of projects which delivered all their planned outputs and outcomes, as well as receiving excellent feedback from the recipient authorities.⁴⁵

That said, with the exception of ICD training programs, the scores for effectiveness⁴⁶ across the JSA portfolio were somewhat lower than for relevance and coherence (Figure 10).

Figure 10: BDO reviewers’ scores for the effectiveness of JSA-funded projects, by lead department

Lead Department	FAD	ICD	MCM	STA	JSA overall
Effectiveness (mean assessment score)	2.3	3.3	2.5	2.7	2.5

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

These differences in average effectiveness ratings are to be expected, given the greater complexity of FAD and MCM projects that involve deep and often lengthy institutional reforms.⁴⁷ By contrast, many STA programs had more straightforward objectives, such as helping officials to organize statistical databases according to international standards, and providing training on how to maintain them.

The higher scores for ICD-led programs are partly because many consisted solely of training, which involves structured and repeatable processes of knowledge transfer, with clear and consistent metrics for success (e.g.,

⁴⁴ 52%. Margin of error at 95% confidence level = $\pm 18\%$.

⁴⁵ For example, two country projects in the ICD Macroeconomics Program ICD_IMF_2021_05 achieved outstanding results in terms of delivering all planned outputs and outcomes and receiving excellent feedback from the authorities. In one country, the project trained authorities in Debt Dynamics Tools (DDTs), thus enabling them to develop their own tool - which they continue to use for public debt forecasts and scenarios. This project is also a good example of the benefits of integrating TA with surveillance/lending, since the development of a debt forecasting tool was a priority under the country’s 2022 ECF. A different country project in the same program helped the authority to customize forecasting tools, and develop sophisticated models for use in policy-making. All outputs and outcomes were achieved, and the authorities told us that the TA helped catalyze and accelerate development of these tools. The success of this project illustrates that TA can be very effective in moderate- to high-capacity countries when the objectives are shared by the authorities.

⁴⁶ NB - the multi-country programs in our sample contained a few projects that were either never launched, or were cancelled early in their implementation. As these projects consumed minimal resources, and did not run long enough for their effectiveness to be assessed, they are excluded from the analysis and scores above.

⁴⁷ For example, the program FAD_IMF_2021_10 (‘Public Financial Management Covid-19’) successfully introduced the authorities in one African country to new tools for risk identification and management, but made only limited progress with the wider outcome of [strengthening] ‘central fiscal oversight and analysis of sub-national government - largely due to the complicated political situation and associated difficulties in institutionalizing the reforms. Similarly, the program MCM_AFR_2021_02 (‘Strengthening Debt Management Operational Frameworks in African Countries’) failed in several countries to achieve objectives such as integrating cash/debt management and implementing risk management frameworks, due to low capacity compared to the scale of implementation required.

participant feedback ratings, ‘before and after’ test scores). The high effectiveness scores for ICD training also reflect the success with which these programs adapted their design and delivery to maintain effectiveness despite the COVID-19 pandemic. At the start of the pandemic, JSA support allowed training to shift to virtual and online courses, as ICD scaled up its online program in response to a large increase in demand.⁴⁸ These new modalities led to a significant scaling up of reach, while the cancellation of in-person training and TA led to significant underspending. Overall, there was a significant reduction in the average cost per participant while maintaining or improving participant satisfaction and learning outcomes - indications of effectiveness. User surveys on utilization and institutional absorption of knowledge also indicate that the longer-term benefits of training (such as continued use of knowledge by staff in a relevant role) were not affected.⁴⁹

One key reason for some of the low scores was the presence in our sample of some ‘outlier’ projects which failed to achieve most or all of their objectives, and a significant number which achieved their objectives only partially.⁵⁰ The main factor affecting the effectiveness of non-training programs in particular was the COVID-19 pandemic, which impacted significantly on implementation across most of the JSA portfolio. Although some TA programs later adapted to online or hybrid modes of delivery, the initial impact of COVID disruption nonetheless affected their ability to deliver their planned outputs, outcomes and objectives as scheduled.

Notably, while the *de facto* improvements achieved by CD differed between regions and between fragile and non-fragile countries, this was not reflected in significantly different effectiveness ratings for the corresponding JSA-funded interventions. One reason for this is that our reviewers scored effectiveness in line with RBM guidelines, i.e., in terms of achievement of the planned objectives of each project or program, rather than as a measure of absolute progress or improvement. As JSA-funded projects in African and/or fragile countries sometimes set more modest objectives than those in Asian countries, their scores for effectiveness are roughly the same - since the extent to which they achieve these objectives is broadly similar (Figure 11 and Figure 12).⁵¹

Figure 11: BDO reviewers’ scores for the effectiveness of JSA-funded projects, by region

Region	African Countries	Asia Pacific Countries	Global / Mixed ⁵²	JSA overall
Effectiveness (mean assessment score)	2.5	2.4	3.2	2.5

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 12: BDO reviewers’ scores for the effectiveness of JSA-funded projects, fragile vs non-fragile states

Classification	Fragile states	Non-fragile states	Global / Mixed ⁵³	JSA overall
Effectiveness (mean assessment score)	2.4	2.5	3.2	2.5

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Our work suggested that the main drivers of effectiveness are:

⁴⁸ For example, the development and launch of Massive Open Online Courses (MOOCs) were accelerated, allowing many more officials and the public at large to take these courses. ICD developed short modular courses that allowed senior officials to learn while still doing their jobs, and launched new blended online/virtual courses.

⁴⁹ The number of total active learners in the online program rose grew by 90 percent (from 89K to 169K) between FY21 and FY23. Completion rates, participant satisfaction, and learning gains were at least as high as before the fast expansion (68 percent, 95 percent, and 20 percentage points respectively).

⁵⁰ For example, one country project in the program ICD_IMF_2021_05 made little or no progress in delivering the planned outputs, even though its objectives were modest, and its relevance and coherence were rated good. The project was implemented intermittently because of COVID-19, national elections, and the authorities focusing their scarce human resources on other issues. Similarly, a country project in the program FAD_APD_2020_02 was launched when country’s 2017 EFF-supported program was about to go off track, when IMF staff were attempting to resuscitate the program by addressing outstanding actions. In the event, no TA was given during the first year because the authorities did not agree to engage with virtual TA. During FY 2022, work resumed. But decentralization reform introduced in August 2021 required a realignment of the various PFM processes and suitable sequencing of reforms. A change in government in June 2021 led to reprioritization within the MOF, and ultimately the project’s objectives were not achieved (even partially).

⁵¹ NB - training programs are treated here as ‘Global / Mixed’, hence their high scores for effectiveness increase the average for this category.

⁵² Global /Mixed refers to programs such as online training, which could include participants from more than one region.

⁵³ Global /Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

- country ownership and commitment to reform (proxied by relevance, and in particular by who selected the particular project and the authorities' contribution to the project design),
- the initial conditions in the recipient country and agency (countries with better institutions and agencies with more resources used JSA-funded CD most effectively).
- better internal coordination within the Fund, e.g., between two CD departments,⁵⁴ or between a CD department and the corresponding country team - including integration of CD with surveillance and lending programs.⁵⁵

Both IMF staff and recipient authorities consider the key drivers of effectiveness to be the quality of the experts, their relationship with the participants and the delivery modality. Respondents also emphasized the importance of 'buy-in' from authorities, noting that poor 'buy-in' was a frequent contributor to reduced effectiveness, and - by extension - sustainability and impact. This tallies with the findings of our detailed review of sample programs,⁵⁶ where low scores for relevance and coherence (of which 'buy-in' is a key element under the IMF Common Evaluation Framework⁵⁷) frequently correlated with poor scores for effectiveness and sustainability.

In general, program and project complexity diminished effectiveness, as it taxed both the authorities' and staff's implementation capacity. Similarly, multi-country programs may sometimes afford returns to scale, but often complicate coordination within the Fund, and may lead to uncertainty on the availability of resources for each country-project.⁵⁸

2.3.2 QUALITY OF REPORTING ON EFFECTIVENESS

As part of our evaluation, we were asked to consider whether the effectiveness of JSA-funded projects and programs as reported by IMF Assessment Reports and Results-Based Management (RBM) data was borne out by other sources and our own judgement.

Unlike the previous evaluation of JSA,⁵⁹ our reviewers found few significant discrepancies between the effectiveness of projects as reported by IMF Assessment Reports and other sources (e.g., interviews, other documents and e-surveys). In some cases, the feedback from recipient authorities on effectiveness was actually more positive than IMF staff's view in the assessment reports - although this data should be treated with caution, since authorities were asked to comment in general on the effectiveness of their JSA-funded project, rather than with specific reference to its individual objectives.

We found that RBM was used in the design of all programs in our sample, and that there was almost universal compliance in entering the required information and ratings. More generally, in triangulating the evidence on effectiveness, we found that the implementation of RBM has contributed to better defined milestones, outputs and outcomes - even if there are still a few cases of confusion between inputs, outputs, and outcomes.⁶⁰ It must be noted, however, that such conceptual issues are inherent in the RBM framework, and caution would be needed in evaluating and comparing effectiveness across programs or countries even if the framework were perfectly implemented. More generally, we found that the RBM database remains limited as a tool for assessing each program's progress against milestones, outputs, outcomes and objectives.

⁵⁴ One notable example of good coordination between two CDDs leading to a superior outcome include was in Cambodia, where synergies were generated by simultaneously tackling GFS and PFM reforms. In this case, coordination between the two CDDs was made even more effective by placing the two respective LTXs at the same regional centre. In another APD country, by contrast, this synergy was absent, which hampered the progress of the PFM reform.

⁵⁵ One such example of coordination between a CDD and an Area Department was in the context of a lending program in an APD country, where an MCM program provided support to the IMF Area Department with implementation of structural benchmarks. The MCM program was designed to provide technical support to the country's EFF program, and made good initial progress against milestones and outcomes. However, when the EFF program went off-track around one year into TA implementation, the IMF lost active engagement with the authorities, which also impacted on progress with the TA program.

⁵⁶ See Section 2.1.

⁵⁷ See, *IMF New Common Evaluation Framework for IMF Capacity Development* (2017), p.24. <https://www.imf.org/en/Publications/Policy-Papers/Issues/2017/04/27/pp040717new-common-evaluation-framework-for-imf-capacity-development>

⁵⁸ See also Section 2.4.

⁵⁹ ENCLUDE and SEO Amsterdam Economics, *Final Report - Evaluation of the Japan Technical Assistance Subaccount (JSA) Capacity Development Programs*, 2017.

⁶⁰ For example, one project in the Macroframeworks program (ICD_IMF2021_05) used as an outcome indicator 'percentage of staff trained', which is at most an output (since the outcome would relate to the actual acquired knowledge and performance of these staff resulting from the training). See also Section 2.5 of the current report.

2.3.3 EFFECTIVENESS OF DIFFERENT DELIVERY MODELS

As part of our evaluation, we were asked to consider the relative effectiveness of different delivery models, including the use of Long-Term Experts (LTXs) versus Short-Term Experts (STXs) in different contexts.

We identified a small number of cases where the recipient authority was unhappy with the quality of the JSA-funded LTX, which impacted on both buy-in and ultimately effectiveness - particularly given the long-term nature of the interaction, and the pivotal importance of the LTX in achieving objectives. There may be scope to give recipient authorities a greater say in the drafting of the LTX's TOR, if not in the selection of the LTX him- or herself. It might also be helpful to consult with recipient authorities when deciding whether to extend or terminate the LTX's contract at the end of the first year. Recipient authorities' greater ownership and involvement in such decision-making could (at least in the case of middle-income countries) be reflected in them making a financial contribution towards the cost of deploying the LTX.

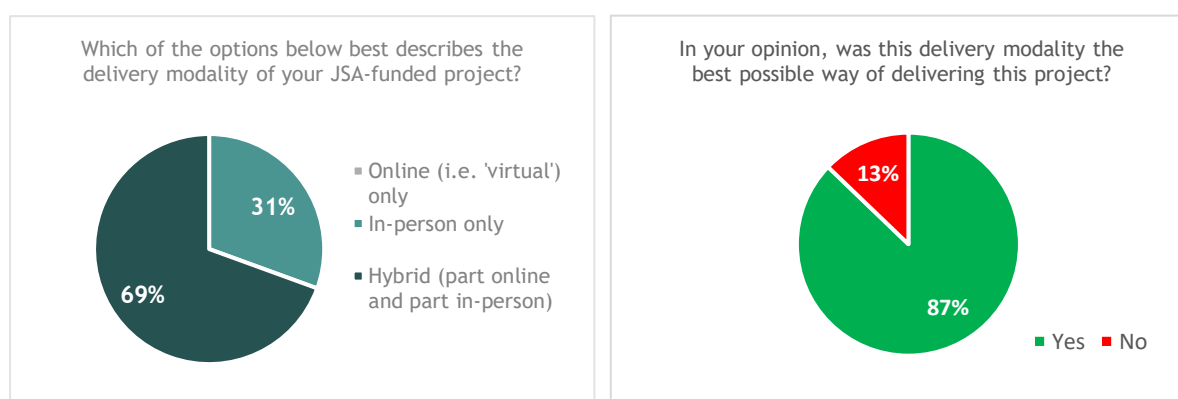
It should be noted that the choice of whether to use an LTX or an STX is sometimes determined by the political context and preference of the recipient country. For example, some countries do not want an LTX in certain sectors because of concerns over access and confidentiality. Conversely, others prefer an LTX, but insist that he/she be based permanently in their country, rather than carrying out occasional visits from a regional centre such as the IMF Capacity Development Office in Thailand (CDOT). A frequent reason for preferring an in-country LTX is that this ensures continuity of support, and the chance for the LTX to be more integrated within the recipient authority team - both for ongoing technical input and for fruitful working relationships. Our work also suggests that this ongoing presence can help to improve engagement and 'buy-in' by recipient authorities.⁶¹

As part of our evaluation, we were also asked to consider whether there were any significant differences in effectiveness of TA between in-person and virtual delivery, or indeed a combination of the two.

The JSA-funded TA programs covered by our evaluation included examples of all three delivery models - not least as the impact of COVID-19 forced several projects to switch from in-person to virtual or hybrid delivery in order to continue.

Our e-survey suggested that recipient authorities were generally happy with the chosen delivery modality, with only 13%⁶² stating that they would have preferred a different way of delivering the project (Figure 13). Despite COVID-related travel restrictions, a significant proportion of respondents (31%) had received their TA in-person only. Many emphasised the continuing importance of face-to-face approaches for some training activities and fact-finding missions and meetings.

Figure 13: Recipient authorities' views on the delivery modality of their JSA-funded project



Source: BDO e-survey of recipient authorities (n=39)⁶³

That said, the enforced switch to online, virtual and hybrid modalities showed that these approaches can be both effective and sustainable for certain activities and settings - particularly training. With the outbreak of the

⁶¹ See Section 2.1.1.

⁶² Margin of error at 95% confidence level = $\pm 10\%$.

⁶³ Note that these figures and feedback exclude pure ICD training programs, since these were not included in our recipient authority survey.

COVID-19 pandemic, the Fund was well positioned to substitute in-person training with online and virtual courses, thanks to its experience with distance learning (which had been largely financed by JSA).

The Fund's experience with distance learning also helped to substitute in-person TA with virtual missions, and to launch virtual missions for surveillance and programs.⁶⁴ Once established, these new modalities contributed to the Fund's strategic goal of integrating TA and training with surveillance. For example, some programs now require authorities to attend initial online training as a prerequisite for receiving virtual or in-person TA. Also, some surveillance activities have integrated virtual training, sometimes preceding or following the surveillance mission. The negative effect of the pandemic served to accelerate the integration of online training, synchronous virtual missions, and in-person deliveries by resident LTXs. In several instances, we observed that TA recipients were asked to take an online course before in-person TA missions, and on other occasions an in-person mission was followed up by virtual consultations. Together, these modalities appear to have been as effective and less costly than the "traditional" one-off missions by short-term experts (STXs) or regional LTXs.

The Fund and JSA's adaptability were particularly visible in the design of programs aimed at addressing the challenges caused by the pandemic, e.g., programs to strengthen Public Financial Management, Revenue Administration, and Macroeconomic Frameworks during COVID.⁶⁵ All of these featured a combination of modalities to maintain delivery despite considerable physical and administrative challenges.

2.4 EFFICIENCY

2.4.1 OVERVIEW

Efficiency is the extent to which an intervention delivers, or is likely to deliver, results in an economic and timely way.

It is currently extremely difficult to assess the efficiency / cost-effectiveness of JSA-funded TA (and IMF programs more widely) using available information, because:

- It is not possible to 'monetize' activities, outputs and outcomes robustly and comparably, even where projects/programs have similar themes and methodologies.⁶⁶ For example, the 'value' of outputs from similar activities will differ depending on the structure and role of the recipient authority; the country's size, economic circumstances and degree of development, and the nature and extent of any resultant changes in resource use and financial flows. Similarly, it is difficult to compare across activities where results can be seen immediately (e.g., developing forecasting models or launching a database) and where results may take a long time to emerge (for example, strengthening customs administration).
- Until recently, cost data was generally not collected or reported by project, only program - and even then on a limited basis. More complete information is now being collected via CDMAP, the IMF's new integrated management system which was launched in late 2021. However, not all CDMAP's planned functions are currently operational. For most of the programs covered by this evaluation, project documentation does not disaggregate budgets between country projects, and there is limited information on how the budgets were actually used.
- Budget figures are frequently adjusted or reiterated without a full 'audit trail', and are not consistently updated on a rolling basis - meaning that the budget figure provided in program documentation may not be current at the time it is reviewed or referenced by studies such as this.
- Assessing JSA projects' 'timeliness' (another element of efficiency as per the OECD-DAC criterion) is also challenging, as up-to-date information on programs' delivery against milestones is not routinely available - particularly at individual project level. Moreover, the target dates for milestones can often justifiably change as part of ongoing project management. This was a common occurrence during the evaluation period, due to the impact of COVID.

⁶⁴ Virtual missions consist of a concentrated period of engagement with the TA-recipients that mimics as much as possible the dynamics of a TA mission.

⁶⁵ See in particular, FAD_IMF_2021_10 (Strengthening PFM during COVID), FAD_IMF_2021_09 (Strengthening Revenue Administration during COVID), and ICD_IMF_2021_05 (ICD Macroeconomic Frameworks, COVID 19 Window).

⁶⁶ One exception is JSA-funded training, where there are estimates of the variable cost of individual courses and the number of participants. The 2019 Project Assessment of the Japan-IMF Initiative on Online Training states that "Average participation is 234 government officials per course, compared with 30 officials per face-to-face course. Cost of delivery is about one-tenth that of a face-to-face course."

- The COVID-19 pandemic impacted significantly on IMF teams' ability to deliver TA, affecting both cost (due to breakdown of planned delivery models and non-availability of recipient authorities) and timeliness (due to extended lockdowns and resultant slippage on key milestones). Hence the two key elements of any evaluation of efficiency (cost and time) were both subject to exceptional disruption during the evaluation period.

In the absence of data to support robust quantitative analysis of cost-effectiveness and efficiency, we used our detailed review of 15 programs and our e-surveys to gather qualitative data on risks to and drivers of efficiency in implementing JSA-funded TA. The key issues identified are discussed in more detail below.

2.4.2 EFFICIENCY OF SINGLE VS MULTI-COUNTRY PROGRAMS

Multi-country programs have been a flexible and generally efficient way to deliver JSA-funded CD in the evaluation period. That said, we identified some concerns relating to efficiency/cost-effectiveness, and - linked to this - accountability and transparency.

Multi-country programs allow the IMF to launch a CD program without having to invest in advance the same amount of time and resources that would be required for a single country program, because they do away with the need to ascertain authorities' full commitment to implement reforms in each and every country. Once the program is in place, the Fund can shift resources from one recipient country to another in response to changing priorities and circumstances. The employment of an LTX is much more cost-effective in a multi-country program than in a single-country program, to the extent that a single LTX can be shared by multiple countries. As an additional advantage, a multi-country program, by bringing a diversity of countries together, can provide opportunities for peer learning.

On the other hand, a multi-country program may make it difficult to identify and hire appropriate experts if the recipient countries are very heterogeneous, and also makes program management more difficult. For example, we found cases where a country manager⁶⁷ could not respond to a request because of uncertainty about the availability of resources, even though the overall program budget had an underspend. These difficulties arise in part because there is insufficient information on the allocation of budgeted resources across the different country-projects (a problem that is expected to diminish once all CDMAP's functions are fully implemented). Finally, related to the points above, multi-country programs make it difficult to evaluate completed programs when some countries have achieved results while others have not.⁶⁸

There may be scope to maximize the benefits of multi-country programs, whilst ensuring that risks to transparency, accountability and program management are managed. For instance, in a full-fledged reform program (as opposed to training or diagnostic programs), an upper limit to the number of recipient countries could be considered; restrictions could be considered for regional coverage (e.g., APD and AFR could not be combined); explicit criteria could be established for when to reallocate resources from one country to another, and an indicative budget could be set for each country at the outset of the program.

2.4.3 REALLOCATING RESOURCES TO OPTIMISE EFFICIENCY

One question which we were asked to consider in our evaluation was whether there is scope to identify 'failing' projects and programs earlier, so that these could potentially be cancelled, and the funds transferred to activities more likely to deliver results - and hence improve efficiency and VFM.

Our detailed review of a sample of JSA activities identified very few programs where there was a clear-cut case for cancelling the work and transferring the funding. More generally, there is little evidence that it would be practical or useful to try to identify programs for early cancellation. Currently, the vast majority of programs are relatively short. Therefore, it would not be practical to cancel them ahead of time, considering the processes and documentation that would be needed in both the Fund and JSA. In most cases, the Fund and JSA would need to prepare a thorough document to explain the cancellation to the recipient authority - the annual interim assessments currently being produced would not be adequate for that task. Moreover, our reviewers found that the amount and quality of information provided in interim assessment reports was frequently limited, and would not have supported a robust decision to cancel or reorientate the program at that stage.

⁶⁷ This refers to the country "backstopper" in CDDs who oversees all topical CD to that country.

⁶⁸ Fund program managers whom we interviewed noted that a more diverse range of topics, often stretching across different divisions in a CDD, can also hamper program management.

At the country project level, multi-country programs do provide some opportunity for responsive reallocation of resources away from projects which are making no or limited progress. Our reviewers found several multi-country programs where activities had been reduced to a minimum in countries suffering major implementation challenges, and the resources transferred elsewhere within the ambit of the same program. Indeed, several interviewees suggested that transferring resources whilst retaining a largely notional project in a problematic country could be a useful way of maintaining engagement without significant cost. Although such projects contribute little to achieving the program's planned objectives, they can potentially 'keep the door' open to more effective and extensive TA should circumstances in the relevant country improve.

2.5 IMPACT

2.5.1 OVERVIEW

Impact is the extent to which an intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects - for example, faster economic growth, lower fiscal deficits or greater access to international capital markets.

Our reviewers generally struggled to identify clear wider impacts from JSA-funded projects. Hence the scores across most departments (**Figure 14**), regions (**Figure 15**) and country types (**Figure 16**) were generally lower than for other criteria, notably effectiveness. However, it should be noted that these low scores do not reflect positive evidence that programs have failed to achieve their intended impacts, but rather lack of evidence of attributable impacts *per se*. The difficulty of robustly evidencing impacts is discussed in more detail in **Section 2.5.2** below, and relates chiefly to short timeframes, data limitations and difficulties of attribution.

In some cases, our reviewers did identify some modest effects beyond the planned outcomes of a project, and generally recognized these with a higher score than the minimum (1). However, the lack of data meant that we found few examples of far-reaching and attributable impacts, and many projects where there was no evidence of any effects beyond the planned activities such as training,⁶⁹ or preparation of documents and tools. Several MCM projects in particular had modest objectives (e.g., analysis of steps to develop a yield curve) which were unlikely to have generated visible impacts. Indeed, several such projects had also been disrupted by severe political unrest - further limiting their scope for significant impact.

Figure 14: BDO reviewers' scores for the impact of JSA-funded projects, by lead department

Lead Department	FAD	ICD	MCM	STA	JSA overall
Impact (mean assessment score)	2.2	2.1	1.9	2.3	2.1

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 15: BDO reviewers' scores for the impact of JSA-funded projects, by region

Region	African Countries	Asia Pacific Countries	Global / Mixed ⁷⁰	JSA overall
Impact (mean assessment score)	2.2	2.1	2.2	2.1

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 16: BDO reviewers' scores for the impact of JSA-funded projects, fragile vs non-fragile states

Classification	Fragile states	Non-fragile states	Global / Mixed ⁷¹	JSA overall
Impact (mean assessment score)	2.2	2.1	2.2	2.1

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

⁶⁹ The Fund has conducted follow-up 'quasi-impact' evaluations of training programs, which indicate that officials who participated in training had utilized and disseminated the knowledge that they acquired. However, these activities do not constitute 'impacts' as per the OECD-DAC definition (see next footnote), and hence could not be taken into account in the scores in **Figures 14-16**.

⁷⁰ Global /Mixed refers to programs such as online training, which could include participants from more than one region.

⁷¹ Global /Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

2.5.2 CHALLENGES TO IDENTIFYING AND REPORTING IMPACTS

Identifying genuine impacts from JSA-funded projects (and this kind of CD *per se*) is challenging because:

- Carrying out evaluations and other assessments shortly after the end of a program (or indeed mid-implementation) means that timescales are often too short for impacts to be visible.
- There is generally no targeted follow-up by the IMF or recipient authorities to identify, measure and report impacts over a longer timeframe.
- Both recipient authorities and IMF staff are sometimes unaware of exact definitions, and hence report as impacts activities or results which are in fact outputs or outcomes.⁷² They also sometimes report *potential* impacts (which intuitively should follow from the CD, but which are not yet evidenced) as actual.
- There are frequently confounding / exogenous factors - e.g., macro-economic trends or the influence of other CD or reforms - making attribution of impacts to a specific project or program difficult.

Both our detailed review of sample programs and our e-surveys sought evidence of impacts attributable to JSA-funded projects and programs. However, for the reasons listed above, these were difficult to find and evidence robustly. That said, we did identify a few reported larger-scale effects which meet the definition of an impact, and can be attributed at least partly to JSA-funded CD projects and programs. For example:

- Continuing collection of customs revenue despite COVID-19 disruption, thanks to improved organizational capacity resulting from the JSA-funded project.
- Enhanced technical and policy analysis capacity at a central bank, leading to timely and clearly communicated monetary policy actions - in turn increasing central bank credibility among financial markets, opinion-makers and the public at large, as well as ability to control inflation based on active response, clear projections and frank risk assessment.
- Increased tax revenue (and hence resources available to the public purse) due to JSA-funded implementation of electronic systems and procedures.
- Increased capacity for bank supervision, including on-site and off-site monitoring, leading to improved risk management and greater financial stability.
- Wider improvements to resource mobilization due to policy advice and skills learnt from JSA TA.
- Strengthened capacity of officials and improved transparency of central bank communications with both public and private sectors - including greater media interest in the annual financial stability review, and a more informed public regarding risks and policy measures to ensure financial stability.
- Fiscal risk evaluations informed by the JSA-funded project were used to review the country's portfolio of state-owned enterprises (SOEs), which led to the privatization of several companies and a restructuring of the financing of public universities.

More generally, identifying and evidencing impacts requires a long timeframe and bespoke follow-up data collection - in some cases many years after the intervention itself. One way for the IMF to improve the availability of such data in future would be to commission or carry out periodic follow-up surveys of recipient authorities. This data collection could also be used to gather more robust and timely data on sustainability (see next section).

2.6 SUSTAINABILITY

2.6.1 OVERVIEW

Sustainability is the extent to which the net benefits of an intervention will continue or are likely to continue.

As **Figure 17** shows, the evidence from our work is that the benefits from JSA-funded TA are fairly sustainable, although - as with development interventions and CD work more broadly - subject to a number of risks. As with

⁷² For example, delivery of training to recipient authority staff is an *output*, increased staff knowledge and capacity is an *outcome*, and any resultant wider benefits (e.g., increased tax revenue, economic growth, improved governance) are *impacts*. However, we frequently found outcomes such as 'improved staff knowledge' being reported in surveys and project documentation as an impact.

impact, the evaluation team had only limited evidence on sustainability after the conclusion of the project. Hence our ratings were based mostly on proxies for the likelihood that benefits would continue - e.g., whether there was strong recipient authority ‘buy-in’, whether knowledge, tools and systems had been absorbed by the recipient organization, and whether staff who participated in the CD were likely to stay in post and transmit the experience to colleagues. ICD conducts participant surveys a few years after the conclusion of training activities, and we also used these data in assessing the sustainability of ICD training programs.

Figure 17: BDO reviewers’ scores for the sustainability of JSA-funded projects, by lead department

Lead Department	FAD	ICD	MCM	STA	JSA overall
Sustainability (mean assessment score)	2.2	2.8	2.5	3.0	2.4

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

The notably higher rating for STA programs is partly because statistics-related CD frequently relates to establishing systems for collecting, analysing and publishing data. Once established - usually supported by manuals and processes embedded in the recipient authority - these systems can be retained and operated year on year with minimal further intervention by the IMF team. By contrast, some FAD and MCM programs - even if they delivered their planned outputs and activities - did not succeed in establishing the structures and systems likely to maintain reforms. In some cases this was because significant institutional or sectoral reform would have been needed to ensure sustainability, but was not achieved - for example, due to political resistance⁷³ or the sheer scale of restructuring required.

Our reviewers noted some marked differences in sustainability of JSA-funded CD between regions. Whilst effectiveness (as measured by our reviewers, i.e., achievement of objectives) is similar between African countries and Asian countries, sustainability appears noticeably lower in African countries. This is mainly due to more severe lack of resources (both human and financial) and staff continuity in African recipient authorities, as well as a greater risk that implementation will be disrupted by political upheaval or regime changes - particularly in fragile states (Figure 18 and Figure 19).

Figure 18: BDO reviewers’ scores for the sustainability of JSA-funded projects, by region

Region	African Countries	Asia Pacific Countries	Global / Mixed ⁷⁴	JSA overall
Sustainability (mean assessment score)	2.3	2.6	2.5	2.4

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Figure 19: BDO reviewers’ scores for the sustainability of JSA-funded projects, fragile vs non-fragile states

Classification	Fragile states	Non-fragile states	Global / Mixed ⁷⁵	JSA overall
Sustainability (mean assessment score)	2.4	2.5	2.5	2.4

Source: BDO review of a sample of 15 JSA-funded programs (1 = poor and 4 = excellent)

Our e-surveys suggest that IMF implementing staff are generally confident that the results of their JSA-funded project will be sustainable. For example, 39%⁷⁶ of IMF managers responding to our survey thought that results would be ‘very sustainable’ over the next five years, and more than a third over the next ten years. That said, our surveys of recipient authorities suggest that they are less optimistic about the sustainability of their results than IMF staff, particularly in the short term (Figure 20).⁷⁷

⁷³ For example, one MCM country-program achieved initial milestones and outcomes related to reforming the regulatory framework, but achieved little in bank restructuring and resolution. In other words, it succeeded in implementing “technical changes” to the regulatory framework, but failed in delivering “hard reforms” which appear to have been at odds with vested interests in the banking and political establishment.

⁷⁴ Global /Mixed refers to programs such as online training, which could include participants from more than one region.

⁷⁵ Global/Mixed refers to programs such as online training and certain regional TA programs, which are not country-specific - and hence could include participants from both fragile and non-fragile states.

⁷⁶ Margin of error at 95% confidence level = ±9%.

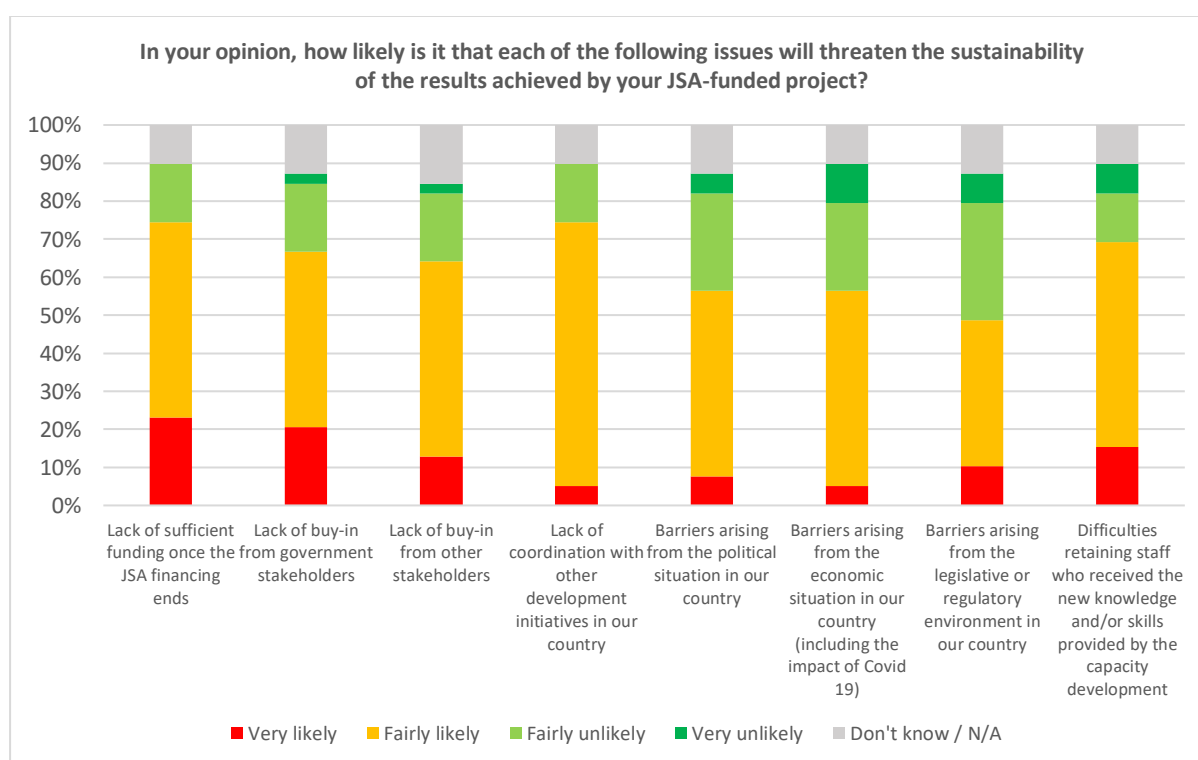
⁷⁷ 33%. Margin of error at 95% confidence level = ±13%.

Figure 20: Views of IMF staff and recipient authorities as to the sustainability of JSA-funded projects

Timeframe	% of respondents who state that the results from the JSA-funded project will be 'very sustainable'	
	IMF Project Managers	Recipient Authorities
In the short term (i.e., up to 1 year)	61%	33%
In the medium term (i.e., 1-5 years)	39%	28%
In the longer term (i.e., 5-10 years)	35%	33%

Source: BDO survey of IMF implementing staff (n=54) and recipient authorities (n=39)

The biggest risks to sustainability reported by recipient authorities include: lack of financial resources to maintain progress once the CD intervention ends; lack of coordination with other development initiatives, and failure to retain staff who have acquired skills and knowledge from the CD (Figure 21). IMF staff reported similar risks to sustainability, although - as per the table above - they were generally more optimistic than recipient authorities that such risks would not threaten sustainability.

Figure 21: Recipient authorities' views on the sustainability of their JSA-funded project

Source: BDO e-survey of recipient authorities (n=39)

More generally, it should be noted that both our reviewers' scores and survey respondents' views are mainly based on the presence or otherwise of *prerequisites* for sustainability, rather than retrospective evidence of whether sustainability was achieved. This is an inevitable limitation of the short timeframes covered by an evaluation such as this, and the lack of available data on what happened after the end of the IMF intervention.⁷⁸ As mentioned in the discussion of impact, to address these constraints, the IMF could carry out or commission follow-up research on the longer-term results of the activities they fund. As a minimum, they could request that recipient authorities report on results periodically after the completion of the project.

⁷⁸ See Section 2.5.

2.7 LESSONS LEARNED

This section presents some key lessons learned from our review of JSA-funded programs, brigaded by relevant OECD-DAC evaluation criteria. While these observations do not represent fully fledged recommendations,⁷⁹ they may be useful for IMF teams to bear in mind when designing and implementing future CD interventions.

2.7.1 ON IMPROVING RELEVANCE

- Sustained and intensive engagement by CD staff is particularly crucial in low-income, low-capacity countries - particularly where there is a risk of poor ‘buy-in’ and ownership. For example, one project in an African fragile state struggled initially with ‘buy-in’. However, intensive engagement by the LTX built strong relationships with authority staff, whose ownership improved as the project progressed.⁸⁰

LTX modalities are the most obvious way of delivering such intensive engagement; however, the same principle can apply to virtual or hybrid delivery as well. For example, consistent and ongoing remote support by an IT expert helped a lower-middle-income African country to establish a new IT system for tax administration - despite capacity limitations at the recipient authority.⁸¹

- IMF lending activities can be a powerful lever for achieving engagement and ownership by authorities - not only because of complementary technical aspects (i.e., coherence) but also because authorities’ existing engagement through mechanisms such as ECF can also be harnessed for CD.⁸² Conversely, loss of engagement from the authority on the lending side can have a knock-on effect on their engagement with IMF CD.⁸³ Both these factors need to be borne in mind by teams in their assumptions and risk assessments regarding authorities’ readiness to engage with proposed TA.

2.7.2 ON IMPROVING COHERENCE

- Various administrative mechanisms and structures can enhance coherence and coordination, particularly in settings where multiple TA interventions and providers are operating at once. For example, in one country the PFM Reform Steering Committee is chaired by the Deputy Prime Minister - which ensures a high degree of coordination, both between the authorities and the CD providers (including the IMF), and among donor partners. In the area of PFM, donor coordination is achieved by periodic consultative meetings, while beneficiary-donor coordination is achieved by regular meetings between the secretariat of the steering committee and donor representatives.

2.7.3 ON IMPROVING EFFICIENCY AND EFFECTIVENESS

- The optimum modality for delivering TA can be determined by both practical issues (such as connectivity and staff capacity) and cultural aspects. For example, during COVID some recipient countries rejected virtual TA on principle,⁸⁴ even where connectivity might have been sufficient to support it. In others, the TA was able to proceed remotely, but with negative impacts on efficiency and effectiveness - not always due to technological issues. In several cases virtual delivery was challenging not because of connectivity, but because of difficulties in building new relationships remotely rather than in-person.⁸⁵
- Within a multi-country program, small-scale TA projects - even if these have minimal expenditure or deliverables - can be a useful way of maintaining relationships with recipient authorities who are currently unwilling or unable to engage with the Fund more substantially. For example, one country-project in a multi-country program was effectively put on hold due to lack of progress; but keeping it open allowed the IMF to gradually resume work once the authorities indicated that they were ready.⁸⁶

⁷⁹ Our full recommendations from the evaluation are presented in **Section 3.2**.

⁸⁰ FAD_AFR_2019_01

⁸¹ FAD_IMF_2021_09

⁸² Cf. various country projects in MCM_AFR_2021_02.

⁸³ E.g. STA_APD_2020_01

⁸⁴ FAD_APD_2020_02

⁸⁵ E.g., FAD_APD_2019_01

⁸⁶ ICD_IMF_2021_05

3. CONCLUSIONS AND RECOMMENDATIONS

This section presents the main conclusions and recommendations from our evaluation. It draws together the main findings from the previous section, and presents higher-level conclusions and lessons on relevance, coherence, effectiveness, efficiency, impact, and sustainability. It concludes with recommendations that could help improve JSA-funded CD, many of which could also be helpful to improve other, non-JSA-funded, IMF CD.

3.1 CONCLUSIONS

Our main conclusions from the evaluation are as follows.

- 1. JSA-funded TA and training constitute a key part of overall IMF CD, and contribute greatly to developing capacity and strengthening institutions in Asia and some African countries, particularly fragile states.**

JSA is the main source of funding for many CD activities including online training and support for JSA priority Southeast Asian countries, namely Cambodia, Lao PDR, Myanmar, and Vietnam. Naturally, given the size of the JSA portfolio, in terms of financing and activities, the results of JSA-funded CD and IMF CD overall are closely interlinked. Still, some differences exist that enhanced the impact of JSA-funded activities - for example, JSA's greater involvement in online and other training, and differences in country coverage.

Through its support for training, JSA contributed to developing online and blended (online, virtual, in-person) activities that have been highly effective. Blended activities also enabled the IMF to continue delivering training and TA during the COVID-19 pandemic.

The evaluation found that JSA-funded CD was generally well aligned with IMF surveillance and lending, as per the aims of the Fund's 2018 strategy. There is scope to further enhance the relevance, effectiveness and impact of JSA-funded programs - for example, by building on the progress made with new delivery modalities, and tailoring activities more closely to each country's needs and absorption capacity.⁸⁷

- 2. The period under review was dominated by the pandemic, and the Fund and JSA deserve credit for adapting to the circumstances.**

The pandemic restrictions halted all CD activities overnight, and they resumed only gradually - more successfully in some countries than in others. The Government of Japan recognized the need for flexibility to adapt JSA-funded programs to the difficult circumstances and accommodated the necessary changes. The Fund was able to replace in-person training with virtual and online courses, and in-person TA with virtual missions, as well as launching virtual missions for surveillance and programs. These new modalities contributed to integration of TA and training with surveillance - one of the Fund's strategic goals.

The pandemic led to innovations in the design and delivery of training that helped maintain effectiveness and impact, while reducing average costs.⁸⁸ A few months into the pandemic, CD delivery (particularly training) started recovering, thanks to innovations by IMF staff, and a significant increase in demand. Many programs were redesigned; some were cancelled, and new initiatives were launched. Initially, with JSA support, training shifted completely to virtual delivery and online courses, as ICD scaled up its online program in response to a large increase in demand. Overall, there was a significant reduction in the average cost per participant while maintaining or improving participant satisfaction and learning outcomes - indications of effectiveness.⁸⁹ User surveys on utilization and institutional absorption of knowledge also indicate that the longer-term benefits of training (such as continued use of knowledge by staff in a relevant role) were not affected.

TA recovered more gradually, first through "virtual missions," and later through sustained engagement of resident LTXs, some of whom were based at RCDCs. In many countries, by the end of 2021 TA delivery had reached pre-pandemic levels. The pandemic served to accelerate the integration of online training, synchronous virtual missions, and in-person deliveries by resident LTXs. In several instances, we observed

⁸⁷ See Section 2.2.

⁸⁸ See Section 2.4.1.

⁸⁹ See Section 2.3.

that TA recipients were asked to take an online course before in-person TA missions, and on other occasions an in-person mission was followed up by virtual consultations.⁹⁰ Together, these modalities proved as effective and less costly than “traditional” one-off missions by short-term experts (STXs) or regional LTXs. The past three years have provided rich experience on how technology can enhance effectiveness and reduce costs.

3. Overall, JSA-funded programs score highly for relevance and coherence.

The majority of JSA-funded programs were relevant, in that they addressed important capacity weaknesses in the recipient countries, and were generally well aligned with priorities identified in IMF surveillance and program work. Similarly, coherence was rated as “good”, as CD teams made an effort to coordinate across agencies in recipient countries, and tried to build synergies between different IMF workstreams. In most, but not all, cases, project managers engaged with other donors to avoid duplication of efforts, and in some cases to partner with them to achieve better results on the ground. Increasingly, RCDCs are enabling better synergy across CD programs as well as donor coordination, leading to greater relevance and coherence of CD programs.⁹¹ These efforts have led to more relevant and coherent JSA-funded programs. The average of the programs for relevance and coherence overall was around 3.0 (on a scale from 1 to 4).

4. There is scope to improve awareness among recipient authorities and IMF Mission Chiefs / Resident Representatives that JSA activities are financed by the Government of Japan.

In general, we found that awareness of the funding source for JSA CD amongst IMF project managers was very good. But there was scope for improvement among both IMF Mission Chiefs / Resident Representatives and recipient authorities. More than a third of recipient authorities responding to our survey were not aware that their JSA project was funded by the Government of Japan. Of those who were aware, two thirds had learned this through discussions with IMF staff rather than ‘branding’ or other sources. Hence there may be scope to improve visibility by increasing JSA ‘branding’ on project documents and - equally importantly - ensuring that IMF staff make clear to their counterparts at recipient authorities who is funding the CD.

5. Most JSA-funded programs were relatively effective, in that they largely or at least partially achieved their milestones, outcomes, and objectives. But there were significant differences across sectors and countries.

The findings from document reviews, RBM indicators, interviews, and surveys are consistent in showing that effectiveness has recovered since the height of the pandemic, and that projects are generally meeting their milestones and intended outcomes - albeit with some delays because of the COVID-19 pandemic.

There were significant differences in our assessments of effectiveness across IMF CD departments, which are mainly due to differences in implementation - given that relevance and coherence were much more similar. ICD and STA have the highest average, with 3.3 and 2.7 respectively. The average ratings for FAD and MCM were 2.3 and 2.5 respectively. These differences are largely to be expected, given the greater complexity of FAD and MCM projects that involve deep and lengthy institutional reforms. There is little difference between Fragile States (mostly African) in the RBM ratings, which are based not on the relative ambition of the milestones and outcomes, but purely on whether or not they are achieved.

The main drivers of effectiveness were:

- country ownership and commitment to reform (proxied by relevance, and in particular by who selected the particular project and the authorities’ contribution to the project design),
- the initial conditions in the recipient country and agency (countries with better institutions and agencies with more resources used JSA-funded CD most effectively).
- better internal coordination within the Fund, between the delivering CD department and the corresponding country team (working on surveillance or lending programs), which usually leads to better designed and more effective programs.

In general, program and project complexity diminished effectiveness, as it taxed both the authorities’ and staff’s implementation capacity.

⁹⁰ See Section 2.3.

⁹¹ See *The IMF and Capacity Development*, IEO Evaluation Report (2022), p. 2 (<https://ieo.imf.org/-/media/IEO/Files/evaluations/completed/10-04-2022-the-imf-and-capacity-development/cd-chapeau-web-v2.ashx>).

- 6. Caution must be exercised in using the RBM framework to compare the effectiveness of one program with another, and more so between different countries. This is particularly the case when assessing the effectiveness of programs in fragile states.**

The proxies for effectiveness in the RBM framework, i.e., objectives, outcomes, and milestones, are set relative to the expectations of what the project could achieve in the corresponding context, and differ depending on the country's and agency's initial capacity. Even within the same program, what constitutes a good outcome for a fragile state may be considered as not achieved for an emerging market economy.

The evaluation period has seen significant progress in the use of the RBM framework for planning and managing projects - including better defined milestones, outputs and outcomes, and improved compliance with entering the required information. However, gaps remain, and there is still sometimes confusion between inputs, outputs, and outcomes. We found that the RBM database remains limited as a tool for assessing each program's progress against milestones, outputs, outcomes and objectives. In part these limitations are inherent in the RBM framework, as milestones, outputs and outcomes are adjusted during project implementation to reflect the realities on the ground.

- 7. Multi-country programs have been a flexible and often efficient way to deliver CD, which can sometimes deliver returns to scale. However, they can also complicate coordination within the Fund, and can pose challenges to effectiveness, transparency and accountability.**

Multi-country programs allow the IMF to launch CD country-level projects without having to invest in advance the same amount of time and resources that would be required for a single-country program. This is because they do away with the need to ascertain in advance authorities' full commitment to implement reforms in each and every recipient country. Once the program is in place, the Fund can shift resources from one recipient country to another in response to changing priorities and circumstances.

The employment of an LTX is more cost-effective in a multi-country program than in a single-country program, as a single LTX can be shared by multiple countries. As an additional advantage, a multi-country program, by bringing a diversity of countries together, can provide opportunities for peer learning.

On the other hand, a multi-country program sometimes makes it difficult to identify and hire appropriate experts if the recipient countries are very heterogeneous, and makes program management more difficult. Our evaluation found cases where a country manager could not respond to a request because of uncertainty about the availability of resources even though the overall program budget had an underrun. This is in part because until recently there has been insufficient information on the allocation of budgeted resources across the different country-projects (a problem that is expected to diminish with greater use of CDMAP and all its functionalities).

Finally, multi-country programs make it difficult to evaluate completed programs when some countries have achieved results while others have not.

- 8. Assessing the efficiency and VFM of JSA-funded activities is difficult, due to the lack of complete cost data and difficulty in monetizing outcomes. However, qualitative evidence suggests that VFM has improved since the pandemic - as new delivery modalities are less costly and appear equally effective.**

Estimating the cost of IMF CD during the evaluation period was difficult because the systems for budgeting and monitoring costs were in transition to CDMAP. This was further complicated by the effects of COVID on costs and volume of deliveries, and the need to change delivery modalities. In 2020 and 2021 there were significant budget underruns in most programs, as travel expenses collapsed. The pandemic forced the IMF to scale up the use of technology, and most activities were delivered through blended (online/virtual/in-person) modalities. Thus, once deliveries restarted, the average cost of activities was reduced because of lower travel costs. Our surveys and interviews indicate that staff and some recipients see benefits in these blended modalities beyond cost reduction, as they allow more continuous interactions between recipients and Fund experts.

- 9. There is limited evidence that JSA-funded programs are having wider impacts beyond their initial outputs and outcomes. However, this is largely due to short timeframes, the fact that data is not routinely collected after project completion, and difficulties with identifying and attributing possible effects to the JSA-funded intervention.**

Although we found a few individual examples of larger-scale impacts attributable to JSA-funded CD, programs generally scored fairly low for impact (average across all projects 2.1). As with donor-funded

interventions and CD more generally, identifying and evidencing impacts requires a longer timeframe and bespoke follow-up data collection - in some cases many years after the intervention itself. There may be scope for the IMF to organise periodic follow-up surveys of recipient authorities to obtain longer-term evidence on impacts (and indeed sustainability - see point 10 below).

10. Sustainability of JSA-funded programs appears relatively modest, although better in Asian emerging economies than in low-income fragile states in Africa.

Overall, we gave an average rating for sustainability of 2.4 (out of 4) across all programs reviewed. There was a significant difference between programs in the Asia and Pacific region (average 2.6) and Africa (average 2.3), which reflects the challenges faced by low-income and fragile states in maintaining staff and resources, as well as political and administrative continuity, once the JSA-funded program ends.

Our assessment of sustainability draws on information available at the completion of a project (or of the latest year in the case of an ongoing program). Hence it is based on the presence (or otherwise) of factors likely to support sustainability, rather than retrospective evidence of what actually happened. We gave high sustainability ratings to activities where there was strong recipient authority 'buy-in', where knowledge, tools and systems had been absorbed by the corresponding organization, and where staff who had participated in the TA activity were expected to stay in the position and transmit the experience to colleagues. This was much more common in emerging economies than in low-income and fragile states.

11. We do not consider that it would be practical to try to optimize VFM by identifying programs for early cancellation and reallocation of funds.

Currently, the vast majority of programs and country projects are relatively short. Therefore, it would not be practical to attempt to cancel them ahead of time, considering the processes and documentation that would be needed in both the Fund and JSA. In most cases, the Fund and JSA would need to prepare a thorough document to explain the cancellation to the recipient authority - the annual interim assessment currently being produced would not be adequate for that task. Our reviewers found that the amount and quality of information provided in interim assessment reports was frequently limited, and would often not have supported a robust decision to cancel or reorientate the project at that stage. At the country project level, multi-country programs allow projects that are not working to be effectively cancelled, without necessarily severing engagement with the country concerned.

3.2 RECOMMENDATIONS

JSA has supported the IMF in expanding the reach and impact of its CD, which is recognized as best in class on issues at the core of its mandate and expertise. JSA-funded activities have been effective at building capacity in many countries, especially in Asia and parts of Africa.

Notwithstanding these achievements, there are areas where there is scope to further improve the performance of the JSA and its programs. We therefore make the following recommendations to the IMF and/or the JSA.

1. The Fund should maintain or adopt important innovations and adaptations catalysed by the pandemic - notably those relating to technology and remote delivery - across all relevant activities. (Priority: High)

For example, the pandemic heightened the role of RCDCs to enable in-person and virtual TA support on a regular basis. With the possible exception of fragile states, JSA should consider prioritizing virtual and blended in-person/virtual activities to build capacity over one-off missions. Brief missions may be a good method to prepare a diagnostic. But implementation generally requires regular follow-up, for which virtual modalities are generally better suited, because of costs as well as coordination. The IMF should continue to favour in-person LTX engagement as a crucial delivery model for fragile states.

2. Regarding training, the Fund should seek to maintain the progress made in using online and virtual courses rather than reverting fully to former modes of delivery. (Priority: High)

The Fund should consider shifting to online formats (and particularly to MOOCs) for imparting basic general knowledge, including on policy issues such as financial programming. It should also consider shifting to virtual mode of CD delivery for upper-middle-income countries, and for lower-middle-income countries with good connectivity. Virtual and in-person training is more effective and much more likely to lead to institutional strengthening when delivered to teams working on the corresponding issues in a single country, rather than when delivered off-site to audiences from many countries. Such training facilitates integration

with TA, and makes for a less strict distinction between training and TA. Regional training has its own benefits in permitting peer learning and helping to build networks.

3. The Fund should seek more and earlier inputs from recipient authorities in deciding on a program, its design, and on the selection of an LTX. (Priority: High)

Before approving a project, JSA could require that proposal documentation include a brief explicit official request by recipient authorities, to confirm their buy-in to supporting delivery. This documentation should also include confirmation from IMF staff that other donors and implementing organizations have been consulted to identify or rule out opportunities for synergies or joint working.

The selection of LTXs should be based not only on their technical expertise, but also on the rapport that they are able to build with the recipient authorities. For this reason, it might be helpful to give the authorities a greater say in the drafting of the LTX's TOR, if not in the selection of the LTX him or herself. Similarly, while not current policy, JSA could require that recipient authorities from middle-income countries requesting support from an LTX contribute to its funding, possibly by providing housing and other local costs. This would provide an indication of their commitment, as well as defraying part of the cost of this most expensive mode of TA delivery.

4. JSA should consider reallocating some of the funding from Asian emerging market economies to other regions, including African countries, which have greater needs for capacity building. (Priority: High)

Many recipient countries in Asia have attained middle and even upper-middle-income status. Their human capital and state capacity have therefore improved tremendously, diminishing the value added of CD, and creating an opportunity to reassign resources to countries in greater need. This could be done without sacrificing the actual scale of activities in Asian countries, if the Fund were to deliver more CD virtually in countries with good connectivity. Absorption capacity in low-income and fragile states remains an issue. But our work suggests that reallocating some of the resources currently spent in Asia to African fragile states may be a risk worth taking. The Fund and JSA could consider creating a joint IMF-JSA facility similar to the Infrastructure Governance Facility (IGF), to allow dedicated resources for CD in fragile states.

5. The Fund and JSA should explore ways to structure multi-country programs that maintain their flexibility, while also supporting transparency and accountability, and reducing managerial challenges. (Priority: Medium)

For example, it may be helpful to limit the number of country projects within a single program, and ensure that recipient countries are relatively similar in their profile and need for the relevant intervention. More generally, some indication of how resources are allocated across different country-projects should be given at the inception of a program, and there should be greater accountability for allocation of resources over the life of the program. Restrictions could be considered on coverage (e.g., not combining projects across regions, or countries with very different initial capacity). An indicative budget should be set for each country at the outset, and explicit criteria could be established for when to reallocate resources from one country to another.

6. The Fund should make greater efforts to raise awareness among recipient authorities and IMF Mission Chiefs / Resident Representatives that the funding source for JSA CD is the Government of Japan. (Priority: Medium)

For example, as part of protocols for engaging with recipient authorities and in-country IMF colleagues, IMF implementation staff could be required to highlight the source of the funding. This approach could be enhanced by ensuring that JSA 'branding' is present not just on final project/program reports, but also on proposals and planning documentation. The IMF should also consider holding specific workshops or sessions during Annual Meetings, showcasing JSA work to Fund staff and recipient countries.

7. The Fund should consider ways to gather evidence routinely on the *de facto* sustainability and impacts achieved by JSA-funded programs after implementation. (Priority: Low)

Ideally, the Fund should carry out periodic follow-up surveys with recipient authorities to gather information on post-completion results - for example through an e-survey similar to the one employed by the current evaluation. A more modest approach would be to require authorities to track and provide a statement on the results of each project periodically (for at least a year) after their completion. In any case, as part of program documentation, the Fund should ensure that it has up-to-date contact information for the counterpart person or unit in each country, both during and after implementation.



FOR MORE INFORMATION:

CAROLINE HAGMANN

Caroline.hagmann@bdo.co.uk

This report has been produced at the request of the International Monetary Fund (IMF). It is intended solely for the use of the IMF for the purposes stated herein. We do not accept or assume responsibility to any third party to whom the Report is shown or into whose hands it may come. If others choose to rely on the contents of the Report, they do so entirely at their own risk.

BDO LLP, a UK limited liability partnership registered in England and Wales under number OC305127, is a member of BDO International Limited, a UK company limited by guarantee, and forms part of the international BDO network of independent member firms. A list of members' names is open to inspection at our registered office, 55 Baker Street, London W1U 7EU. BDO LLP is authorised and regulated by the Financial Conduct Authority to conduct investment business.

BDO is the brand name of the BDO network and for each of the BDO Member Firms.

BDO Northern Ireland, a partnership formed in and under the laws of Northern Ireland, is licensed to operate within the international BDO network of independent member firms.

© March 2024 BDO LLP. All rights reserved.

www.bdo.co.uk

