GFA CONSULTING GROUP

FINAL REPORT

MULTI-COUNTRY

Thematic Pilot Evaluation of The Africa Regional Technical Assistance Center West 2 (AFW2), the Regional Technical Assistance Center for Central America, Panama and the Dominican Republic (CAPTAC-DR) and the Africa Training Institute (ATI)

Prepared by: Soren Kirk Jensen

Presented to:
INTERNATIONAL MONETARY FUND
700 19TH STREET N.W.
WASHINGTON D.C.
USA

AUGUST 2023

Your contact person
within GFA Consulting Group GmbH is
Halima Abdi

Thematic Pilot Evaluation of The Africa Regional Technical Assistance Center West 2 (AFW2), the Regional Technical Assistance Center for Central America, Panama and the Dominican Republic (CAPTAC-DR) and the Africa Training Institute (ATI)

MULTI-COUNTRY

Final report

Address

GFA Consulting Group GmbH
Eulenkrugstraße 82
DE-22359 Hamburg
Germany

Phone +49 40 6 03 06 – 270

Fax +49 40 6 03 06 – 199

E-mail halima.abdi@gfa-group.de

TABLE OF CONTENTS

LIST OF TABLES

LIST OF BOXES

ACRONYMS

AD	Area Department
AFR	Africa Department
AFRITAC	Africa Regional Technical Assistance Center
AFW2	Africa Regional Technical Assistance Center West 2
ATI	Africa Training Institute
BIC	Border and Inland Control
CAPTAC-DR	Regional Technical Assistance Center for Central America, Panama and the Dominican Republic
CD	Capacity Development
CDMAP	Capacity Development Management and Administration Program
COMIECO	Consejo de Ministros de Integración Económica
CSO	Civil Society Organizations
CR	Costa Rica
CRM	Customs/Compliance Risk Management
CSN	Country Strategy Notes (CD)
DR	Dominican Republic
DRM	Domestic Revenue Mobilization
ECF	Extended Credit Facility
ECOWAS	Economic Community of West African States
EFF	Extended Fund Facility
ES	El Salvador
EU	European Union
FAD	Fiscal Affairs Department
FCDO	Foreign, Commonwealth & Development Office
FGD	Focus Group Discussion
FY	Financial Year
GA	The Gambia
GDP	Gross Domestic Product
GH	Ghana
GIZ	German Agency for International Cooperation
GRA	Ghana Revenue Authority
HQ	Head Quarter
HR	Human Resources
IADB	Inter-American Development Bank
ICD	Institute for Capacity Development
ICT	Information and Communications Technology
IMF	International Monetary Fund
ISOCA	International Survey on Customs Administration
ITAS	Integrated Tax Management System
LEG	Legal Department
LTX	Long-Term Expert
OECD-DAC	Organization for Economic Co-operation and Development - Development Assistance Committee
PCA	Post-clearance Audit
PFM	Public Finance Management
RA	Resident Advisor
RBM	Results-Based Management
RCDCs	Regional Capacity Development Centers
RMTF	Revenue Mobilization Thematic Fund
Res Rep	Resident representative
RSN	Regional strategy notes
RTACs	which includes Regional Technical Assistance Centers
RTC	Regional Training Centers
SC	Steering Committee
SIECA	Secretariat for the Economic Integration of Central America

SL	Sierra Leone
STX	Short-term Expert
TA	Technical Assistance
TADAT	Tax Administration Diagnostic Assessment Tool
UN	United Nations
USAID	United States Agency for International Development
VAT	Value Added Tax
WAIFEM	West African Institute for Financial and Economic Management
WCO	World Customs Organization
WTO	World Trade Organization
WHD	Western Hemisphere Department

EXECUTIVE SUMMARY AND RECOMMENDATIONS

Revenue administration is of critical importance to IMF member countries and a key function of IMF CD consists in strengthening this effort across tax and customs. This evaluation assessed the extent to which the three IMF Regional Capacity Development Centers contributed successfully to this effort in terms of its relevance, coherence, effectiveness, efficiency, impact and sustainability. This was done primarily by evaluating projects and courses in the relevant workstream(s) but also assessing the overall performance of the three Centers.

The evaluation covers FY20 to FY22 (May 2019 – April 2022), a period that was marked by the global Covid-19 pandemic which altered the way CD was delivered and also greatly affected the recipient government entities. This affected performance across the areas assessed with the evaluation criteria mentioned above. While challenging CD delivery and monitoring, nevertheless, the pandemic also opened up new avenues for moving towards virtual and hybrid approaches to CD.

The evaluation found that CD to revenue administration had generally made a positive contribution in the sampled case countries with room for improvement in some areas. The sampled TA projects and training courses were broadly relevant and coherent, which was reflected in the reasonable level of effectiveness and cost-efficiency in terms of achieving outcomes that were above average already halfway through the evaluation period. There were good cases of potentially strong impact in terms of improved revenue administration and even increased revenue collection. Some of the case countries also stood out in terms of potential for impact over time. Successful efforts were driven by key factors such as political will and commitment to reform, consistency of CD delivery over time and ability to demonstrate tangible results to key decision-makers. The sustainability of the impact will depend on issues such as navigating absorption capacity constraints, extent to which critical capacity gaps within the institutions supported are mapped out and targeted, existing mechanism for broadening ownership of reforms and institutionalizing stronger follow-up on training and TA.

Across the areas of relevance and coherence many good practices and significant efforts were identified by the evaluation as detailed in this report. Adaptability, flexibility and responsiveness can be used to capture the overall effort during this turbulent period. Not only the changes imposed by the pandemic, which was met swiftly through significant changes in work processes and the format of delivery, but also political transitions, shifts in commitment and closing and opening of windows of opportunities were generally met in an apt manner. Significant efforts were made to ensure coherence in a field with many actors and stakeholders both internally and externally although sometimes falling short of meeting the expectations of some development partners.

Findings indicated a few concrete areas of relevance and coherence that require further improvements; including making full use of excellent TADAT based diagnostics, the need for a diagnostic framework similar to TADAT for customs and mapping of critical capacity gaps at the personnel level within recipient government agencies, formalization of strategic planning of CD at country level through regularly updated Country CD Strategy Notes (CSNs) to strengthen synergies between training and TA, ensuring systematic engagement of resident representatives in planning of training in particular and more proactive, in-depth engagement by long-term experts for tax and customs administration to strengthen coordination of CD to revenue administration.

The effectiveness of CD at project and workstream level measured by outcome indicators was rated between partially and largely achieved with very limited variation between the RBM ratings and evaluation ratings. This indicates good progress but also outstanding work to be completed in the current phases. At country level there was slightly more variation, but this evened out indicating that at workstream level, CD to revenue administration works and that the RCDCs delivering this CD are on track to deliver strong outcomes by the end of the evaluation period.

Two areas of concern were identified concerning the RBM approach which included the need for more consistent reporting both in terms of how and when to report on results and in terms of strengthening the requirement for evidence of attribution through the results chain from inputs to outcomes. More importantly than the quantitative ratings, evidence from the various case countries provided numerous detailed examples of CD to revenue administration that worked well as well as examples of areas that could be improved.

CD that had been effective at case country level included sector-based tax audits in AFW2 countries, support to the establishment of an internal affairs units (aligned with a structural benchmark under the IMF program), compliance risk management in one of the CAPTAC-DR countries where there was a strong correlation between diagnostic work (from a recent TADAT), a structural benchmark and TA. In customs, particularly good progress was made in the area of risk-based strategic management and customs clearance and trade facilitation was highly effective. In terms of training, particularly the gender budgeting course stood out as having been effective in line with the RBM framework outcome indicators and was also a standout example in terms of synergies between training and TA.

In some projects, activities were delayed or fell behind schedule, partly as a consequence of the difficulties imposed by the shift to virtual delivery. Focus areas, such as compliance risk management, had more traction in some countries than others in the CAPTAC-DR region. While generally the RCDCs were responsive, there was one case where the focus of support could have changed more swiftly to respond to the priorities of the authorities. Consistency in the delivery of CD was of central importance and reinforced by the pandemic where building trust became ever more important, and the evaluation identified at least one case where too many different short-term experts involved without adequate use of tools to ensure the consistency of the TA hampered progress. Overall, training and TA were equally effective at country level but more structured content training in revenue administration delivered in synergy with TA in the AFW2 region would be beneficial especially if it was targeted according to specific capacity gaps and sustained through follow-up. CAPTAC-DR delivered substantial amounts of training in revenue administration with good results.

Overall, revenue administration lends itself to more in-depth cost-effectiveness and impact assessments as results are often quantifiable but achieving this requires more work on the application of the results-based management approach across the different RCDCs. Based on a novel and explorative approach based on recent CDMAP data, the evaluation produced findings concerning the cost of CD delivered, cost-efficiency and cost-effectiveness. At the most basic level, this approach indicated in two cases that data was either missing or reflecting significant outliers in comparison with the cost of CD in other cases. The approach showed cost of CD with and without backstopping for example indicating that the more structured approach to CD to customs administration in AFW2 appeared more cost-efficient than in CAPTAC-DR. Cost of courses also varied and while virtual delivery was clearly more cost-efficient due to the absence of substantial travel costs. Evidence from one course with a high number of participants was also quite effective, which suggested that a high number of participants in a virtually delivered course did not necessarily limit its effectiveness and also made this course particularly cost-effective. This type of analysis however depends on more mature CDMAP data in order to produce more robust assessments and even then should be interpreted very carefully and always be complemented by careful qualitative analysis of project performance along the lines of this evaluation.

The potential for impact from CD to revenue administration is high as work progresses towards the end of the current phases and beyond. The evaluation identified five thematic impact stories based on promising courses and projects. Generally, but not always these corresponded with the effectiveness ratings. Moreover, three case countries were highlighted as having stronger potential for impact over the next years. These are detailed in the report.

In addition to their delivery of CD to revenue administration the evaluation also assessed the performance of the three RCDCs more broadly in accordance with nine specific areas of interest.

Generally, the Centers did very well across these areas with particular emphasis on the way they responded to the situation imposed by the pandemic. In the midst of the pandemic, they also managed to facilitate the oversight function of their steering committees, retain a decent mix of modalities even if most delivery naturally shifted to being duty station-based TA. The Centers also performed proactive outreach and fundraising, and broadly secured funding for the current phase through additional contributions and savings from the shift to virtual delivery, introducing new priority areas to the CD offer, allocating resources reasonably among members with prioritization of fragile states (Africa) and members with strong demand for CD (CAPTAC-DR). Centers need to continue their efforts to diversify their expert base and find ways to engage more local and Portuguese speaking experts as well as ways to mitigate uncertainties associated with being significantly dependent on external fundraising. It will also be important to keep refining the use of virtual and hybrid CD and blended learning based on the lessons from this evaluation and other Centers. Finally, the Centers were unanimously commended for their responsiveness and commitment to continue delivery of high-quality CD throughout the pandemic.

Lessons and recommendations

This pilot thematic evaluation has identified a ten lessons and associated recommendations which are listed in section 3 of this report. The following four lessons and recommendations are the most critical ones that have emerged from this evaluation and the hope is that they will serve to inspire not only CD delivered by the three RCDCs under evaluation but also IMF CD more broadly through the strategic review.

Lesson 1: As a diagnostic tool to inform CD to revenue administration, TADAT has great value which should be further exploited to inform TA. It is important that the Centers and authorities are able to draw on up to date TADAT diagnostics that could be prepared either as full repeat TADATs or as self-assessments that are supported by Center experts for example by contracting experienced TADAT assessors as STXs to support such processes. As TADAT does not cover customs, the availability of good diagnostics to inform revenue administration can be skewed in spite of other diagnostics being used to fill this gap.

- **Recommendation**: RTACs, FAD and the ADs to support authorities in keeping TADATs and follow up plans up to date either through repeat TADATs or by supporting and quality assuring TADAT self-assessments. In addition, FAD should continue working with WCO on rolling out ISOCA as a diagnostic tool for customs at country level.

Lesson 2: Country CD strategy notes (CSNs) do not fulfil their intended purpose of guiding CD delivery strategically at country level as they are often outdated and sometimes fail to articulate how structured content training should complement TA at country level. Addressing this would be an important first step towards stronger synergies between structured content training and TA (duty station, in-person and virtually delivered), which is currently lacking in AFW2 and ATI case countries.

- **Recommendation:** Area departments should ensure that CSNs are regularly updated and work with RCDCs to ensure that CSNs set out clear strategies for synergies between training and TA at country level. This should address how structured content courses delivered through ATI can be used to close critical capacity gaps in targeted agencies that also receive TA. This approach should be piloted for revenue administration CD in all countries that currently have outdated CSNs.

Lesson 5: Room for improvement in the use of the RBM approach to CD remains and need to be addressed urgently as CDMAP and the possibility of undertaking cost-efficiency and cost-effectiveness analysis to inform prioritization, planning and delivery of CD increases. Milestone and outcome indicator ratings (both for TA and training) need to be reported consistently to CDMAP and across other means of communication, monitored and ideally outcome indicator ratings should be subject to backstopping by ICD to ensure consistency. There is also a need to invest in strengthening the methodology to rate achievement of milestones and indicators ensuring that these are attributable to the activities supported by the project throughout the results chain (linking inputs, outputs, outcomes and objectives) as well as considering their relative importance to the broader objective. This will help demonstrate impact and value for money for external and IMF own funding of CD.

- **Recommendation:** ICD to develop a step-by-step methodology for ensuring clear and credible results chains to demonstrate attribution of achievements to CD delivery. This could be done by linking inputs (CD days, missions), outputs (milestones), outcomes (outcome indicators) and objectives in a convincing and consistent manner in the RBM, annual reports and TA reports. When possible, complement internal analysis with repeat TADAT assessments regarding results claimed at outcome indicator level.

Lesson 7: The expansion of CDMAP offers great opportunities for using cost-efficiency and cost-effectiveness analysis systematically to TA and training in revenue administration. The novel and experimental analysis carried out as part of this evaluation provided some initial findings indicating outliers in the cost of a CD day in some countries and demonstrated future potential for expanded and more robust use of such tools. However, the credibility of such analysis is inherently linked with assuring that classification of resources is done accurately (and in line with the new classification of CD modalities) as well as the required strengthening of the RBM approach (generally and particularly concerning the relative importance of indicators to the achievement at the objective level) and should always be used with caution as a complementary tool to qualitative analysis.

- **Recommendation:** AFW2 and ICD should look into the costs of the tax projects in two out of the three case study countries (in one of the countries, all costs need to be reviewed while in the other backstopping costs needs to be reviewed) in order to check for reporting and classification issues or other explanatory factors for the high cost per CD day identified. Based on this, AFW2 and ICD should consider developing an application in CDMAP that automatically flags outliers for checking and correction.

1 INTRODUCTION

The purpose of the Thematic Pilot Evaluation of the Africa Regional Technical Assistance Center West 2 (AFW2, phase II), the Regional Technical Assistance Center for Central America, Panama and the Dominican Republic (CAPTAC-DR, phase III) and the Africa Training Institute (ATI, phase II) is to evaluate the quality of the IMF´s capacity development (CD) efforts in revenue administration; tax and customs projects as well as relevant courses delivered by the three centers. It is a mid-term evaluation covering the first three years (FY20-FY22) of the current phases of each of the Centers CD interventions. The evaluation had three key objectives i) using the OECD-DAC criteria, assess the relevance, coherence, effectiveness, efficiency, impact and sustainability of TA and training within the revenue administration workstream (part I of this report), ii) based on key guiding questions, conduct an entity level analysis for each of the three Centers to assess their strategy and operational management in achieving their mandate (part II of this report), iii) identity findings and recommendations that can inform future CD and phases. The final report builds on extensive work carried out over a period of more than a year based on the approach outlined in the evaluation inception note, which was approved in December 2022. The report draws on nine case studies and three entity level evaluations, the summaries of which are available in Volume II of the report.

Methodological approach

This is the first time the IMF has used a thematic approach to an external evaluation of its RCDCs CD efforts. It is also the first time the same methodological approach has been used across the different types of RCDCs (RTACs and RTC). While this new approach provided some interesting results, it was not without challenges (see the limitation and boundaries section below).

To inform the assessment across the DAC criteria, extensive amounts of documents and materials has been reviewed, including hundreds of TA reports, briefing papers, RCDC annual reports, IMF program documents and staff reviews, Article IV staff reports, Regional CD Strategy Notes, Country CD strategy notes, TADAT reports and other diagnostic assessments, country and revenue administration strategies etc. Data from CDMAP and the IMFs RBM system has been carefully reviewed. To ensure consistency of the evaluation, entries of RBM ratings into the central RBM database was used to inform the effectiveness analysis and validation of ratings and performance but most emphasis was put on the qualitative analysis identifying activities that worked well as well as those that had room for improvement. Ratings were validated against all available evidence, including reporting in annual reports against milestones associated with specific activities and focus areas under the project objectives, using basic averages to determining the exact quantitative score.[1] FY22 cost data in dollar terms was used to inform cost-efficiency at country and comparative level which provided some interesting basic findings but requires further thinking and development and will be more useful in the future as CDMAP matures (with data becoming available for consecutive years) and CD settles into a more consolidated post-pandemic mode.[2]

The evaluation relied on a hybrid approach to field work with in-person missions to five of the seven case countries, while fieldwork for the Gambia and Côte d'Ivoire was done remotely. Of the five visited countries, Ghana and Sierra Leone were sampled as members of both AFW2 and ATI. Entity level evaluations benefited from in-person visits to the host countries in Ghana and Mauritius, while Guatemala was covered remotely. The in-person field visits were short and relied on semi-structured, key informant interviews with relevant country authorities and SC members, IMF staff including RCDC personnel and development partners. Evidence from course participants was collected primarily through focus group discussions (FGD). Several follow-up

[1] There is a case for using a more sophisticated approach to quantitatively rating outcomes in the future, for example through weighted approaches, but this would have to be clearly determined at the inception stage.

[2] For many reasons, FY22 was not a representative year, neither for the three fiscal years under implementation, nor for the provision of CD in general. As noted the approach was chosen to use the most recent available format of data and to be forward looking.

interviews were conducted subsequently, for example with course lecturers, long-term experts (LTXs), short-term experts (STXs) and headquarter-based (HQ) personnel. In total this resulted in more than 120 interviews and FGDs. The validity of the ATI findings was also discussed extensively through a substantial series of interviews with current and former lecturers as well as senior staff.

A survey was also conducted. While a large number of respondents were targeted (2247), feedback was received from only 123 individuals (27 IMF personnel, 10 Steering committee and members as well as 86 beneficiaries from all the three regions), translating to a 5 percent response rate. Given this limited response rate, the validity and usefulness of the survey was limited. However, specific qualitative information and comments have been used to corroborate findings from other sources.

Information from the various sources were cross-referenced and drafted into preliminary findings and recommendations as well as an extensive annotated outline of the report, which was shared with the IMF in June 2023. In addition, presentations were made to two of three Steering Committee meetings where emerging findings and potential themes for recommendations were presented. The evaluation team received valuable feedback from these events as well as on the annotated outline and preliminary findings and recommendations. A final review phase included revisions and verification of data and evidence based on feedback from IMF stakeholders.

The evaluation report and case studies are presented in a cascaded manner where different layers of granularity in the analysis used to arrive at this overarching report was produced in various underlying working documents and spreadsheets. The nine case studies and three entity level reports are included as abbreviated versions in Volume II.

Evaluation boundaries and limitations

1. While thematically focused, the scope of the evaluation was broad and required coverage ambitious. Including three RCDCs, using different CD modalities, and seven highly divergent case countries as well as the use of the OECD DAC criteria shaped the evaluation and its thematic depth. There were strict limitations on the length of the two resulting reports (volume I and II) which clashed with the understandable desire for numerous aspects to be covered and details added by the entities evaluated. Many decisions and prioritizations had to be made throughout the long evaluation process and the resulting report reflects this balance.
2. Integrating ATI with the two RTACs in the same evaluation implied that the same evaluative approach and methodology had to be followed based primarily on case studies. This resulted in ATI courses being evaluated primarily at the country level, which is unusual but corresponds to the requirements and design of the evaluation as outlined in the inception note. The evaluators recognized this is a novel approach but find that it has produced interesting (while not fully generalizable) results both for ATI and the role of ATI delivered courses to IMF CD overall.
3. While the RBM offers an opportunity for assessing interventions across the three regions, the use of the RBM approach still faces significant challenges in its consistent application across RCDCs and is not yet fully applied to training at the individual course level. In some cases, outcome indicator ratings do not seem to be an expression of the milestone indicators achievements, it is applied differently and there are some divergences in the views of how RBM is supposed to be measured and reported on. CAPTAC-DR for example uses an approach to rate ongoing activities in their annual reports that goes above and beyond requirements for documentation (could be considered a “gold standard”) which in some cases differs from their entries in the RBM system. They also view outcomes as being beyond their control and exclusively in the purview of the authorities. While attribution of results can represent a challenge given factors outside the control of CD providers, to the extent possible RBM (and theory of change) driven approaches aspire to ensure that inputs are managed through a credible results chain in a way that leads to expected results and must navigate unexpected circumstances to ensure effectiveness, impact and sustainability.

4. Field work was designed to rely on key informant interviews and focus group discussions, but some challenges were experienced in relation to gathering identified participants in ATI courses which in some cases meant that intended FGDs became interviews and in a couple of cases could not be realized despite repeated attempts in further virtual meetings after the end of the field visits and despite support from the ATI secretariat.

The following sections provide detailed analysis of the findings of the evaluation along the evaluation objectives.

2 PART I: PROJECT AND COURSE LEVEL EVALUATION

The report is structured line with the OECD DAC criteria and presents findings from the review of 12 tax and customs projects within the revenue administration workstream and six courses within the revenue administration, fiscal policy and inclusive growth workstreams in detail. Annex 1 provides a summary of the RBM rating for each DAC criteria aggregated at the workstream level.

2.1 Relevance

Alignment with country contexts, strategies, plans and priorities

CD under the revenue administration workstream (tax and customs projects and relevant courses based on the sample for this evaluation) was broadly aligned with context, development plans and strategic plans of government revenue administrations in all the case countries. These were all to a greater or lesser extent affected by the Covid-19 pandemic and global economic repercussions of the war in Ukraine in terms of economic growth, fiscal balance, debt levels and revenue collection. TA and training in the area of revenue administration and fiscal policy was therefore highly relevant to support increased revenue mobilization and often this ambition was concretely articulated in the strategic plans of the national revenue administration and reflected in IMF programs and Article IV reports.[3] The pandemic also became a driver for engagement such as when tax revenue to GDP and customs revenue dropped in two of the CAPTAC-DR member countries.

Overall, except in very few cases, the specific focus areas within the revenue administration workstream were found to be well aligned with government priorities and demand driven.[4] An example of shifts in priorities reducing buy-in and requiring realignment of CD was in relation to support to excise duties management in one of the members of AFW2. In this case, the authorities, led by the Ministry of Finance, decided to pursue the introduction of an excise tax stamp while the original plan was to support an organizational restructuring. While still relevant, the sequencing of the support had to be altered as the revenue authority (RA) became reluctant to move forward with the changes (see volume II for more details). In another AFW2 member country, though service delivery was overall demand driven, there were occasions where the authorities declined repeated TA offers in tax administration on data cleaning for tax ledgers, as they had established their own methodology. In addition, there were few cases, particularly in the CAPTAC-DR region, where it was hard to substantiate the level of demand and ownership at a more granular level due to lack of documentation. This was in part due to the limited information provided in some TA reports which did not provide much information except the fact that the "TA was in line with government priorities".[5]

In AFW2 support to integrity in customs was aligned with a common regional framework based on the WCO Revised Arusha Declaration on Customs Integrity. In one member country, there was strong buy in and high-level political support to counter corruption and promote integrity in customs while in another commitment shifted over time. In CAPTAC-DR support to Customs Risk Management also followed a regional approach.

In the period under review AFW2 overhauled its approach to TA in customs.[6] This relies on a sequenced approach to key reforms in three areas (Post-clearance Audit (PCA); Custom Risk

[3] The AFW2 case countries all had integrated revenue administrations while the CAPTAC-DR case countries had tax and customs directories.

[4] The evaluation uses the term focus areas for specific revenue administration topics that were addressed at country level usually through a sequence of missions and measured at the RBM level by one or more milestones.

[5] The evaluation team was not granted access to BTO reports to supplement the content of TA reports.

[6] The approach was revisited at the time of the outbreak of Covid-19. Not particularly to make TA more relevant to the pandemic situation per se but to make it more relevant in general.

Management (CRM); Border and Inland Control (BIC)), using international best practices, a result-based logframe, and a participatory, focal point-led, process. While following an overall generic and structured approach, it is adapted to reflect the country's specific needs and priorities. The approach has been rolled out to all AFW2 member countries and the regional perspective has been broadly welcomed by authorities and the process has deepened ownership. The amount of time it took for revenue authorities to buy in to the approach varied a bit across countries but at the end of the evaluation period, all members found the approach relevant. In one of the member countries, the approach was integrated in the revenue authorities' strategic plan and provided guidance for a country wide customs integrity plan.

The stated objectives and content of the sampled courses in Tax Policy and Administration, Fiscal Policy and Inclusive Growth delivered through ATI were highly relevant. There is a significant need for knowledge and skills taught in these courses mainly in relation to addressing deteriorating fiscal balances and case countries often aimed for increased revenue mobilization to support fiscal consolidation. The relevance of courses was increased by lecturers using data and cases from the region although the size and diversity across economic, administrative, fragility, linguistic, social and cultural dimensions of the 45 member countries covered clearly challenges the extent to which this can be done. Feedback from participants was broadly appreciative in the use of data and case studies, although some fragile country participants indicated less use of formal examples from similar country contexts. The formal and informal exchanges between participants (through country presentations etc.) also added significantly to the relevance of structured content courses delivered through ATI. The high relevance level is corroborated by data from ATI's post course surveys where participants rated the sampled courses 4.5 (out of 5) on relevance[7]. CAPTAC-DR provided extensive amounts of multi-country training under the revenue administration workstream which related to the TA provided and was considered relevant by participants[8]. Benefitting from the availability of ATI courses in the region, AFW2 provided less structured content training through workshops and seminars within the workstream increasing the importance of synergies between ATI delivered training and AFW2 TA (see page 23 for further reflections on different training modalities).[9]

Diagnostics and capacity needs assessments

Diagnostics are of vital importance to identify institutional capacity gaps, to inform the relevance of CD and can help build commitment to reform. In the case of tax administration, RCDCs have been able to rely extensively on TADAT based diagnostics. These provided central entry points to inform TA in tax administration in four out of six TA case countries and to some lesser extent in two out of six case countries.[10] They were often highly appreciated by the authorities as a useful tool but not always utilized to their full potential. In most case countries the assessment was becoming quite dated at the end of the evaluation period which reduced their relevance and usefulness. Most of the outcome indicators in the tax administration RBM framework are identical to the TADAT performance outcome areas.

Table 1: Most recent TADAT assessments per case study country

COSTA RICA	DOM REP	EL SALVADOR	GHANA	SIERRA LEONE	THE GAMBIA	CÔTE D'IVOIRE
May 2019	May 2017	Oct 2018	May 2017	Aug 2016	Apr 2018	Nov 2021

In some case countries, the authorities indicated that they would like a repeat TADAT while others indicated that they were planning to do self-assessments based on the TADAT methodology. Updated TADATs are critical to further strengthen the relevance of TA but RCDCs cannot directly respond to such requests as they are coordinated by and dependent on the resources of the

[7] However, ratings are very high in general, and relevance rates lower on average than the "overall" and "recommended" ratings which are 4.6 and 4.7 respectively. The outcome indicator for the overall satisfaction rating in the RBM system is 4.4.

[8] CAPTAC-DR delivered significantly more training with a total of 26 courses with 1239 participants

[9] AFW2 delivered two revenue administration relevant workshops and one seminar reaching a total of 78 participants

[10] In these two cases the report and/or implementation plan was not fully endorsed

IMF-based TADAT secretariat. Moreover, the “expiration date” of TADATs also to an extent rely progress against the implementation of recommended reforms.

Box 1: The relevance of TADAT

In the Gambia, the 2018 TADAT assessment and post-planning exercise provided a solid basis to establish a work program with the authorities. In El Salvador, identification of priority areas was strongly based on TADAT findings in two out of three topics (taxpayer registry and compliance risk management). In the Dominican Republic, TADAT served as a basis to lay out CD priorities to increase taxpayer’s compliance and tax revenue and strengthen the tax inspection and audit functions in line with specified national development goals. In Costa Rica TADAT identified weakness particularly in the area of compliance risk management which became a focus area of support. In Ghana, the TADAT was not formally approved by the authorities but has still been used to inform the GRA’s Strategic Plan, allowing an overall good correspondence between GRA priorities and TADAT-based AW2 missions in tax administration. In Sierra Leone an implementation plan was developed on the basis of the TADAT but progress stalled due to lack of approval by the Revenue Authority. The TADAT is most relevant to inform TA but could also be used to tailor the content of the Tax Policy and Administration course. Several lecturers used regional cases and examples to increase relevance of courses. Participants consulted through focus group discussions appreciated this and, in some cases, suggested even more tailored content. So far TADAT reports have not been used in this way

In addition to TADAT, Centers made use of specific scoping studies, evolving analysis, findings and recommendations from previous missions documented in TA reports that were routinely shared with the authorities. Diagnostic material on tax administration produced by FAD was also used but more in Revenue Mobilization Thematic Fund (RMTF) countries, especially in two of the three West African case countries where regular RMTF missions and analysis complemented the analysis of tax administration LTXs and STXs. In at least one CAPTAC-DR member country, which does not benefit from RMTF support, available FAD analysis was very dated, and its continued relevance was questionable.

Approaches, focus and depth of customs diagnostics vary from country to country and across regions. In AFW2 the new approach to TA in customs was based on a gap and needs assessment conducted by the AFW2 Regional Customs Advisor. The needs assessment categorized the countries’ challenges as financial, operational, human capacity, customer/compliance and external risk. In the one CAPTAC-DR member country, time release studies were conducted in 2019 and 2021 to assess the clearance of customs cargo identifying recurring weaknesses in customs declarations and post-clearance audit. Moreover, a regional diagnostic on post-clearance audit was carried out in 2021. Such diagnostics provided valuable tools to inform TA, particularly on post-clearance audit in the three case countries. Still one particular country has outdated customs diagnostics primarily due to a lack of demand from the authorities.

While internally provided diagnostics are important, the advantage of TADAT is the extent to which it is systematic and standardized in its assessment of institutional capacity gaps in tax administration. In customs, the IMF and the World Customs Organization (WCO) have collaborated on the International Survey on Customs Administration (ISOCA). The survey was created to collect quantitative and qualitative data on Customs administrations and to enable comparisons to be made between countries sharing common features. Importantly, the WCO and IMF state that one of six key objectives of ISOCA is to identify specific needs for technical assistance and capacity building. 51 countries participated in the first round of ISOCA, including all the CAPTAC-DR case countries (but none of the AFW2 or ATI case countries). The first survey was finalized in June 2020 and a report presenting some of the results was released in November 2021. As this exercise was completed only halfway through FY22 it was too late to be used to inform TA on customs in these CAPTAC-DR member countries within this evaluation period.[11]

While TADAT and available customs diagnostics are relevant at the level of institutional capacity, the evaluation did not find evidence of any standardized tools and approaches to assess evolving needs and critical capacity gaps at personnel level within the agencies supported in relation to enhancing human capacity, which is the main purpose of structured content training-oriented CD modalities. Such assessments would help identify critical capacity gaps and could help inform and strengthen synergies between training and TA. The relevance of such an approach is further emphasized by staff turnover and retention issues, which further complicates the challenge of identifying and closing critical capacity gaps within supported agencies. Relying more on this type of assessments is a strategic consideration to strengthen synergies between training and TA but would have to be undertaken by various IMF departments in conjunction and through partnerships with other development partners.

Continued relevance

While TA to revenue administration was broadly relevant and demand-driven, commitment and ownership are not static features. Political transitions and governance issues can result in major and sometimes sudden shifts in this regard. It can also result in reduced transparency and willingness to share plans and strategies complicating the possibility of alignment. In Central America, the authorities in one of the case countries published a broader political plan for fiscal policies but were not willing to share a detailed anti-evasion plan with CAPTAC-DR. This hampered a close alignment of support within this area (and the ability to verify it from an evaluation point of view). More commonly, political transitions or other factors resulted in staff turnover in senior government positions, influencing the ability of revenue administration authorities to maintain the reform effort of previous years. Overall, the Centers aligned well to political transitions and the fluctuating levels of commitment by scaling down activities or proactively responding when opportunities emerged. In one West African country, some activities were cancelled in response to inactivity by the authority but at least in one case revived opportunistically when the situation changed with a corruption scandal as the main driver.[12]

The Covid-19 pandemic overshadowed the delivery of IMF CD in most of the evaluation period. The IMF introduced an institution wide suspension of mission travel from March 13, 2020, with all mission activity, including all RCDCs being carried out in a virtual environment. This policy was still in place at the end of the evaluation period in April 2022. The response of the Centers in terms of swiftly transitioning to an unprecedented situation and delivering all training virtually was remarkable. Management and staff put significant efforts into keeping operations going and at the same time agencies and personnel receiving TA and training also made significant efforts to adapt to virtual delivery. In addition, in several countries, the Covid-19 pandemic increased demand from authorities that had been more reluctant to seek support as the importance of retaining revenues became apparent. This was the case in two of the three CAPTAC-DR case countries. In the third case country in that region, the pandemic provided a window of opportunity to address digitalization-related challenges within tax administration. Still in CAPTAC-DR, an interesting effect of the shift to virtual delivery was that the LTX could accompany more missions in the region. TA broadly continued to support the core focus areas identified in the revenue administration workplans especially as the initial shock of the pandemic was overcome and restrictions started to ease at country level. While this did not negatively affect the relevance of TA overall, some topics such as digitalization could perhaps have benefitted from more attention and swifter integration in workplans.

[11] CAPTAC-DR has informed that the diagnostics has been used in FY23 and FY24, outside the evaluation period. However, feedback on preliminary recommendations indicated that confidentiality could be an issue in terms of using WCO diagnostic work. It is critical that such issues are overcome with reference to the significant success of TADAT.

[12] This case is further discussed below.

In training, significant efforts were invested in finding the best approach to virtual training, which was challenged on a range of fronts including participants and some lecturers being in different time zones. The main challenge was that more ground had to be covered in a shorter time as the duration of sessions had to be reduced without the content being significantly reduced. This implied that there was more “lecturing” and less time for interaction. Participants appreciated the use of Mentimeter and other interactive virtual tools in some courses, which possibly could have been used even more. This lesson could be valuable for the post-pandemic virtual or blended training.

Content of the sampled courses remained the same overall, but lecturers and participants confirmed that many courses managed to incorporate some Covid-19 relevant examples and cases. Beyond the sampled courses, ATI complemented their general offer with webinars and courses based on IMF’s Special Series Notes on COVID-19 in areas like public finances, commodity prices, external and monetary accounts, the real estate sector among others.

As noted, a key aspect of ATI course delivery that strengthens relevance for participants is the networking with peers from countries in the region and both formal and informal sharing of experience from a huge variety of country contexts. This naturally suffered during the pandemic as it is very difficult to emulate individual networking in a virtual environment. Participants suggested specific things that could have been done such as reducing content/lectures (or recorded these for individual access), prioritization of more interaction through breakout rooms, linking participants from the same country with each other.

2.2 Coherence

Internal coherence

The planning and delivery of IMF CD to revenue administration is complex with area departments (AFR, WHD) having responsibility for the strategic approach while the technical or CD departments (FAD) have responsibility for coherence and quality of TA and some of the courses delivered through ATI within the thematic area (tax policy and administration course). ICD is responsible for training in some areas based on the catalogue including the sampled fiscal policy and inclusive growth courses.[13] The governance structure of the RCDCs reflect these different roles with the Center Directors reporting to the area departments and the long-term experts (LTX) being backstopped by the CD department.

Furthermore, revenue administration is supported through different funding vehicles. The RCDCs are funding vehicles as most of their resources are fundraised specifically for them. Significant amounts of CD to revenue administration is also delivered through the RMTF and relevant diagnostics are delivered through the Tax Administration Diagnostic Assessment Tool (TADAT) .[14] In effect this means that CD to revenue administration is delivered by a mix of support from HQ (including RMTF financed support), the RCDCs and in some cases by Resident Advisors who are LTXs that are based in one specific country and do not report to the RCDC but to HQ and also often are separately funded. Importantly for coherence, the IMF has a declared ambition of integrating CD with surveillance and programs, which are areas that are led by the area departments. IMF Resident Representatives (Res Reps) are very central to the coherence across all these dimensions at country level. Res Reps report to the area departments and are considered the “face” of the IMF on the ground. Not all countries have Res Reps, however, as this depends on the level of IMF engagement in the country. In the case countries selected for

[13] Coherence of revenue administration covers tax and customs projects and the findings in this section applies to both areas both is most prominent in relation to their function to support revenue mobilization. Customs still play a significant role in this regard in many countries, but it is also having broader objectives such as facilitation of legitimate trade, prevention of smuggling and illicit trade, and security of the borders. Such issues are rarely covered in IMF programs and surveillance and therefore possibly less coherent with the broader fiscal agenda. TADAT is obviously tax administration focused and RMTF also dedicates more attention to tax than the broader customs functions. At country level, this is reflected in a somewhat different configuration of external partnerships as addressed in the subsection on external coherence.

[14] Both RMTF and TADAT have recently been evaluated separately.

this evaluation, Dominican Republic and El Salvador were represented by a regional Res Rep based in Guatemala.

Table 2: Different IMF actors with relevance to Revenue Administration in the case countries

COUNTRY	RCDC	RMTF	RA	RES REP
Costa Rica	CAPTAC-DR	No	No	Yes
Dominican Republic	CAPTAC-DR	No	No	No
El Salvador	CAPTAC-DR	No	No	No
Ghana	AFW2/ATI	Yes	No	Yes
Sierra Leone	AFW2/ATI	Yes	No	Yes
The Gambia	AFW2/ATI	No	Yes	Yes
Côte d'Ivoire	AFW/ATI	Yes	No	Yes

There is some internal coherence in the IMFs delivery of TA to revenue administration, but CD could be better strategically coordinated through regular and detailed CD Country Strategy Notes (CSN), especially in relation to synergies between TA and training. The following outlines the findings of the evaluation across the different entities involved.

There is coherence between TA delivered through the RMTF and the RTACs in countries where the former is active. In Ghana for example it was clear that there was an overall clear division of labor structured around the RBM framework objectives and outcomes. This was reflected in terms of the complementarity between the RMTF focus on support for transformative reforms and AFW2 having a permanent presence in the country, as well as a long history of engagement with the revenue authorities enabling it to focus on implementation in a continuous manner. LTXs are backstopped by HQ experts and STXs are backstopped by the LTXs. Overall, the backstopping relationship was reported as being constructive and fulfilling its function of support and quality assurance of CD.

Given the central importance of TADATs to inform TA to tax administration, efforts could be made to strengthen coherence between RCDS and the TADAT secretariat. This could help increase the number of up-to-date TADATs in RCDC member countries (RCDCs could discuss need for updates with member countries and help communicate this to the TADAT secretariat) or include as a concrete activity in workplans to support and quality assure TADAT self-assessments. TADAT uses only certified assessors but LTXs and regional advisors typically participate in TADAT missions and are involved in pre-assessments and assessments. Building on this engagement, LTXs and STXs could support country led self-assessments and, if needed, with some supervision from the TADAT secretariat.[15]

As stated above, Res Reps played a central role in relation to both internal and external coherence in the case countries. In the CAPTAC-DR, a Regional Res Rep is based in Guatemala and covers most of the CAPTAC-DR member countries. One of the three case countries has its own Res Rep given that the country has a program with the IMF and coordination was found to be highly effective. There was a fluent dialogue, both formal and informal, for the purposes of managing the revenue administration CD projects. This level of engagement was lacking in the two remaining case countries in the region. In all the AFW2 case countries, the role of the Res Rep was viewed positively as well in relation to supporting internal coherence. Res Reps are also a very important resource in relation to the creation of synergies between training and TA mainly with regards to the selection of participants for ATI delivered courses and also potentially in relation to follow up. Lecturers sometimes reach out to Res Reps and consult them in the process but not in a systematic manner. One Res Rep reported to have been consulted only on a handful of occasions, others indicated that it happened more frequently but not systematically.

[15] Adding to the importance of increasing coherence, the use of TADAT performance outcome areas as outcome indicators in the tax administration RBM framework implies that progress can only be fully verified based on a repeat TADAT.

One case country in AFW2 has a domestically based Resident Advisor (RA) in revenue administration and there is also an IMF Res Rep in country providing a very strong presence and access to government institutions. Given this presence and absorption capacity issues, coherence is particularly important. Both AFW2 and the resident advisor for example worked on the new Integrated Tax Management System. Coordination was generally very good but could be further strengthened, for example through more regular sharing of workplans and better coordination of missions.

Table 3 provides an overview of the status of case countries in their programmatic relationship with the IMF, their most recent Article IV (surveillance) and CSN, and finally, if the Article IV has a summary of the CD strategy to the country.

Table 3: Overview of relations with the IMF

COUNTRY	PROGRAM	ARTICLE IV	CSN	CD SUMMARY
Costa Rica	EFF, Mar 2021 - Jul 2024[i]	Mar 2021	2022	No[vi]
Dominican Republic	None	Jun 2023	None	Yes (2p)
El Salvador	None	Jan 2022[v]	None	No[vii]
Ghana	(ECF May 2023)[ii]	Jul 2021	2017	Yes (3p)
Sierra Leone	ECF, Nov 2018 - Nov 2023	Jul 2022	2017	Yes (5p)
The Gambia	ECF, Mar 2020 - Jun 2023[iii]	Dec 2021	2017	Yes (2p)
Côte d'Ivoire	ECF, Dec 2016 - Dec 2020[iv]	Jun 2022	2022	Yes (4p)

i) FY22 only, ii) Not within evaluation period, iii) FY21 and FY22 only, iv) FY20 only, v) Also 2023 but not publicly available, vi) Only a table listing TA missions but extensive CSN available for Costa Rica and vii) a two-page summary made available for the evaluation team (under confidentiality agreement).

In countries with active IMF programs, CD in revenue administration tends to be broadly consistent with program priorities although generally the priorities of programs are at a higher level. In the CAPTAC-DR region alignment was notably stronger in the only country with an IMF program (from the final year under evaluation) than in the other countries in the region. This was partly due to the presence of an IMF Res Rep in the country. The area department (WHD) and CAPTAC-DR work closely together, including through mission participation, to ensure CD is consistently integrated with the authorities' reform priorities and effectively supports program implementation. Concretely, one of six structural benchmarks in the current program addressed revenue mobilization requiring the approval of a new Tax Compliance Improvement Plan, in line with the IMF staff recommendations. In one of the AFW2 case countries, there was also strong complementarity between revenue administration reforms targeted in the IMF program and AFW2 support indicating that these were informed by TA advice. This coherence also involved supporting structural benchmarks with targeted TA supporting the setup of the internal affairs unit.

At a broader level, there is also coherence between CD and surveillance. Article IVs often identify improvement of revenue collection as an important part of fiscal consolidation (in a context of rising fiscal deficits and debt levels). However, at this level, isolating revenue administration from tax policy in terms of evidence and findings is not easy as evidenced by at least two of the case countries (One in CAPTAC DR and one in AFW2). In the CAPTAC DR one, the authorities pointed to improvements in tax administration while IMF staff were sceptic of the expected yields from tax administration measures materializing in the short-term and recommended a series of tax policy measures instead.

Article IVs increasingly provide coverage of capacity development activities but not comprehensively or in depth. As demonstrated in table 3, CD summaries are available in five out of the seven most recent Article IV staff reports but range between two and five pages of which most is a table of priorities for all workstreams supported. Similarly, TA reports also rarely make reference to Article IV reports. The limited coverage leaves a gap in terms of documenting the strategic approach to CD at country level, especially when there is no updated CSN in place.

CSNs are supposed to complement the broader CD Regional Strategy Notes (RSNs) and provide a more detailed strategy for CD delivery at country level including analysis of the political context, engagement with the IMF, prioritization of CD, engagement strategy among others and should be updated every three years. As per current policy, CSNs are only required for “heavy users” of IMF CD which is defined as countries where CD spending is greater than US$1.5 million annually.[16] Only two of the seven case countries had an updated CSN with reasonable coverage of the required areas. While not a requirement, it is also worth noting that none of the CSNs corresponded well with the current phase of the three RCDCs which inhibits the role they could potentially play in guiding CD at the country level in a more strategic manner. The lack of updated CSNs in some countries leaves a gap in terms of defining a strategic approach to CD delivery, prioritization, consistency and effective communication on CD at country level. RSNs were available and updated on an annual basis for both WHD and AFR but while providing a broad strategic dimension to CD they are too generic to guide CD delivery at country level.[17]

On a positive note, the most recent RSN for WHD (2023) states that the preparation of CD CSNs to top CD users and program countries in the region will be expanded. This is important for two key reasons. First, both RSNs and available CSNs indicate revenue administration as a priority area which strengthens the need for coherence to increase effectiveness and impact. Secondly, in relation to engagement strategy, currently neither CD summaries (in Article IV reports) nor CSNs set out specific approaches to how TA and training can be used in a complementary manner to strengthen the effectiveness and impact of CD at country level. This is particularly important for strengthening the link between human and institutional capacity strengthening measures.[18] Currently none of the reviewed AFW2 and ATI sampled case country CSNs make reference to training and the role of ATI or structured content training which suggests a gap in the extent to which it is considered alongside TA as a key component of CD. One of these CSNs does make reference to two workshops on Financial Programming and Policy noting that the program was “extremely good” but “because of lower capacity, the tool is yet to be put to good use”. In CAPTAC-DR, one CSN makes few references to training including a reference to a virtual workshop in tax compliance risk management. Despite these sporadic references, in all the case countries, there is no *strategic* approach to integrate TA and training at country level, including through the use of capacity needs assessments. Currently this represents a glaring missing connection to ensure better synergies between training and TA in a formalized manner. Revenue administration would be a perfect area in which to initiate a more integrated approach. AFW2 experts and Res Reps could play an important role in this and, for example, be introduced and should be granted access to ATI’s extensive existing database on all participants in ATI delivered training and use this in their support in related areas of TA.

External coherence

Few development partners have policies and support specifically to revenue administration and therefore external coherence in this area tends to be structured around the broader domestic revenue mobilization agenda.[19] The case studies indicated that many partners look to the IMF for this type of specialized support and many indeed provide funding to IMF CD through the RCDCs, other IMF vehicles (RMTF, TADAT) or through bilateral arrangements (e.g., in case of

16 The IEO’s evaluation of IMF CD provides a comprehensive overview of the background and requirements of CSNs concluding that coverage of heavy CD users is patchy.

17 And are indeed not intended to do this as they focus on higher-level department CD strategic objectives.

18 As part of this evaluation the most recent Article IV reports and CSNs available for all 45 members of ATI were reviewed and this exercise showed very few references to training or ATI activities.

19 This is true both for tax and customs administration although some of the non-revenue mobilization features of customs do lead to other types of partnerships at country level.

the Gambia).[20] Considering this important role, development partners expressed clear expectations that the IMF would provide detailed, technical briefings through dedicated PFM or preferably DRM coordination forums (beyond information sharing) and proactively consult development partners that are active in the broader DRM and PFM agendas, especially those agencies that provide budget support and look to the IMF to inform associated policies.

Across the regions covered in this evaluation, the main international development agencies working on supporting the DRM agenda are the FCDO, GIZ, the World Bank, the multilateral development banks IADB and AfDB, the European Union and USAID. The Inter-American Center of Tax Administrations (CIAT) and the OECD are the main actors on international taxation issues. In most case studies, several positive partnerships exist (outlined in detail below) and the Centers make extensive efforts to share information and engage with various development partners. Nevertheless, specifically in relation to revenue administration, interviews indicated that external coherence could be strengthened especially in relation to sharing of more in-depth information and effective coordination with development partners. Several partners interviewed held clear and strong views on this issue.[21] There is no doubt that efforts are being made by the RCDCs and numerous examples outlined below provide examples that can be built on but evidence suggests that overall the effort is not always effective. Positive examples indicate that the more concrete collaboration is made, the more effective it tends to be.

Positive examples of bilateral partnerships include the collaboration with the IADB in the CAPTAC-DR region although it was not consistently positive across the three case countries. In two of them, collaboration structured around the digitalization agenda was strong but in one case study country, there was less contact with IADB despite both agencies being active in supporting revenue administration. The Gambia represents a positive example of close collaboration with the World Bank in relation to supporting the rollout of ITAS, which is funded significantly through the World Bank.

Perhaps given the distinct features of customs, a number of positive examples of collaboration in this area emerged from the case studies. In CAPTAC-DR a series of regional workshops on risk management in coordination with the IADB, SIECA and COMIECO led to the adoption of a regional strategy on customs risk management. The collaboration between the IMF and WCO on country specific issues and in relation to ISOCA is also noteworthy. The latter seeks to address a gap in diagnostics but as noted in the diagnostics subsection, the pilot exercise took place towards the end of the evaluation period and has only subsequently been used to inform TA. In Sierra Leone, a comparative analysis with WCO to inform a workshop related to the Customs Integrity Framework provided a positive example of external coordination. The WCO's Integrity Assessment Tool complemented AFW2's efforts at a regional level in terms of shared analysis, joint conclusions and a high level of collaboration. Lessons from this exercise could be used to take the work on ISOCA to the next level. AFW2 has regular meetings with the World Bank on trade facilitation work and with ECOWAS and the Africa Continental Free Trade Area Secretariat on regional standards and priorities in customs. Bilateral development partners that are most active in the customs area are USAID and GIZ, in the CAPTAC-DR region there were examples of good collaboration with these agencies.

Good development partner coordination at country level is often associated with the existence of a coordination forum on DRM with regular meetings that is led by, or at least has active participation of, government revenue agencies. Most often, however, DRM becomes a subset of Public Financial Management (PFM) coordination and addressed in the PFM coordination forum,

[20] This report does not make an explicit distinction between “financial” partners that fund the work of the RCDCs and development partners more broadly that support the authorities with CD that is directly or indirectly relevant to revenue administration as coordination with these is equally important.

[21] This issue is not novel. A recent IEO background paper on coordination and collaboration with partners in IMF Capacity Development (Radelet, 2022) found that “…in most cases, in-country coordination is limited to basic information sharing and avoiding duplication of activities, and is rarely more strategic or more deeply collaborative.” And: “Regarding collaboration on thematic issues, aimed at leveraging the expertise of other CD providers with complementary skills, there has been relatively little progress. This kind of collaboration is difficult, costly, and complex, but in certain circumstances it has the potential to bring significant benefits.” The findings of this evaluation is in line with the findings of the IEO background paper.

which is not ideal given the importance of both issues. Only one of all the case countries had a dedicated DRM coordination group (including the FCDO, GIZ, the World Bank, and the European Union, USAID, GIZ). A matrix of activities has been developed to support coordination and avoid the risk of duplication, and the revenue authority played an active role in these coordination efforts. Other countries had broader or differently focused coordination forums such as “PFM development partners working groups or budget support coordination groups. Other case study countries relied on bilateral coordination only and in one case country, the authorities directly lamented lacking coordination by development partners in the area of revenue administration. Finally, there was evidence from at least one case study country that the Covid-19 pandemic had negatively impacted on coordination more broadly around PFM issues (mainly due to virtual “fatigue”).

Given the key role of the RCDCs in providing support to revenue administration development partners, it was found that in most cases more could be done to go beyond information sharing and engage development partners in more in-depth discussions about progress and challenges in this area. Many noted in interviews that they knew little about the ongoing TA delivered by the RCDCs to revenue administrations and even less about training in related areas. In one case study country that otherwise is a best-case example, at least one partner reported knowing very little about AFW2 support to revenue administration. CAPTAC-DR distributes an overview of capacity development activities to partners in the region (mainly the financial partners), but it was seen as superficial, with key development partners being interested in more in-depth information. RCDC experts do not systematically engage in the coordination forums, for example to offer debriefings at the end of missions or to share key diagnostics, which would be of significant interest to other development partners. This means that findings from TA missions are not systematically shared although there have been some improvements in providing access to TA reports through Partners Connect.[22] Development partners at times undertake work that could help better inform IMF CD such as a tax gap assessment for VAT and import duty by FCDO on Ghana.

The lack of a more in-depth approach to coordination with development partners represents a missed opportunity both in relation to development effectiveness and to better showcase CD delivered in the area of revenue administration to increase visibility, which is also important for fundraising. In terms of effectiveness and impact, coordination and generation of synergies, particularly with agencies providing budget support, could be beneficial. While such engagement happens occasionally, it could be implemented systematically at relatively low cost at least in RTACs such as AFW2 and CAPTAC-DR that have relatively few member countries.[23]

While ATI makes efforts to coordinate with external partners, there is still less coordination and information sharing with external partners on training than on TA at country level. It is not part of ATI’s “business model” to engage with partners at country level and hence the evaluators found practically no evidence of this happening at the ATI country case study countries. . Efforts in strengthening this area would depend on a CSN guided strategic approach to better align and integrate the availability of ATI training and AFW2 TA at the country level. ATI could benefit from stronger partnerships with development agencies working on institutional strengthening with focus on human resource management. Res Reps could also be requested to share information about the numbers and areas in which country officials have received training through ATI. Participants in courses with relevance to specific issues on the agenda could also be invited to coordination forums to foster the use of skills acquired. Such efforts would have the additional benefit of increasing ATI visibility, which is currently low at country level.

22 The evaluation team was not able however to identify TA reports systematically on Partners Connect despite being granted access. This suggests the need for improvements in the use of the platform.

23 It is estimated that the cost of this effort would be low relatively to other more resource intensive efforts to disseminate information through newsletters or bilaterally as experts would need to invest limited time in preparation and mainly share their knowledge and insights with key partners.

2.3 Effectiveness

The effectiveness of CD to revenue administration in the first three years of the phases under evaluation has been reasonable, particularly when considering the impact of the Covid-19 pandemic.[24] To inform this assessment, the evaluation considered progress across the revenue administration projects and underlying activities or focus areas in the case study countries. Overall, many tax and customs projects progressed well and a number of examples have emerged of specific focus areas within tax and customs projects that are likely to be fully achieved and impact positively on revenue administration at country level. Other focus areas faced more challenges, underperformed, were delayed or cancelled, affecting the level of effectiveness with the respective projects and their level of achievement of milestones and outcome indicators. All projects are still under implementation and effectiveness is likely to have picked up in the aftermath of the Covid-19 pandemic that completely altered the way CD was delivered and challenged effectiveness. ATI courses also suffered overall from the impact of Covid-19 but at the end of the evaluation period there were clear indications that performance was returning to previous levels.

Overall, the RTACs applied the RBM system relatively consistently using it as reference point for the pre-defined objectives in the catalogue but also adapting the outcomes, indicators, baseline and milestones to specific country contexts but there were variations in the way RBM was applied across the Centers. While AFW2 reported in line with ICD guidance, CAPTAC-DR went above and beyond requirements with very granular reporting in its annual reports using more sophisticated approaches, such as weighting of ratings according to the amount of resources used.[25] However, their reporting through the RBM system followed a different approach. It is hence critical that ICD continues working with the Centers to harmonize approaches to rating, including how and when this is introduced to the central RBM system. From an evaluation perspective, it was sometimes a challenge to fully reconcile the results chain and log frame from country specific objectives, outcomes, indicators, baselines, targets and milestones with activities on the ground (missions) and associated reporting (TA reports) based on available evidence, making it harder to verify the IMFs contribution to the results achieved. The evaluation suggested that there is also a need for further discussions and clarification on attribution of results throughout the results chain. At ATI, the RBM is being applied to training for courses that are part of the Results Based Management (RBM) Catalogue, but ratings are not yet being applied systematically. It was possible however to rate the sampled courses in the country cases mainly based on ATI's detailed database of participant performance.

With these caveats in mind, the evaluation broadly concurred with outcome indicator ratings at workstream and project level provided by the RCDCs in the RBM system with no significant variations identified. At the end of the evaluation period, with two years remaining of the current phases, ratings averaged 2.4 (the same for customs and tax projects) thus between partially and largely achieved. The average evaluation ratings are marginally lower at 2.3 (2.3 for customs and 2.2 for tax projects) slightly closer to partially achieved. In the context of a medium-term evaluation and with all the challenges associated mainly with the pandemic, this is a very reasonable level of achievement. The remainder of this section provides a number of positive examples of how results were delivered effectively and based on evidence from the country cases some examples of activities with room for improvement.

Effectiveness of TA revenue administration projects – what worked and what could have worked better

A focus area under the objective to strengthen core tax administration functions that was supported with some success in all three AFW2 countries was to perform sector audits. Work in

[24] As should be clear, this section does not pertain to evaluate the effectiveness of all TA and training delivered by the three RCDCs but focuses on CD to revenue administration. Also the number of case countries is limited. This means that the sample of courses and member countries is low particularly for ATI.

[25] CAPTAC-DR uses a formula to define the weights based on expert-weeks that are applied to the individual ratings given by the LTX.

this area made strong progress in the Gambia where a key ambition was to conduct audits for identified specialized sectors. A successful audit was conducted in telecommunications and since expanded into tourism and hospitality sectors leaving the outcome largely achieved. This Gambian case is further elaborated in the impact section. In another country in AFW2 case study countries detailed preparations to undertake an audit in the telecoms sector were made but since the activity area ran into some delays, the outcome was partially achieved. This was also a focus area in in the third case country for AFW2, but activities took some time to pick up and eventually only started to progress in FY23, which is outside the evaluation period. With good experiences from sector focused audits from the other countries in the region and the traction made so far, the revenue authority from the latter country has fully bought into the focus on auditing of its financial sector. There is a good chance that this focus area will perform better in the last two years of the current phase and potentially have a significant impact in line with the results in the Gambia although the size of the economy and complexity of the associated sectors must be considered in relation to expectations.

Another area that performed well was AFW2's support to the establishment of an internal affairs unit in the Gambia. The unit has recently been made functional and operational. This was also a structural benchmark under the IMF program which indicates good synergies between CD and program support.

Support to customs has made relatively slow but steady progress in AFW2 case countries following the new approach outlined in the relevance section. In two out of the three case study countries, support to post-clearance audit, custom risk management and border and inland control has made good progress over the last two years with most required documents (including log-frame, mandate, policy, standard operating procedures) finalized for each area, although they were still pending approval by top management in both countries during field visits. In addition, a range of training materials has been developed. Ratings have adequately reflected this and while generally in a good way, progress needs to continue and speed up in the final two years of the current phase to achieve the expected results.

A focus area under the objective to strengthen revenue administration management and governance arrangements that has received substantial attention and been quite effective in CAPTAC-DR is compliance risk management. Significant investments were made through numerous missions. In one of the case study countries, the approach was not immediately welcomed with authorities finding the issue to be somewhat imposed by the IMF. Indeed, a tax and customs compliance improvement plan to reduce compliance gaps in Value Added Tax (VAT) and Corporate Income Tax (CIT) was a structural benchmark under the IMF program. This represents a very good example of alignment between IMF programming and CD. Moreover, the updated TADAT for this country had identified risk management methodologies and the compliance improvement plan as areas of weakness. Eventually, after strong engagement under the difficult circumstances caused by the pandemic, the authorities became increasingly committed to the approach. Further evidence from TA reports indicate that real progress was made in consistent progress in planning, going from risk definition to the design and implementation of compliance plans. The next stage, concerning quantitative verification and validation of compliance risk is work in progress and reflected in a partial achievement to date but based on previous activities also quantitative monitoring and evaluation of tax compliance should be within reach within this phase.

A second case study country in the CAPTAC DR region had a similar turn of events despite a much more complex operating environment for CD delivery. The country has no program with the IMF, the TADAT is outdated and collaboration with the authorities have been complicated by a marked drop in transparency overall and willingness to share strategic plans.[26] Similar to the first country, support under this area had a difficult start. The IMF response to this was a strong recalibration of expectations and positive messaging rather than focusing on gaps and delays.

[26] Interviews and the IMF website on third case country document that the authorities have refused publication of the 2023 Article IV. This is highly unusual and a strong confirmation of findings from field work in this country on the complexity of the operating environment (see case study).

After a core team had been constituted that was dedicated to working on compliance risk management, continuous progress was noted and consistently reported. The team learned fast and also exchanged experiences with a similar team in Honduras. Trust was also established between the STX (working continuously on the project) and the team. By FY22, strong progress had been made in relation to compliance risks and the milestone that had been rolled over from FY21 was classified as partially achieved. The operational context in the above second country creates high risks to the impact and sustainability of these achievements especially in relation to working with a relatively narrow team that depends on senior management buy in and politicized, hierarchical decision making.

In customs, good progress was made in one CAPTAC DR case study country where TA to risk-based strategic management and customs clearance and trade facilitation was highly effective. Progress was registered in relation to the enactment of the Customs Law (replacement of obsolete legislation) and regulations and in compliance risk management. Notably, these have contributed to the creation of a Risk Management Unit and a Customs Intelligence Unit within the Customs Agency, and the establishment of a national trade facilitation committee (composed of public and private sector organizations), among other organizational reforms in the sector. TA was delivered in close collaboration with IADB and WTO, was aligned well with national priorities and delivered consistently with significant involvement of the regional customs LTX and an STX. Achievements included adoption of medium-term plans with key performance indicators (KPIs), application of a segmentation strategy of traders based on risks levels, and an increased level of awareness of the importance of sectoral post clearance audits based on risk management and in line with the WTO Trade Facilitation Agreement. As noted in the relevance section, the diagnostic framework for customs in this country was considered better than in other countries which might have contributed to the progress registered. The main question, however, is whether all these achievements can in fact be attributed to the CD provided as the support consisted in only four missions during the three-year long evaluation period. While achieving impressive results with very limited resources is possible in a highly favorable context, with realistic and achievable indicators, the issue of attribution of results within the IMFs RBM approach need further scrutiny.

Support to customs also performed well in in one of the other case countries while there was limited traction in the third case study country. CAPTAC-DR successfully supported the development of the Aduana Fácil initiative, aiming to tackle digital related challenges and modernize the customs and telecom services to business operators. As a result, methodologies and guidelines for the identification and mitigation of operating risks in sensitive sectors was developed. A successful contribution was also made to the amendment of the National Customs Law, based on analysis and sharing of international good practices. The limited traction in the third case study country was mainly in response to a drop in demand from the authorities although this changed to an extent as a result of the pandemic. In some areas, such as increasing voluntary declarations through more effective control during the process of goods clearance, new plans were drawn up to respond to the changed circumstances.

There were also examples of areas where TA was less effective as for example in relation to support to strengthen excise duties management capacity in one of the AFW2 case study countries. It was already referred to in the relevance section as an example of a focus area where buy in from the Government was limited due to changes in priorities at the time. This resulted, however, in none of the recommendations made in FY20 being adopted by the authorities and in FY21 support for the manual and staff training was suspended in part due to the Covid-19 pandemic. The authorities requested more practical guidelines and were interested in being supported in the introduction of a digital tax stamp. Moving forward, a decision has been made to appoint another STX to work in this area with the authorities.

Several activities in AFW2 appeared to be delayed and behind schedule, which is partly a consequence of the difficulties imposed by the shift to virtual delivery for most of the period under evaluation. This was the case for most focus areas related to strengthening revenue administration management and governance arrangements such as the support to the customs integrity framework in most countries, which is also recognized as a highly complex and politically sensitive area. In one case study country, support to the ICT strategy and VAT compliance was postponed. However, there are signs that support to tax administration will produce better outcomes in the second part of the current phase assuming that support to the ICT strategy and excise are carried out as planned. This is partly due to a change in commitment to tax reform, the agreement of the new ECF program with the IMF and increased support from the RMTF.

In another AFW2 case study country, planned work to strengthen stakeholder engagement did not lead to the expected results, in part because it coincided with an intense period dedicated to the launch of the new Corporate Strategic Plan. Absorption capacity is also a known issue in this country(as noted in the coherence section) and stakeholders pointed out that delays are likely to take place due to limited human resources, competing priorities and lack of skills (including sometimes computer literacy). These issues also common in a in the third AFW2 case study country where a CD activity incorporated data analysis competencies including training in advanced Excel use, audit analysis, and data matching.

In two out of the three CAPATC DR case study countries, limited progress was made in the area of tax inspection and tax audit. The area had limited traction but still utilized significant resources, especially in FY21 without clear evidence of how inspection coverage through the adoption of a declarations management plan has been improved. The achievement of milestones was postponed to FY22 and while it was dropped in one of these two countries, continued efforts were made in the other. A key issue in relation to the limited effectiveness was a transition in the use of STXs. In the above country where continued efforts were made, there was limited progress in FY21 but a clear plan and TA reports adequately reporting on the limited progress in the use of new models. In the second year, however, the focus changed from tax inspection to audit of large taxpayers after a different STX took charge. The focus of the TA hence appeared disconnected from what was provided by the previous STX without reporting on previously planned activities in the TA reports. In FY22 a third STX took over this area again with a significant change in the style and approach to TA with no references to the work of the previous STXs. Challenges in relation to sustaining the use of STXs are inevitable but in this case backstopping by the LTX could have focused more on continuation of the support provided. The LTX for tax administration also changed in this period which might have added to the issues identified in relation to TA in this area. It is important to emphasize that the issue was largely isolated to this focus area as there was great consistency in other focus areas such as for example compliance risk management.

Effectiveness of training

Although the predominant focus of this evaluation is on ATI delivered training, RTACs are also mandated to use training as part of their CD delivery.[27] Over the course of the evaluation period AFW2 delivered two revenue administration relevant workshops and one seminar reaching a total of 78 participants.[28] CAPTAC-DR delivered significantly more training with a total of 26 courses with 1239 participants over the three years. Most of these courses were multi-country but this is within a limited number of fairly heterogeneous member countries. CAPTAC-DR does not, however, capture the results of training in as much detail as ATI but were generally viewed as an effective supplement to TA.[29] It is possible that additional customized training both from CAPTAC-DR and AFW2 could have gone under the radar for example if it has been integrated with TA delivery.[30] The registration of different types of CD delivery by the six modalities is still

[27] The focus was determined by the design of the evaluation.

[28] These numbers were supplied by ICD at the start of the evaluation period.

[29] Hence conducting a more detailed analysis of the effectiveness of CAPTAC-DR training was challenging.

[30] AFR’s Regional Strategy Note (RSN) for Capacity Development for FY21-FY23 for example mentions that in AFW (not AFW2!) TA missions in tax administration are systematically combined with training sessions. The Covid-19 pandemic might have impacted this type of CD delivery. In the future it will be important for RTACs to record and report systematically on the delivery of different CD modalities. AFW2 has started doing this (as discussed in the entity level note) but still it was only possible to identify limited training on revenue administration in the region.

quite recent and the disturbances to CD delivery caused by the pandemic could have further complicated this effort.

The effectiveness of ATI delivered training (within the revenue administration workstream) at country level was assessed following a stringent methodology outlined in the three ATI case studies. While the RBM approach has been rolled out to be applied to training, it is still not systematically applied.[31] None of the sampled courses had been rated by IMF staff but the RBM structure and data collected was used to assign evaluation ratings to the effectiveness of the six sampled courses at country level. The RBM framework for courses consists of an objective and three outcome indicators. The objective, which is generic across most courses, is that participants effectively acquire knowledge and skills taught in the courses and use them subsequently on the job or in their interaction with the Fund. From an evaluation point of view this consists of two objectives that need to be assessed separately. First if participants acquired knowledge and skills and secondly if these skills were applied. The results framework has three outcome indicators to measure if the objective has been met of which two rely on results from the pre- and post-course tests and one on the post-course survey results. The targets are that 60 percent of participants should obtain more than 60 percent of post-quiz answers correctly, that learning gain should average at least 15 percent and, for the one based on the end-of-course survey, that post-course feedback should be at least 4.4 (out of 5).

These outcome indicators do not measure the second part of the objective, which would require obtaining information from participants after they have returned to their home countries and returned to work. Due to resource constraints, such follow-up does not currently happen, which in the view of the evaluators implies that the objective, as it is formulated, cannot be assessed.

In the following, the results of the evaluation on the effectiveness of ATI delivered courses at country level based on data from ATI’s extensive participant database (first part of the objective) and focus group discussions (second part of the objective) are presented. As stated in the relevance section, the assessment of ATI at country level provides some very interesting findings in relation to strengthening the impact of IMF CD at country level through a more integrated and strategic approach. ATI structured content training has an extremely important role in this regard to complement TA and RTAC delivered training that is often embedded in TA or more customized.

ATI training at country level was relatively effective. In all three countries it was found to have more than partially achieved the objective but with variations in how this result was achieved. The evaluation of the effectiveness of courses at country level has provided granular feedback from a small sample of participants and generally paints a relatively promising picture in terms of effectiveness.

In all countries participants acquired knowledge and skills evidenced by pre- and post-course test results and information shared by participants in focus group discussions and interviews. There were significant limitations on the data available, especially for one of the three ATI case study countries and for the tax policy and administration course across the three countries.

In two out of three ATI case study countries, participants had significant learning gains, with both countries almost meeting the target for outcome 1 and significantly exceeding the target for objective 2. The third case study country did not perform as well against these outcomes (results in table 4). In relation to outcome 3, all three countries had an average rating of 4.6.

[31] The evaluators received a spreadsheet with RBM data for ATI structured along the same lines as RBM ratings for TA. The spreadsheet included 13 courses but only one had outcome and milestone ratings filled per end of April 2022. This course was not sampled for the evaluation (AFR_FAD_Anti-corruption course_FY22).

Table 4: Country level effectiveness of ATI training in the three case study countries

	COUNTRY 1	COUNTRY 2	COUNTRY 3
Outcome 1	54%	53%	31%
Outcome 2	31%	29%	21%
Outcome 3	4.6	4.6	4.6
1st part of the objective	Largely achieved	Largely achieved	Partially achieved
2nd part of the objective	Partially/Largely achieved	Partially achieved	Largely achieved
Combined 1st and 2nd rating	2.5	2.5	2.5

Participants in the focus group discussions verified these findings with the overarching message that courses were considered of high overall quality, and their perception was that they had acquired significant gains in knowledge and skills from the courses. In comparison with other providers (e.g. WAIFEM), ATI was found to provide training of high quality, with professionalism and excellent administrative arrangements. Learning seemed to have been very strong in particular from the gender budgeting course and the three fiscal policy courses. On the latter, gained learning reported by the participants was on the correlations between budget execution and sustainability and between budget deficit and public debt. Financial programming of cost (how a budget framework is created) and macroeconomic forecasts according to the different models were emphasized as areas where participants felt they had acquired skills. A participant in the tax policy and administration course confirmed that he made improvements and learnt for example, how to ensure stakeholder engagement, and different challenges in the field of tax (e.g. tax payer behaviors, evaluation standards etc.). The case studies have more examples from focus group discussions and interviews that corroborate the findings based on test results and surveys.

Evidence of knowledge and skills being used in work or relations with the Fund is less robust for the above mentioned reasons and based exclusively on focus group discussions and interviews. The picture emerging was more mixed. One participant stated clearly that the courses enabled them to perform the tasks they are assigned better. Participants from one of the case study countries had a tendency to see participation in the courses from "personal benefit" point of view and transversal skills were appreciated as they continue to be relevant if a person changed jobs. In the other two case study countries, there were stronger indications of learning being applied, especially in relation to the gender budgeting course. The course stood out throughout the evaluation in terms of its relevance, coherence and effectiveness and provides a good example of strong synergies between the work of AFW2 and ATI. Consequently, participants in one of the latter two countries reported that they were able to use the training to develop recommendations on gender sensitive budgeting and had presented them to their senior management in the Ministry of Finance. This has led to the development of pilot projects for the development and implementation of gender sensitive budgets in 5 key ministries. In relation to the fiscal policy courses, participants noted that they had used skills for projections in relation to adjusting the country's budget. In the other country from the above two case study countries, a participant in the gender budgeting course reported to have used the acquired knowledge and shared it with other relevant stakeholders at district level and incorporated what she learned in training for others on local level.

As shown in table 4 above, the aggregate ratings indicate that the objective has been between partially and largely achieved. This result seems to be driven mainly by the gender budgeting and fiscal policy courses, which were examples of courses that worked well. Despite its significance for the evaluation, the tax policy and administration course assessment suffered from significant data limitations as there was limited data from the participant database and the difficulties in realizing focus group discussions. There were also data limitations on other courses and particularly for one of case study countries which is an issue that ATI and ICD need to address if the RBM approach is to have credibility. There is also a need for more systematic follow-up to gather information about the use of knowledge and skills to substantiate the findings in monitoring and future evaluations. According to interviews with ATI senior management the main constraints on implementing systematic follow-up on courses are resources and an overriding concern with the difficulties in meeting high demand for courses in member countries. These are valid concerns, but this evaluation finds that effectiveness and impact of ATI training could be strengthened through follow-up in combination with other measures. Strengthening follow-up was also a recommendation of the previous evaluation. To do this in a cost-effective manner follow-up could be done through a semi-automated approach which could be developed by ICD in collaboration with ATI and other RCDCs.

Overall, despite the prevailing conditions, ATI managed to achieve commendable results with virtual training throughout the pandemic. Nevertheless, both effectiveness and post training data collection seem to have been constrained by the rapid shift to a fully virtual learning environment. The impact on effectiveness was a major topic for participants in the focus group discussions (and the evaluation survey) and should both be acknowledged and used to inform future use of virtual and blended approaches to training. Most of the constraints are, however, well-known and referenced in greater detail in the case studies but primarily concern logistical issues such as connectivity, equipment and physical space. A good practice to mitigate this kind of impacts came from Sierra Leone where the IMF resident representative office provided access to a training space for the gender budgeting course outside the work environment and with better internet connectivity. This was greatly appreciated by participants in this course and potentially contributed towards the relatively high effectiveness of this course. A similar effort was made by the Res Rep in the Gambia offering space for participation in online training.

In addition to the logistical issues, participants have also provided suggestions on other areas for improvement. These include: ensuring that selection process is completed at least one month in advance of the course to allow planning, increased use of Africa-specific case studies (fiscal policy workstream), more interactive learning (e.g. with Mentimeter and other tools to engage participants actively), increase interaction between language groups with support from interpreters, bilingual virtual deliveries, informing participants about courses on interrelated topics (e.g. debt sustainability) , incorporate self-study sessions in the course program and give “homework” to enhance learning, strengthen individual interactions and peer engagement after virtual lectures and share information on other participants from the same country.

Some issues relate specifically to the use of pre- and post-course tests which should be considered given their importance for the RBM ratings. Overall, results were not available for a high number of participant pre- and post-course tests. This issue has most likely been exacerbated by the virtual learning environment, but ATI should monitor this closely and seek to bring down the number of errors. While participants agree with the use of the tests, they would like to have access to the results which they do not currently have. This seems like a fair request that might increase dedication and motivation to learning.

Effectiveness of TA revenue administration projects – comparative ratings

Based on the examples and general storyline provided above, the picture that emerges is one of mixed but broadly positive performance in CD to revenue administration (and additional workstreams for training) within the evaluation period although much work remains outstanding and, in some cases, behind schedule. This picture is broadly reflected in the RBM system ratings provided by IMF personnel and the validation of these ratings by the evaluation team, which are listed in Table 5. It is important to emphasize that the evaluation was undertaken based on results achieved by mid-term, leaving time still to pursue targeted outcomes. The levels of achievement by mid-term also reflected the changed operational environment caused by the pandemic.

Ratings provided by staff suggest that the RBM objectives were between partially and largely achieved and almost identical across the three objectives. The evaluation ratings on average verify the RBM ratings being only very marginally lower. The objective to strengthen core tax administration functions was rated lowest on average by evaluators also indicating the only project with any real discrepancy between IMF and evaluator ratings. Even this should be seen in context of several positive examples of progress under specific focus areas above.[32]

Table 5: Comparison of IMF and Evaluation ratings of effectiveness in meeting RBM objectives

	CAPTAC DR			AFW2			
	Country 1	Country 2	Country 3	Country 1	Country 2	Country 3	Average
Tax administration							
Strengthen revenue administration management and governance arrangements.	4,0	1,0	3,0	2,5	2,0	2,0	2,4
Evaluation rating	3,0	3,0	2,5	1,3	2,0	2,2	2,4
Strengthen core tax administration functions	1,0	3,0	3,0	3,0	2,3	2,0	2,4
Evaluation rating	2,0	2,0	2,1	2,0	1,6	2,0	2,0
Tax administration	2,5	2,0	3,0	2,8	2,2	2,0	2,4
Evaluation rating	2,5	2,5	2,3	1,6	1,8	2,1	2,2
Customs administration							
Strengthen revenue administration management and governance arrangements.	3,0	2,7	3,0	2,0	2,0	2,0	2,5
Evaluation rating	3,0	3,4	2,0	1,3	2,0	2,2	2,3
Strengthen core customs administration functions	3,0	2,5	3,0	2,0	2,0	1,8	2,4
Evaluation rating	2,8	2,5	2,3	2,0	2,0	2,5	2,3
Customs administration	3,0	2,6	3,0	2,0	2,0	1,9	2,4
Evaluation rating	2,9	3,0	2,2	1,7	2,0	2,3	2,3

Overall IMF							2,4
Tax IMF							2,4
Customs IMF							2,4
Overall evaluation							2,3
Tax evaluation							2,2
Customs evaluation							2,3

Figure 1 illustrates the ratings across countries under the tax administration project were verified on average but with slightly more variations at the country level. The higher evaluation ratings in the two CAPTAC-DR countries (explained above) are visible as well as lower evaluation ratings. Generally, evaluation ratings were slightly lower under the objective to strengthen core tax administration functions primarily based on limited progress in the area of tax inspections and audit. Two of the CAPTAC DR case study countries and one AFW2 case study country had the biggest divergence overall between IMF and evaluation ratings. This is also a good reflection of the examples provided above with slower progress in relation to tax inspection and audit and the support to excise operational procedures.

[32] It is important to emphasize that particularly two ratings stand out in terms of reflecting the somewhat different approach followed by CAPTAC-DR to report back to the central RBM system. In both cases, the rating does not reflect the level of achievement that the Center has made and in both cases the evaluation rating is substantially higher than the RBM rating. CAPTAC-DR submitted information for clarification on this topic but the data was provided in a format different to the original file that contained RBM ratings for all sampled projects. This does not significantly affect the overall picture although higher ratings would have brought the IMF ratings under tax administration marginally closer to largely achieved. .

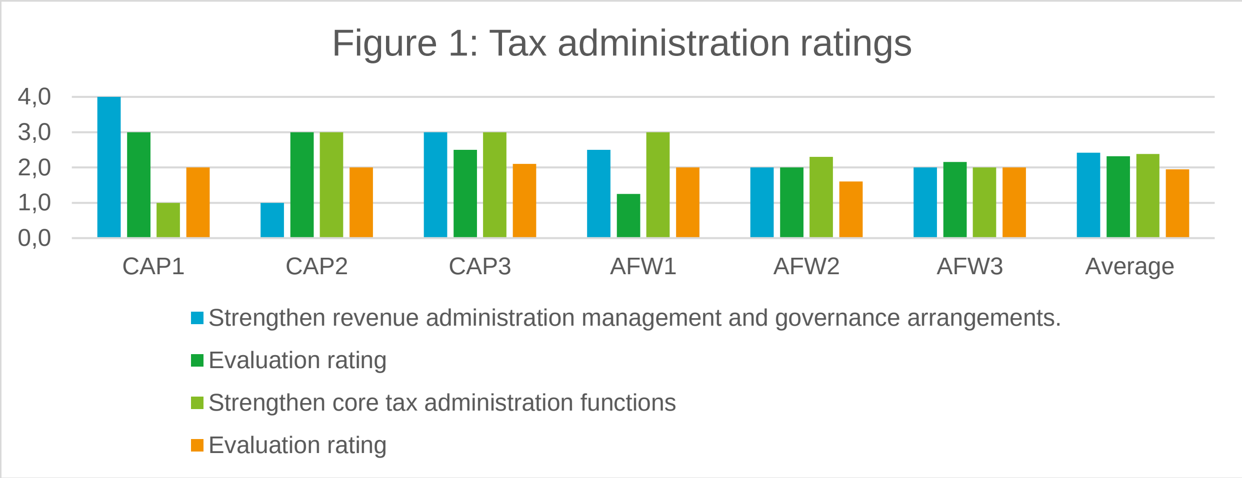
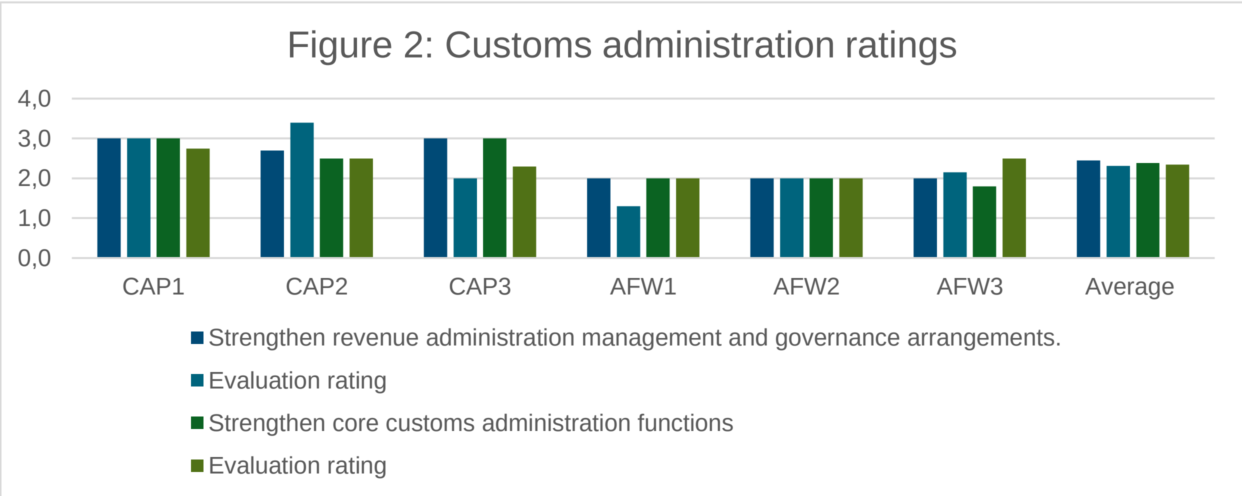


Figure 2 illustrates the ratings across countries in the area of customs administration largely confirming the alignment of IMF and evaluation ratings. Only one country in CAPTAC DR and one AFW2 country had evaluation ratings that were marginally lower than the IMF ratings while one CAPTAC DR country and one AFW2 countries had evaluation ratings that were higher than the IMF ratings (as reported in the RBM system).



From an evaluation perspective it was challenging in some cases to associate and substantiate the assigned RBM ratings with evidence available from TA reports. Sometimes even mapping and matching TA reports with objectives, outcomes and milestones was challenging and required significant effort. It is important in the future that milestone and outcome ratings rely on clear evidence that their achievement can be directly associated with project interventions/activities through a logical results chain. This needs to be done consistently across all RCDCs, and will require additional guidance and training. There is also a need for a systematic, arms-length review of ratings which could be done through some form of backstopping mechanism with ICD.

The effectiveness of TA often relied on the consistency and continuity of its delivery, a tendency that was exacerbated by the Covid-19 pandemic as trust became harder to establish. In one CAPTAC DR case study country, there was a marked difference in the performance between areas that were delivered by the same STX and the area where various STXs were involved over the three years (more details below). This is partly due to inadequate use of tools to increase consistency and requirements of new STXs to use these. In another CAPATC DR case study country, having more permanent accompaniment provided by the LTX in combination with more specific assistance provided by the STX was highly effective. In the third CAPTAC DR case study country, LTXs and STXs cultivated a close dialogue with and accompanying the beneficiaries at every stage of the project lifecycle are considered among the contributing factors that enabled high performance and impactful results in the modernization of the customs process and taxpayer services.

In some case countries, TA reports included a table summarizing actions agreed with the authorities. However, in subsequent TA reports, this table simply introduced new actions without reporting on the status of the implementation of the action from previous missions. LTXs and STXs noted that they had received varying guidance on how to use these tables and the most recent guidance was to not use them for follow-up. This represents a missed opportunity to ensure consistency and continuity in the delivery of TA and is low-hanging fruit in terms of improvements. In the case of the above country where there was a marked difference in the performance between areas that were delivered by the same STX and the area where various STXs in the CD delivered in the area of tax inspection and audit, this would have forced the STXs to connect with what the predecessor had done and avoided constant restarts and reinvention of the wheel. Systematic follow-up from previous missions in a granular manner (using the table) should be reintroduced as a requirement and good practice for all TA missions and reports. Importantly, this does not mean plans should be set in stone but rather a living tool to calibrate expectations over time in an informed manner. Each mission report should provide a brief explanation as to why a certain action was not completed and provide a new recalibrated list of actions for the next mission.

2.4 Efficiency, Cost-Efficiency and Cost-Effectiveness

During the evaluation period, the IMF transitioned from earlier financial systems such as the Travel Information Management System (TIMS) and the Analytical Cost Estimation System (ACES) to the Capacity Development Management and Administration Program (CDMAP) which aims to integrate planning, budgeting, managing and monitoring of CD. CDMAP has been rolled out progressively since 2020 and in April 2021 the module on CD execution and resource management went into implementation which means cost data is available for FY22. The evaluation uses data for FY22 exclusively to inform the efficiency dimension of the evaluation, given the complexities of obtaining and comparing relevant cost data from different IMF data management systems and the skewness of the data from before and during Covid. While the system is still new and significant "teething issues" are to be expected, the intention has been to apply a forward looking approach to inform the efficiency dimension of the evaluation.[33] The cost efficiency analysis is in line with the IMF's 2018 CD Review that called for more attention to the cost efficiency of different modalities of CD delivery and increasing interest from donors in cost-effectiveness and value-for-money assessments as well as the 2022 IEO evaluation of IMF CD.

The assessment uses metrics such as average cost of a capacity development day (CD Day) and average cost of a CD participant day (CDP Day) for comparisons across the three RCDCs. The main purpose of this approach is to enable easy identification of outliers in the cost data e.g., a much higher cost per CD Day of a tax project than a customs project in the same country or across similar countries. When an outlier is identified this can be picked up for further analysis to review what drives the extraordinary cost levels and help inform decisions about the project.[34] Finally, the evaluation has made an effort to take forward the approach proposed by the IEO evaluation to assess costs against activities or outputs (efficiency) and outcomes (effectiveness) relying on RBM milestone and outcome ratings. The objective is to explore how data on the outputs and outcomes of CD from this evaluation can be considered alongside estimates of the costs of the CD delivered and produce simple indicators (or metrics) of cost-efficiency—that is, cost per unit of CD output—and cost-effectiveness, i.e., cost per unit of CD outcome.[35]
Again, this approach is more illustrative than conclusive due to the limitations of the cost data but also that the projects are still under execution.

[33] Efficiency analysis has been a challenge in many previous evaluations often due to lack of access to comprehensive cost data.

[34] These metrics have been used in the past by ICD to provide information to development partners financing IMF CD.

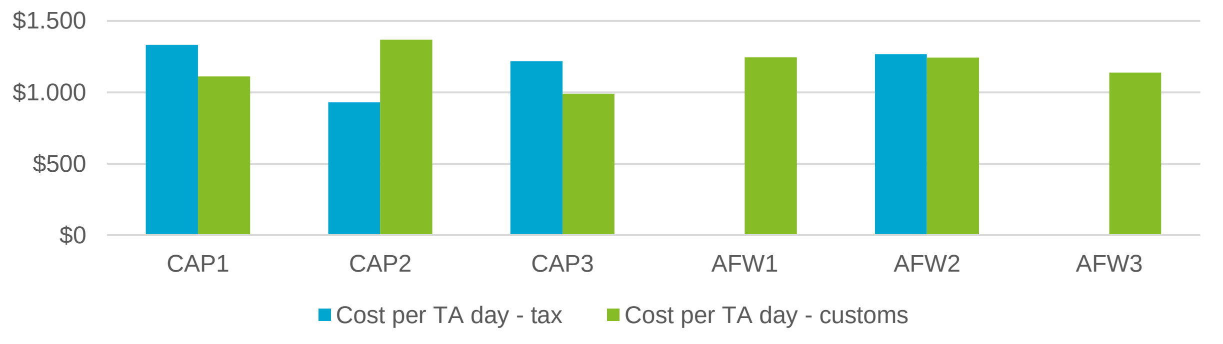
[35] This text has been adapted from Jensen and Kell (2022). While this evaluation lent itself to taking this analysis even a step further and look at the cost of quantifiable result in revenue administration, time and resources available did not allow taking the analysis to this level and is only compared to ratings.

The efficiency analysis is based on FY22 data as received based on a request to the IMF to provide cost data in dollar terms in a disaggregated manner by courses (training) and country and projects (TA) with only one major outlier being reviewed with the ICD financial department and the RCDCs. The data includes cost for delivery (classified as duty-station work due to remote delivery) and backstopping. It is the first time an external evaluation has attempted this kind of analysis and as CDMAP is still a recent instrument and implementation was affected by the continued virtual delivery, data reliability issues (misclassifications etc.) most likely explain the outliers identified.

Cost of TA projects

With these caveats in mind, the evaluation obtained data on the total cost of projects per country and number of resource days in CDMAP. In terms of total spend, the costliest project by far was tax administration in one of the CAPTAC DR case study countries, which amounted to almost US$190,000. Two projects were just below US$100,000 (a tax and a customs project) and three projects ranged between $30-40.000 (one tax and two customs projects). These differences were mostly explained by traction with substantial amount of support being offered to the country (140 resource days).

Figure 3: Cost per CD day across the case countries



Please note that "outlier" data from two AFW2 countries has been omitted from the chart.

In relative terms, the evaluation found that the cost of a CD Day for the tax and customs projects in the six RTAC case countries fell within a range of roughly US$1000 to US$1300 per day. One major outlier was identified (not included in figure 3) in the tax project for one of the AFW2 case study countries, where available data indicated a cost per CD Day of US$2,195 greatly exceeding the range identified in other case countries and for customs in the same country. The divergence from the average range seems to be too marked to be explained by the higher costs of CD delivery to be expected in this country and evaluators suspect this is a classification issue (e.g. in the number of virtual mission days). This is a good example of a project that should be subjected to further analysis. In the case of another country in AFW2, no cost was registered for CD delivery (duty-station work) of the tax project, but cost was registered for backstopping. Inquiries about the missing data revealed that the planned activities had been postponed but did not shed full light on why backstopping costs were still filed (although technically the backstopping could have led the activity to be postponed).

In another AFW2 case study country, the cost of a tax and customs CD Day was almost identical. In CAPTAC-DR, delivery of tax projects was more costly than customs projects in two out of three case study countries but in the third case study country, the customs project was substantially more expensive to deliver than the tax project.

In addition to the costs associated with the delivery of CD overall, CDMAP singles out cost for delivery and backstopping. The backstopping costs varied across countries and projects. Interestingly, the cost of backstopping tax projects in CAPTAC-DR were very similar around 15 percent of the total project cost but increasingly expensive in AFW2 for two out of three case study countries (note that the third case study country´s data is included here to illustrate the relationship between the delivery and backstopping costs although the data is still needs to be verified). On customs, another interesting trend appears with the cost of customs backstopping in AFW2 being consistently low around 13 percent of total project cost. Backstopping of customs projects in CAPTAC-DR were consistently higher than the AFW2 average and much higher in the case of a particular case study country. In AFW2 it is possible that the more standardized approach to CD in customs has lowered the cost of backstopping.

Under normal circumstances, the main cost drivers of CD are the cost of personnel and travel.[36] Given the remote delivery of CD during the period under evaluation, there were no travel costs. Different personnel decisions also affect the cost of projects. Personnel costs are higher for LTXs who are hired on longer contracts and carry significant responsibility for delivering IMF CD including backstopping of STXs. The latter are hired on short-term contracts on a consultancy basis and hence do not qualify for benefit packages. Typically, STXs are therefore less costly than LTXs. This tendency was reinforced by remote delivery as STXs typically incur higher travel costs as compared to LTXs who are based in the region. A project relying heavily on LTX implementation can therefore be more costly than an STX delivered project. Also, the number of experts involved in delivery affects the cost of the project.

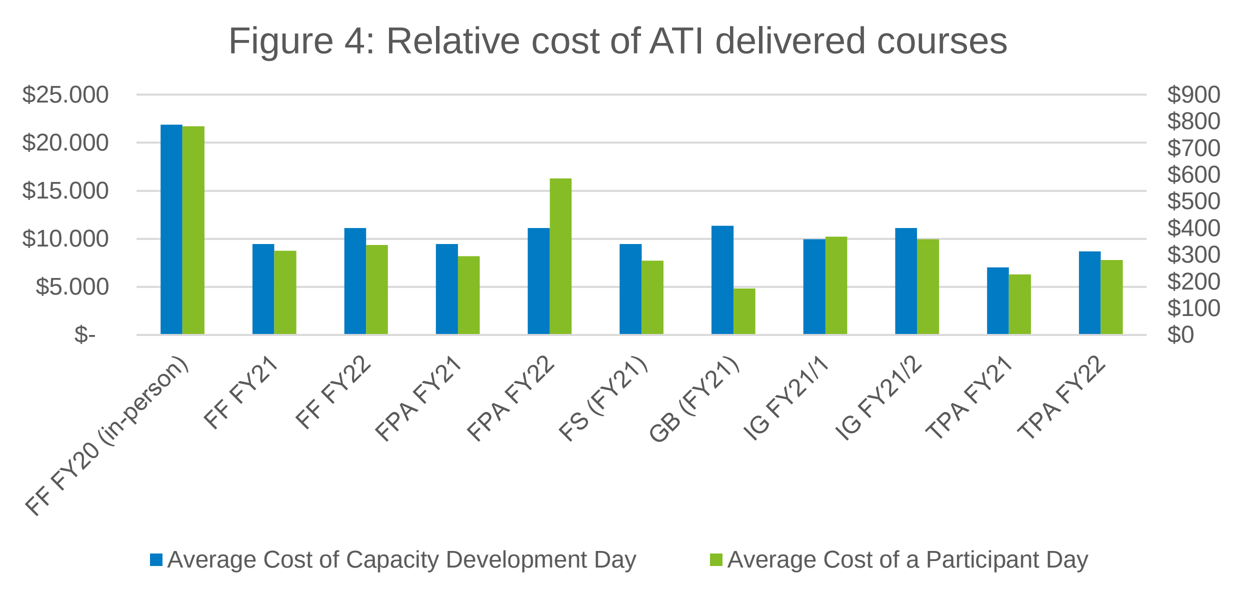
The STX/LTX ratio (number of STX days divided by the number of LTX days) calculated based on CDMAP data showed that for tax projects in AFW2, the ratio was 3.6 while for customs it was 4.5 implying that STXs were used more extensively for customs projects. In CAPTAC-DR, the ratio for tax projects was 2.3 and the ratio for customs was 7.5, indicating a substantially higher involvement of LTXs in the delivery of tax projects than customs projects. This might reflect the transition from one LTX to another although this could not be verified.[37]

Cost of courses

The cost of the delivery of the six sampled courses (of which some were delivered more than once during the evaluation period) varied with the only in-person course delivered in November 2019 (FY20) clearly being the costliest course in the evaluation period. This is a clear illustration of the difference between the cost of in-person and virtual delivery as the cost of travel of participants and lecturers made up 43 percent of the total cost of the course. However, other drivers of the cost of courses include personnel (number of lecturers involved), duration of the course, translation and interpretation. The relative cost in terms of average cost per participant day is significantly influenced by the number of participants in the course. Based on these factors, the average cost of a CD Day was relatively stable around US$9-10.000 for virtual courses although the cost of the tax policy and administration courses was lower due to the involvement of only two lecturers as opposed to the standard three (see Figure 4). The in-person fiscal frameworks course was twice that amount. In terms of the average cost per participant day, the FY22 fiscal policy analysis course was significantly above average as only 19 persons participated. The gender budgeting course, on the contrary included 65 participants bringing the down cost per participant day substantially. This is very interesting as the gender budgeting course also came across as an example of a successful course in terms of effectiveness. The higher number of participants does not seem to have affected effectiveness, which indicates very good value for money from this course.

[36] And these are determined by IMF Board-approved policies covering salaries, benefits, allowances and travel class.

[37] Also, the CDMAP data had to be corrected as three LTX missions were classified as STX.



Based on the cost of a participant day, the allocation of resources by country can be calculated based on the number of participants per country. This calculation shows that the three case countries each benefitted from 3.8 percent (Sierra Leone), 4.4 percent (Ghana) and 4.1 percent (Côte d'Ivoire) of the resources spent on the sampled courses. This is above average, reflecting that these countries were selected as cases for this evaluation among other criteria based on the number of participants in the sampled courses.

Cost efficiency and cost effectiveness

In tables 6 and 7 below, a number of cost, efficiency and effectiveness metrics that can be combined to produce a basic cost-efficiency and cost-effectiveness analysis of the revenue administration projects in the case countries. In this exercise, the outcome ratings from the RBM system and the evaluation ratings are combined to provide a balanced average rating. It is important to note that this approach has an even more preliminary nature than the above given that it is harder to establish robust connections between inputs and outcomes than between inputs and outputs.[38]

The tax administration metrics are interesting in the sense that two out of six countries were outliers or had data issues that should be addressed before any further analysis is conducted. Secondly, one of CAPTAC DR case study countries had the most extensive LTX involvement, which might be reflected in the marginally higher cost per CD day. Four out of six countries used only STXs for delivery. Milestone and outcome ratings corresponded broadly except in the above CAPTAC DR case study country where the outcome rating was significantly higher than the average milestone ratings.

Table 6: Cost, efficiency and effectiveness metrics for tax administration in CAPTAC DR and AFW2 case study countries

TAX ADMINISTRATION	CAPTAC DR			AFW2		
	COUNTRY 1	COUNTRY 2	COUNTRY 3	COUNTRY 1	COUNTRY 2	COUNTRY 3
Total cost	186.465	41.810	96.331	19.186	112.857	63.460
Backstopping cost	29.917	5.276	13.532	19.186	26.974	21.761
Resource days	140	45	79	32	89	19
LTX	34	0	0	0	16	0

[38] Note that this analysis also draws on average milestone ratings extracted from the RBM system.

STX	106	45	79	32	73	19
Ratio	3,1	STX only	STX only	STX only	4,6	STX only
Cost per CD day (FY22)	1.332	929	1.219	Outlier	1.268	Outlier
Outcome rating	2,5	2,0	3,0	2,8	2,2	2,0
Milestone rating	1,7	1,8	2,5	2,5	2,5	2,3
Evaluation rating	2,5	2,8	2,4	1,6	1,8	2,1
Average rating	2,5	2,4	2,7	2,2	2,0	2,0

Figure 5 plots the average cost per CD day against milestone and outcome ratings providing an illustration of the cost-efficiency and cost-effectiveness of the four countries that have sufficient data available. The chart suggests that in terms of cost against milestone ratings (cost-efficiency), one of the AFW2 case study countries performed best but in terms of cost against outcome ratings (cost-effectiveness) one of the CAPTAC DR case study countries performed quite well. The other two CAPTAC DR case study countries also performed well on cost-effectiveness but the higher cost per CD day suggests lower cost-effectiveness. It is important to note that the highest overall spending was in in two countries (one in AFW2 and CAPTAC DR respectively) and the strong milestone rating (efficiency) could indicate that cost-effectiveness will improve in the AFW2 one in the final two years of the current phase. The CAPTAC DR one needs to address the slightly lower milestone ratings but these also reflect that CRM has moved into the next level of complexity. Indeed the latter is a major issue for further consideration in that high effectiveness ratings in the current system can reflect unambitious indicators and/or low ratings and could be an indication of slower progress in more complex reform areas. Overall, the current approach to RBM in the IMF could benefit from a metric indicating the level of complexity of reforms being pursued.

The customs administration metrics are interesting in their completeness suggesting no major outliers or data issues. Secondly, only one country in CAPTAC DR case study countries relied on STX involvement while one country in the AFW2case study countries had the highest level of LTX involvement. Milestone ratings were generally higher than the average rating except in one country in the CAPTAC DR region where they were marginally lower. Cost per CD Day was highest in another country in the CAPATC DR region which is a bit surprising as only STXs were used for delivery.

Table 7: Cost, efficiency and effectiveness metrics for customs administration in CAPTAC DR and AFW2 case study countries.

CUSTOMS ADMINISTRATION	CAPTAC DR			AFW2		
	COUNTRY 1	COUNTRY 2	COUNTY 3	COUNTRY 1	COUNTRY 2	COUNTRY 3
Total cost	98.856	32.819	38.633	79.632	79.576	69.454
Backstopping cost	13.358	8.796	5.948	9.417	9.417	7.778
Resource days	89	24	39	64	64	61
LTX	17	0	5	10	10	11
STX	72	24	34	54	54	50
Ratio	4,2	STX only	6,8	5,4	5,4	4,5
Cost per CD day (FY22)	1.111	1.367	991	1.244	1.243	1.139
Outcome rating	3,0	2,6	3,0	2,0	2,0	1,9
Milestone rating	3,3	3,7	2,1	2,6	2,7	2,5
Evaluation rating	2,9	3,0	2,2	1,7	2,0	2,3
Average rating	2,9	2,8	2,6	1,8	2,0	2,1

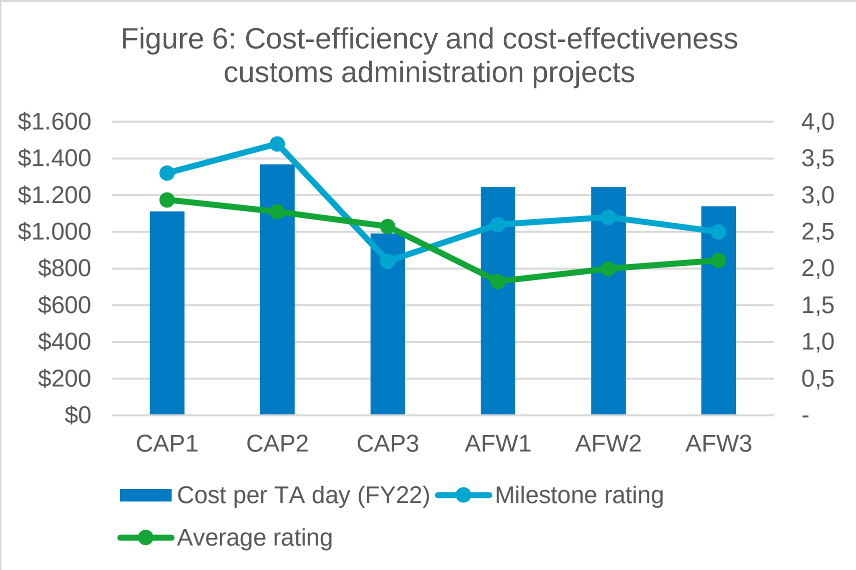


Figure 6 plots the average cost per CD day against the milestone and average outcome and evaluation ratings providing an illustration of the cost-efficiency and cost-effectiveness of the six case countries. The chart suggests that both in terms of cost-efficiency and cost-effectiveness, the one case study country in the CAPTAC DR region performed best, justifying the relatively high cost per CD day. However, another case study country in this region also performed well at a lower cost. Overall, the picture is quite encouraging with higher milestone ratings in all countries except one of the case study countries in CAPTAC DR at reasonable levels of cost per CD day.

The analytical approach and metrics presented above need to be interpreted very carefully and should always be complemented by careful qualitative analysis of project performance along the lines of this evaluation. Moreover, this analysis is based on limited data from the first year that CDMAP has been able to produce cost data at this level of granularity. However, as cost data matures and becomes available for more financial years, this type of analysis will be very beneficial for the IMF, for example to help inform strategic choices and prioritization of allocation of resources. The evaluation team believes that the explorative approach applied here can help inform and improve future evaluations through better cost-efficiency and cost-effectiveness analysis within and across different workstreams.

2.5 Impact and Sustainability

This section addresses preliminary indications of the impact and potential sustainability of CD to revenue administration (and gender budgeting). The mid-term evaluation approach looking at the first three fiscal years of the current RCDC faces does not lend itself to evaluating impact as all the projects and course activities evaluated are still ongoing and results should be expected mainly at the end of the five-year period.[39] There is also the attribution issue mentioned in the limitation sections that affects how impact is analyzed. The potential effect from CD to revenue administration on revenue mobilization is for example hard to isolate from other factors such as developments in the macroeconomic context. It should also be emphasized that at the outcome level the RBM framework for tax administration uses TADAT indicators. This implies that the best evidence of impact would be ascertained through repeat TADATs or quality assured TADAT self-assessments. This adds another compelling reason for carrying out such assessments and for ICD to consider if this represents the best use of indicators to inform the RBM approach. However, this evaluation has identified five thematic impact stories based on projects and courses that seem particularly promising and have potential to produce impact. It also points to three case countries where there is a strong potential for impact over the next years. These findings are combined with considerations about sustainability in terms of the expected continuation of the projected results over the medium and long term.

[39] Indeed, a longer timeframe could be necessary to evaluate impact as sustainable capacity development is a long-term effort but this would require an entirely different evaluation strategy to the one used in this evaluation.

Impact stories

The main impact story is based on tax administration support to sector-based audits particularly in the Gambia but with work on this progressing both in Sierra Leone and Ghana too. The TA reportedly had a direct impact on revenue collection, in particular in the telecom sector, where it contributed to a significant increase in tax revenue from the telecom companies. This exercise was expanded to the tourism sector with less success as this sector suffers from lower capacity and was hit hard by the Covid-19 pandemic. TA was based on a step-by-step methodology jointly prepared by AFW2 and the Gambia Revenue Authority, including data / gap analysis, data matching using third-party information. The examples of what worked well in this area in Sierra Leone and Ghana suggests that impact from this approach could be achieved in these countries too although especially in relation to Ghana, the size of the economy and complexity of the associated sectors must be considered in relation to expectations. Building on the cost-effectiveness analysis above and more granular data on the size of the increase in revenue collection from the telecoms sector and overall, the analysis could be extended to a more comprehensive value-for-money assessment. Even a marginal increase in revenue to GDP would greatly exceed the investment in CD to the country but a more robust assessment would look at the changes in revenue collected from the sector before and after the intervention and seek to isolate other impacts.

Other stories were less clear in their potential impact but CD to customs administration in Costa Rica and the Dominican Republic led to the adoption of new customs laws. In the Dominican Republic, TA supported the preparation of a comprehensive new customs law (replacing a 68-year-old version) that set a path to strengthen post-clearance audit capacities and modernize and expedite customs procedures, including electronic declarations and provisions on advance rulings. Progress in customs was generally good and this could result in significant impact in relation to revenue collection and trade facilitation. In Costa Rica TA was highly relevant to the development of the Customs Law reform in May 2022. This is expected to further promote trade facilitation and to help increase productivity growth in trade-related sectors as well as improve revenue collection through more effective risk-based enforcement. It should be possible to identify metrics associated with the potential impact in this area too both in terms of increased revenue collection from customs and trade or specific key performance indicators based for example on time release studies.

Another story related to support by AFW2 to customs integrity frameworks in the three case studies. This area of support is probably the most challenging and complex identified with potential impact as it deals with efforts to tackle corruption in customs and would benefit from close coordination with the LEG around the IMFs new governance priority area. The adoption of these frameworks has been relatively slow but in the case of one of the case study countries, it provided an example of strong responsiveness of AFW2 to opportunities provided by a changing operating environment. While support had discontinued due to lack of demand, a corruption scandal in customs resulted in the special prosecutor’s office requesting the adoption of an integrity framework. AFW2 responded swiftly to the window of opportunity and provided agile support outside the established work program. The framework was developed by a deadline set by the prosecutor and a team was put in place to implement the policy, but it was not possible to gather further information about its adoption.

A positive story from an operating environment that has challenged the delivery of CD, concerns the support to compliance risk management in one of the CAPTAC DR case study countries. While progress was initially very slow in this area, substantial progress was made after a recalibration of expectations, formation of a core team, exchanges with neighboring countries (peer-to-peer), positive messaging and building of trust. These are all necessary prerequisites for impact in relation to the use CRM to improve registration, timely filing, timely payment, and correct reporting and identify drivers of non-compliance. CRM made good progress in the case of Costa Rica where authorities came around to embrace the approach after recognizing its relevance. It is now applicable to all tax control processes carried out by the tax authorities and has led to a more progressive and equitable approach. Thanks to the approach, it would be easy to identify tax administration metrics across the registration, filing, payment and reporting cycle to inform cost-effectiveness and value for money assessments.

Finally, the ATI delivered gender budgeting course constitutes an impact story related to training. While anecdotal, there are indications that knowledge and skills from this course has been put into practice and led to concrete gender budgeting initiatives in some of the case countries. Participants from Sierra Leone reported that they were able to use the training to develop recommendations on gender sensitive budgeting and presented them to the management. The subsequent course led to the development of pilot projects for the development and implementation of gender sensitive budgets in 5 key ministries. In Ghana the content of the course had been applied to gender budgeting efforts at the local level. While only indirectly relevant to the thematic focus of the evaluation, the gender budgeting course was highly relevant to the case country contexts and constitute an important priority area for the IMF. It represents an example of strong synergy of training and TA provided by ATI and AFW2. Moreover, this course was highly cost-efficient due to the virtual delivery, relatively high number of participants and low cost per participant day. It would be interesting and relevant to carry out a more in-depth study on the use of the learning from this course across participating countries to inform future work in this area.

Promising case countries

Overall, countries that stand out in terms of the potential for impact in the remainder of the current phase are Costa Rica, the Gambia and Ghana. In Costa Rica and the Gambia, RCDCs already have very good traction both in tax and customs administration, demand for support and commitment to reform. Costa Rica has an ongoing ECF program with the IMF and the Gambia passed a final review of its current program in June 2023. Both had strong complementarity between revenue administration reforms supported by the IMF under the programs including structural benchmarks with targeted TA to support implementation and both countries were characterized by stable political contexts. The Gambia also receives additional support from the IMF resident advisor in tax administration. In both cases, milestone ratings exceeded the average outcome and evaluation ratings (except for tax administration in Costa Rica) reflecting good progress and suggesting that the latter could pick up. The success in the area of sector specific tax audits in the Gambia is also likely to increase appetite for further strengthening. One of the AFW2 case study countries is included more as a dark horse, but the country has a strong potential to increase impact. A program with the IMF has just been agreed with very strong emphasis on the need for revenue mobilization. Commitment to reform has increased and as a consequence the country will also receive substantial support through the RMTF. The country also has strong incentives to increase its lackluster revenue to GDP figures. All these also form a good foundation for sustainability of the promising practices identified above.

Box 2: Drivers of change

A few drivers of change were identified in the evaluation and summarized in this box.

Political will and commitment to reform are key factors that have contributed to driving change in areas identified as having potential impact. While this is a well-known issue it is important to factor into the planning and delivery of CD to get the greatest possible impact and value-for-money for the scarce resources invested in CD. Moreover, and as noted in the coherence section, political will is not static and can be influenced by and through engagement of stakeholders both within and outside government. This can help increase the visibility of key decision-makers' commitment to necessary reforms and holding them accountable for key decisions. One case study country in the AFW2 region is a good example of how commitment to different aspects of revenue administration has changed over time. Clear diagnostics, joint identification of key issues and priority areas as well as development of plans to address them can contribute to change.

Consistency of CD delivery over time with good sequencing of missions, conscious effort to build from previous recommendations, follow up and limiting changes in personnel are also critical factors of driving change. Finally external shocks most evidently the impact of the pandemic and the war in Ukraine on revenue mobilization but also events like the cyberattack and a change of government in one case study country in the CAPTAC DR region shows appetite for reforms could change with such events.

The ability to demonstrate tangible impact as in the case of sector tax audits can be a key driver as well as peer-to-peer exchanges at senior management level to share such results. Both technical and soft skills of experts can be central drivers of change as they help build credibility and trust from key partners at country level. Change management skills such as recalibrating expectations and identify alternative workable solutions as well as communication skills and cultural competencies are also critical.

Key sustainability issues

As noted, sustainability is an extension of impact and can therefore only be assessed on a very preliminary basis. Some of the key challenges (and potential mitigation) include: i) absorption capacity and general skill levels within institutions supported in all country contexts, which is further impacted by staff turnover, and to be sustainable continuous assessments of critical human resource capacity gaps are necessary; ii) reliance on technology, costly data analysis software, as well as the existing IT knowledge and ability to manage digital systems and reluctance to utilize such software; iii) working in isolation with dedicated teams that while a good entry point need to seek continuous endorsement by senior management; and iv) limited post-course follow-up with participants which reduces the sustainability of training. A cluster of activities that could strengthen the sustainability of training includes increased collaboration with development partners, partnerships with domestic (administrative or PFM focused) training centers, provision of practical tools to facilitate feedback, sharing of learning and training of trainers approaches.

2.6 Conclusion – Project and Course Level

CD to the revenue administration workstream provided by the three RCDCs covered by this evaluation was broadly aligned with context, development plans and strategic plans of government revenue administrations in all the case study countries. Both TA and training was relevant to the countries context, and this was retained through an unprecedented time in history due to the Covid-19 pandemic. However, while TADAT was a very useful tool in most case countries, up-to-date diagnostics and capacity needs assessments were lacking to better inform the planning of CD delivered and create synergies between TA and training.

Despite of complexity in the delivery of CD to revenue administration the CD offer was relatively coherent internally, especially in countries with an IMF program. The main instrument for planning CD strategically at country level was however either not available or dated in most case countries. They also were found to not clearly articulate structured content training as an integral component of the CD to revenue administration. This constrains a more complementary and strategic approach to CD delivery to revenue administration.

Proactive Resident Representatives play a critical role both in relation to internal and external coherence while engagement and consultations on TA and training took place, it was not done in a formalized and systematic manner. RCDCs invest considerable resources in making information available to development partners but this falls short of the expectations in relation to the expected depth and details of information sharing. LTXs and STXs should be more involved in sharing information through debriefings with interested development partners. Countries rarely use a coordination forum for partners on domestic revenue mobilization which the RCDCs could help promoting more proactively.

The effectiveness of CD to revenue administration in the first three years of the phases under evaluation has been reasonable, particularly when considering the impact of the Covid-19 pandemic. At the end of the evaluation period, with two years remaining of the current phases, workstream level ratings averaged 2.4 (the same for customs and tax projects) between partially and largely achieved. This level of achievement was largely confirmed by the evaluation with the independently assessed ratings only marginally lower at 2.3 (2.3 for customs and 2.2 for tax projects). In the absence of RBM ratings for the sampled courses, the evaluation found the level of effectiveness to be 2.5, between partially and largely achieved. Areas of work that stood out overall included support to compliance risk management in CAPTAC-DR, sector audits in AFW2 and the ATI delivered course on gender budgeting, which was also supported by AFW2. There were strong indications that this average level of achievement of outcomes would continue to improve over the remainder of the current phases
The indicative cost-efficiency and cost-effectiveness analysis calculated the average rates of cost of a CD Day roughly between US$900-1400 and identified outliers against these averages to be checked in two case study countries. The indicative cost-effectiveness analysis suggested that in terms of the tax projects, the Dominican Republic performed best while for customs projects, Costa Rica performed best. This type of analysis needs to be further developed however as CDMAP and RBM data mature and becomes more consistently reported. A clear lacking metric to inform such analysis is the level of complexity of the reform effort.

There is significant potential for the revenue administration projects to have an impact over time. For example, the tax administration project in AFW2 has a clear potential to help increase revenue from the sectors targeted for sector audits and risk-based enforcement of new customs legislation in the CAPTAC-DR region could have a similar impact. Good progress was also made on compliance risk management in CAPTAC-DR which could improve tax administration across the registration, filing, payment and reporting cycle. The gender budgeting course led to concrete gender budgeting initiatives in some of the case countries in a cost-effective manner. Generally, there is strong potential to identify metrics that could inform an in-depth impact assessment after the end of the current phases. Case study countries that stood out in terms of potential impact were Costa Rica, the Gambia and Ghana.

3 PART II: EVALUATION OF STRATEGY AND OPERATIONS MANAGEMENT OF THE THREE RCDCS

This part of the report provides a summary of the findings from the three entity level evaluations carried out as part of the thematic pilot evaluation structured along the lines of nine evaluation areas.

3.1 Role of the steering committees

On balance, while room for improvement remains, the three RCDCs have made efforts to strengthen the ability of SCs to provide strategic guidance, help set priorities and provide oversight. The pandemic introduced new ways of collaborating that with time could contribute to more regular and detailed engagement and other adjustments could be made in line with the findings of this evaluation.

The Steering Committees (SCs) of ATI, AFW2 and CAPTAC-DR managed to deliver broadly on their advisory role on strategic guidance and oversight even in the difficult context of the Covid-19 pandemic.[40] The FY20 SC meetings were cancelled as the Centers were not yet in a position to carry them out virtually. Inputs from SC members were instead collected bilaterally. The FY21 SC meetings were held virtually in June 2021 and the joint ATI/AFW2 FY22 SC meetings were held back-to-back in June 2022 applying hybrid formats while the CAPTAC-DR FY22 SC meeting was held virtually in June 2021.

The hybrid format boosted participation with more participants for example of IMF staff and development partners with mixed experiences across the RCDCs in terms of enabling in-depth discussions of key issues. In some cases, a relatively compressed agenda with many presentations, left limited time for discussions while in others this was less of an issue. In the case of AFW2, member countries and not necessarily AFW2, took on stronger leadership roles and reversed the trend of development partners dominating the discussions. In CAPTAC-DR virtual meetings were found to be effective, allowing discussions of budgetary constraints and new CD delivery modalities during the pandemic. The ATI SC endorsed various key decisions throughout the pandemic. Although challenging for ATI given the size of the membership, more could be done to facilitate discussions about programming, budgets, results, cost-effectiveness, and impact. Challenges related to high turnover of country representatives are persistent in both AFW2 and ATI.

In AFW2, the focus of SC attention in the last two meetings have been on topics of interest to members rather than the workplan and results (including RBM). In CAPTAC-DR the Centre had not provided hands-on training or supplemented tutorials to SC members on RBM to strengthen their ability to appreciate and contribute actively planning and oversight of this approach to CD delivery. The limited number of meetings was found to be a constraint in all three regions and decisions have been made to realize bi-annual SC meetings in ATI and AFW2. The biannual meetings could be further enhanced by the introduction of specialized subcommittees with a rotating membership to deal with more technical issues complemented by working groups on a need's basis. This should be complemented by improving members' access to relevant information particularly on RBM metrics and availing it in an easily accessible manner. The latter seems currently to be lacking across the RCDCs (and particularly in CAPTAC-DR). The information sharing could build on the experience with Partners Connect and provide relevant data and documentation through a single central repository that is shareable and secured in a cloud software platform. Data and information should include but not limited to briefing notes, diagnostic studies, scoping missions, TA reports, back-to-office reports, country reports and RBM databases.

40 The SCs are not mandated to make binding decisions on behalf of the RCDCs, and they do not have fiduciary responsibility.

AFW2 and ATI held a joint SC meeting with a thematic focus on revenue administration in 2022 which catered for participation of revenue authority representatives. This allowed ATI to have more direct engagement with members and the combination of this approach with increased direct outreach to member countries represents a promising approach to build stronger relationships.

3.2 Coordination

The evidence gathered in internal CD coordination provides a picture of a relatively coherent CD offer to member countries despite a highly complex approach to delivering CD in revenue administration. IMF RCDCs have a complex governance structure with members and financing partners having an advisory role, area departments (AFR and WHD) responsible for determining the overall strategic direction and priorities, CD departments (FAD and ICD) being responsible for ensuring the quality and consistency of TA and ICD for training and the RBM approach. Moreover, big trust funded programs like TADAT and RMTF are critical vehicles for CD in revenue administration, adding to the overall complexity within this workstream.

Regional strategy notes (RSN), country strategy notes (CSN), article IV and staff reports, program reviews, country reports, the RBM framework and CDMAP are key instruments in setting the strategic direction and implementation of CD, especially in relation to the objectives to ensure complementarities between IMF programs, surveillance and CD, and strengthening synergies between training and TA at the country level. There is significant room for improvement in the extent to which these instruments help to achieve the two objectives. RSNs are too generic to be useful at country level. CSNs are not regularly updated and rarely factor in the important complementary role of training. Article IVs have some coverage of CD but only provide superficial summaries of country CD strategies with almost no attention to training. This situation impacts negatively on strengthening synergies between the structured content training delivered by ATI and TA delivered by AFW2.

The division of labor with HQ leading on strategy and policy issues concerning CD and the RCDCs being responsible for implementation works reasonably well although the lines between the two can get blurry. Key personnel and reporting lines between ADs and RCDS are mission chiefs, center directors and resident representatives (Res Reps). For CD departments and the RCDCs key relationships are between the HQ based backstoppers and regionally based long-term experts (LTX) and lecturers. The backstopping process works well, is fluid with regular communication. ATI and ICD coordination is strong through joint development and delivery of courses and backstopping. Invitation of Center representatives to participate in program or surveillance missions has been applied and worked well in sustaining internal coordination. Likewise, key HQ personnel participate in the SC meetings of the Centers.

The move to virtual and hybrid working has facilitated coordination. The role of res reps is critical for both internal and external coordination and follow-up at country level. In the CAPTAC-DR region there is a regional Res Rep and recently there is also a country-based res rep in Costa Rica. A regional Res Rep cannot cover as much ground at country-level as the ones who are country-based and countries without an Res Rep are at a disadvantage in terms of coordination. Res Reps are occasionally but not systematically consulted by lecturers on selection of participants in ATI training. Clarifying this as a clear requirement for all lead lecturers (both ATI and HQ based) and require inclusion of a paragraph on the outcome of this consultation in back to office or end-of-course reports would help strengthen coordination and effectiveness of ATI delivered training at country level.

The RCDCs have broadly avoided duplication and overlap of assistance with other development partners supporting revenue administration and DRM. AFW2 and CAPTAC-DR have maintained an open dialogue with development partners and regional organizations both in West Africa and Central America through some online meetings, providing access to CD material through Partners Connect and through some technical working groups for example on trade facilitation. ATI has a handful of close partners across the region (including regional bodies) and have collaborated with UN Women in the development of a course on gender equality and macroeconomics. A planned partnership with OECD about international taxation that would have increased the relevance of the course offer in the thematic focus area of this evaluation did not materialize.

Despite the generally positive approach to coordination with external partners, as already pointed out in the coherence section, this is an area that could be further strengthened. The RCDCs have reasonably good communication and outreach strategies with various products and presence on social media. However, more could be done to reach out to the local media, academia and other relevant civil society organizations (CSO) dialogue platforms to further promote the visibility and importance of CD to national development and sustainable development goals.

3.3 Funding model[41]

Overall, the Centers are sufficiently funded for the current phases with significant cost savings from the impact of Covid-19. At the end of the evaluation period AFW2's phase II financing requirement of US$46.5m faced a gap of US$13.7m. However, by FY22 only 58 percent of the budget was spent. Member/financing partner split was US$7m/US$25.9m. CAPTAC-DR's phase III financing requirement of US$40m still faces a financing gap of US$6m. However, also in CAPTAC-DR there were savings in relation to the planned budget. Member/financing partner split was US$9.5m/US$30.5m. ATI financing requirement for phase II was US$34.4 with a gap of US$8.9m between budget and pledges. However, while by FY22 only 65 percent was spent, the FY24 work program is not fully funded highlighting the severe budget constraints at ATI. The member/financing partner split in ATI was US$2.4m/ US$25.5m (but this excludes a significant contribution from the host country Mauritius). Member contributions are low in ATI and the large membership provides opportunities for an increase in the share of financing from member countries even through relatively limited contributions. Future financing requirements are however also significant if ATI is supposed to meet the huge demand for its courses, contribute to better integration of training and TA and strengthen follow-up with participants. Part of this will require capital investments to upgrade existing facilities for which currently no financing is available.

Overall, fundraising efforts and savings from the impact of Covid-19 has allowed the RCDCs to respond to the immediate demands for TA, but ATI faces a significant level of applications for its courses, exceeding its current capacity to admit participants. In CAPTAC-DR, some financing partners frontloaded contributions during the pandemic although there is still a gap between pledges and received amounts. AFW2 managed to attract one new donor but also experienced reductions in the contribution from others. Dependency on fundraising from financing partner contributions remains very high. Improved coordination with development partners and increased visibility in the form of proactive engagement could significantly increase the chances of securing more funding agreements.

Member country contribution is seen as an important part of the RCDC funding models both to finance activities but also to promote ownership through having 'skin in the game' and ensure sustainability. AFW2 has made some progress in getting financial contributions from some but not all its member countries. The outreach visits and participation in AFRITAC SC meetings are important in this regard. CAPTAC-DR managed to almost triple member pledges from phase II to III. Ten ATI members have pledged support and ATI has ongoing negotiations with another seven members about contributions. The contrasting experiences from the different regions indicate that a member contribution model should be aligned with key economic indicators of member countries. Fragile countries should be asked to make token contributions (with potential for exceptions for resource rich fragile countries while emerging economies should be required to pay significantly higher contributions. The global and regional economic outlook in Africa, which is facing the highest debt levels since the early 2000s, however complicates efforts to significantly increase member contributions.

41 This section refers to the financial position at the end of the desk and field work phases of this evaluation. As this situation evolves continuously, the numbers might have changed from FY23 onwards.

A recent change in IMF policy suggests an increase in the allocation of its own resources (IMF01) towards the Centers. The exact implication of the policy was not clear at the end of the evaluation period, but it will most likely remain of marginal importance and only take effect from FY24. As long as RCDCs are not considered a permanent part of the IMF organizational setup, RCDCs will have to face the uncertainty associated with the volatility of overseas development aid and donor dependency. While probably manageable given the IMF's unique position to provide CD in technically complex areas, this could have implications for the long-term impact of IMF CD. The IMF and its shareholders should consider the risks of this model in the long-term and make strategic decisions on the current organizational setup of CD delivery as well as level of IMF01 financing in relation to the overall financial position of the Fund.

3.4 Allocation of resources

Allocation of resources in AFW2 is remarkably even, apart from Cabo Verde that has only received between 8% and 11% of AFW2 total resources, against an average of 17%. According to AFW2, this situation primarily reflects the country's low CD demand and absorption capacity but the lack of Portuguese-speaking experts is also an acknowledged issue. Cabo Verde also distinguishes from other members in being further removed from the Center's base in Ghana, and it is one of the three countries in the region that also receives support through the RMTF.

Distribution of resources among the CAPTAC-DR members is remarkably uneven. Costa Rica has taken the lions share with more than 20 percent of resources being allocated. Honduras is another major recipient of around 18 percent of resources. Nicaragua has received less than 10 percent of resources and the Dominican Republic only slightly more. Although no resource allocation formula was in place to give greater weight to CD performance and country specific needs, the greater allocation of resources to Costa Rica clearly reflects greater demand, reform commitment and absorptive capacity. Costa Rica also has a program with the IMF and has performed well in terms of effectiveness and potential impact. Similar patterns could perhaps be expected in AFW2 as countries vary in both commitment and absorptive capacity.

Analysis based on detailed ATI's course participant database (a proxy for resource allocation) show significant variations in number of participants per country and hence the allocation of resources.[42] The number of admitted participants over the three fiscal years under evaluation varied from 2 (Eritrea) to 157 (Nigeria). The average number of participants per country was 61 and the mean was 58. There is no clear pattern to the variations. Some of the countries in the membership with bigger economies and larger populations tend to have a higher number of participants (e.g., Ghana, 127, South Africa, 102, and Angola, 100) but other big economies had much fewer participants (e.g., Ethiopia, 45, Kenya, 42 and Tanzania, 22). The main outlier is Mauritius that benefits disproportionately from ATI course participation. ATI lecturers do make efforts to ensure a fair distribution of course participants across member countries as this is one of the selection criteria. ATI has done well in strengthening participation across the three main language groups with the number of Portuguese speaking participants rising steeply from nine in FY20 to more than 100 in both FY21 and FY22. The distribution across language groups was Lusophone (8 percent), Francophone (41 percent) and Anglophone (51 percent).

The allocation of resources across the three RCDCs is broadly reasonable. An entirely even distribution of resources should not be the goal, especially not for ATI. For many reasons, it makes sense that more officials from Nigeria than Eritrea receive training. It also makes sense to invest more in countries that perform strongly. It is important that language issues do not constrain access, but ATI has made great strides to ensure that is not the case. In AFW2, however, the issues underlying the low allocation of resources to Cabo Verde should be further scrutinized. In ATI, countries that fall below or above a certain threshold (for example the lowest and highest quintile) could be reviewed in more detail which would include countries like Tanzania (low) and Liberia (high) as well as Mauritius (high).

[42] Resource allocation data for ATI is available from FY22 but CDMAP is unable to disaggregate spending multi-country CD activities by country and therefore it is not shed light on resource allocation by country for ATI delivered training.

3.5 Modalities

The vast majority of CD delivered in the evaluation period was duty station-based TA, remote structured content courses and some remote workshops given the virtual delivery of CD. AFW2 demonstrated significant commitment to identifying and using the right mix of modalities across in-person, remote, national and regional/peer learning activities. However, Covid-19 greatly influenced these plans in the first three years of phase II. As noted in the relevance section, AFW2 delivered limited structured content training in the area of revenue administration and in general the vast majority of personnel resources have been allocated to TA missions (of which some may have been delivered as embedded or specialized training). Training was mainly through virtual/hybrid regional workshops which as demonstrated in Volume II of this report have been on the rise. Due to the pandemic, attachment/mentoring was used to a very limited extent and was not possible at all in FY21. However, it was noted in interviews that not all TA missions are followed up with training or remote continuous support. The classifications used by AFW2 do not correspond with the six main CD modalities recently adopted by the IMF. The training in AFW2 was complemented by ATI training as member countries have some of the highest rates of participation (Nigeria #1, Ghana #3, Liberia #10, Sierra Leone #14 although the Gambia and especially Cabo Verde were much further down the list of 45 countries). Yet, as discussed throughout this report, the capacity gained at the individual level from ATI training was not taken into consideration in the delivery of TA and synergies need to be strengthened.

The main modality used by ATI is multi-country, structured content courses but it has also aspired to provide non-course CD activities to diversify its modalities of CD delivery. The aim was to include customized and/or tailored in-country training, and peer-to-peer learning, conferences and high-level seminars but Covid-19 largely disrupted these plans. During the evaluation period, webinars have been used as an effective way of disseminating learning associated with the Covid-19 pandemic and new emerging themes such as gender, climate change and governance. In the area of peer-to-peer learning, the Peer-to-Peer Research Seminar Series (P2PRS) stands out as being innovative and potentially adding value to the ATI product. While these represent exciting additions to ATIs product, they should not take away the focus from the crucial course activities. ATI’s expertise in delivering structured content course to the continent already adds tremendous value and could contribute even more if better integrated with the TA delivered by the AFRITACs.

CAPTAC-DR differs markedly in the significant amount of structured content training offered in the highly prioritized area of revenue administration but also across other workstreams. Regional training and webinars were the predominant modalities (also not reported against the six main CD modalities) that were used to build technical and institutional capacities of member countries in a cost-efficient manner. A total of 49 regional and 50 bilateral training events were delivered over the past three years, with participation of 5,123 officials from all seven member countries. Reportedly, the number of officials trained in all workstreams supported increased, from 675 in FY20 to 2,705 in FY22, thus denoting the rapid response to demands from member countries. The courses were gender balanced and received an overall average score of satisfaction of 4.8 out of 5. The numbers were greatly increased due to the virtual delivery, but this has the disadvantages of limited possibility for relationship building and networking and is not a perfect substitute for face-to-face interactions. CAPTAC-DR hence operates more like a fully-fledged RCDC and not as an RTAC or RTC. This makes sense as there is no RTC equivalent to ATI for the Western Hemisphere.[43]

3.6 Human resources and expert diversity

All RCDCs were committed to diversifying the expert base, but results were mixed. Except for ATI that had a fully gender balanced use of lecturers, there is significant room for improvement in the other RCDCs. In AFW2, two LTXs (and the Director) are female and in terms of STXs, 19

[43] CAPTAC-DR could consider adopting some of ATI’s approaches to tracking the effectiveness of training. Due to the lack of this, and to an extent the low response rate to the survey, it was not possible within the resources available to conduct an in-depth evaluation of the effectiveness and efficiency of CAPTAC-DR training in revenue administration.

percent were female. In CAPTAC-DR from FY22, there were three female and four male LTXs (plus the Director). 20 percent of all STXs contracted over the period were female.

Diversity and the use of experts from the regions is also mixed. ATI management and lecturers in the courses sampled for this evaluation make up a highly diverse group. In AFW2 more than half of LTXs and more than a third of STXs working on tax and customs are from the region (although Africa as a region is also highly diverse). In CAPTAC-DR a very low number of experts were from the region, but most experts were recruited from the broader Latin American region.

Linguistically, there is some room for improvement. This is however an issue mainly concerning AFW2 and ATI as the CAPTAC-DR region is homogenous in its language profile and all recruited experts are Spanish speakers. In Africa, however, the Lusophone countries remain underserved in their official language as exemplified in the case of Cabo Verde. At ATI, one STX was identified as Portuguese speaking in addition to the Deputy Director who is a native Portuguese speaker. Although ATI has invested heavily in the use of interpretation this has limitations in relation to engaging participants.

There is appetite especially in AFW2 and ATI to further strengthen diversity of experts and lecturers but there are certain constraints. There is full recognition of the added value local experts bring in terms of understanding and navigating the cultural and political context which can help create an enabling environment for needed reforms as long as these qualities are combined with expertise and experience. Cultural competencies can also be acquired to a certain extent by experts who are not necessarily born and raised in the region but have extensive experience from working and living in different contexts. This should be a requirement in cases where beneficiaries are specifically looking to the IMF for non-regional, international TA.

A particular constraint is that all STXs have to be recruited through the IMF roster and often local experts lack the exposure and experience required to be successfully vetted by the IMF. As LTX are responsible for backstopping STX, it can be risky to recommend new people to be admitted to the roster in case they end up not having the required expertise or need substantial support in delivering expected outputs. Apart from the STX route there are very limited options to bring in local expertise such as local practitioners even as guest lecturers for ATI courses. To further grow the number of local experts involved in delivering CD, it needs to be combined with a mentoring program that would incur additional costs and might require changes to IMF HR policy.

3.7 New priority areas

The RCDCs have made considerable efforts in delivering content that corresponds with new IMF and global priority and growth areas, especially on gender and governance while also being conscious of the need for demand and ownership of the CD offer. ATI and AFW2 developed and delivered courses on gender equality and macroeconomics as well as the sampled budgeting course. This course was also delivered as an activity in Nigeria, the Gambia, Sierra Leone and Liberia. Gender-based budgeting was also done in Costa Rica. In CAPTAC-DR a regional WCO-FAD study on gender in customs administration was carried out.

Governance and anti-corruption featured prominently in ATI plans and activities. A very high-level conference on governance and corruption issues had to be cancelled on very short notice due to the onset of the Covid-19 pandemic. Instead, a course on building institutions to fight corruption in Africa was delivered both in FY21 and 22, including senior officials from key institutions. In AFW2 and CAPTAC-DR, governance issues and anti-corruption are addressed through the TA workstreams such as the support to customs integrity frameworks. In CAPTAC-DR, this area is addressed mainly through training but there are also TA activities that deal with the issue (e.g., on post clearance audit procedures).

While certainly an area of increasing attention, climate change and financial inclusion has received much less attention, especially during the pandemic, but ATI delivered a course on macroeconomics of climate change in FY22 which is supposed to be delivered regularly in the future.

3.8 Conflict and fragility

Since March 2022, the official policy on the identification of fragile countries has been to use the World Bank's annual list of fragile and conflict-affected situations. This complicated the identification of fragile states in the evaluation period and the primary methodology used was to use the World Bank list for the entire evaluation period as outlined in the approved inception note. In line with this approach, the fragile countries covered by this evaluation include: none of the countries covered by CAPTAC-DR, Nigeria (classified as fragile all three FYs), Liberia and the Gambia (both classified as fragile in FY20 and FY21, then graduated off the list) and 19 countries among the members of ATI (of which the three former are included). However, Sierra Leone was also considered fragile according to the IMF's definition through March 2022.

The findings in this section therefore only concern AFW2 and ATI and currently Nigeria is the only AFW2 country defined as fragile. With the broader definition from before March 2022, fragile states have received more than 50% of the center's resources which indicates that they have been prioritized from a resource allocation perspective.

ATI identified 18 of its 45 members as fragile and admitted 1098 participants from these countries in the evaluation period, corresponding to 40 percent. This provides clear evidence of ATI's prioritization of ensuring that officials from fragile settings receive training on equal terms. However, ATI's dataset indicates that participants from fragile countries are at a disadvantage in terms of learning gains from virtual courses. Data on learning gains were successfully collected from a lower number of participants (almost 50 percent registered as not available in the database), negative or zero learning gain (15 percent) and 39 percent who registered a positive learning gain. These metrics are a cause of concern, but ATI is already planning to work on better tailoring course materials to the specific context and absorptive capacities of fragile states.

AFW2 experts (LTX and STX) have a long experience in providing CD in fragile settings but the project-level evaluation on tax and custom projects illustrate little difference in the way CD is delivered in Ghana, Sierra Leone and the Gambia across political economy issues, prioritization (to reflect low capacity but still follow a demand-led approach) and sequencing (to set realistic objectives and milestones, including in the RBM) ignoring some important structural barriers to reforms in fragile settings. This calls for AFW2's model of delivery to be further adjusted in fragile countries.

3.9 Covid-19 response and post-Covid changes

The IMF introduced an institution wide suspension of mission travel from March 13, 2020, with all mission activity being carried out in a virtual environment. This policy was still in place at the end of the evaluation period in April 2022. All three RCDCs report that the majority of activities were cancelled or postponed until the end of FY20 in April 2020 (about six weeks to the start of the first FY of the new phases). In FY21 all mission activity shifted rapidly to virtual delivery. By FY22 some activities could be delivered in person (mostly in the host countries) and the FY22 AFW2 SC meeting in June 2022 was held in a hybrid format with in-person participation as travel restrictions in the region had been lifted.

The rapid swift and commitment to continuation of CD delivery under unprecedented conditions by all three centers was remarkable and beneficiaries unanimously commended the centers for their responsiveness and flexibility in adapting to the conditions imposed by the pandemic. In CAPTAC-DR the Center started implementing a new training modality shortly after the outbreak of the pandemic. This consisted of short virtual seminars (typically lasting 2 hours), mainly on topical issues of interest to the authorities. Many of these events were COVID-19 related and served to identify good practices to respond to the different challenges posed by the pandemic in different policy areas. In AFW2, the workplan was adjusted to respond to changing priorities and new CD needs emerging from the pandemic. Four areas of critical importance were identified: (i) protecting domestic revenue; (ii) ensuring health and safety of staff, taxpayers and other stakeholders; (iii) business continuity planning; and (iv) immediate measures that could be taken to support taxpayers. ATI used the situation as a catalyst for further digitalization and managed to reconvert all CD delivery into a virtual platform. As a result, by FY21, ATI offered a total of 62 remote events boosting the number of participants in that year.

The legacy and main takeaways from the Covid-19 era identified by this evaluation are:

- The switch to remote working has been more problematic in some countries than others generally following fragility and level of income indicators.
- Cultural preferences and many other factors such as local conditions and logistics (not just connectivity) matter and have made it difficult for many to benefit fully from CD. Evidence indicates more negative impacts of this on ATI delivered training than TA.
- IT connectivity and familiarity with remote working has improved facilitating a hybrid and blended approach to CD. In relation to this it is important to note:
  - A very strong preference for in-person versus virtual CD delivery in AFW2 and ATI
  - In CAPTAC-DR there was more immediate support for a blended approach to CD in the post-pandemic stage
- The switch to remote working has had some beneficial impacts: follow-up tends to be more regular (for TA); more cost-efficient and more inclusive events with the ability to reach much larger audiences at no additional cost.
- Virtual TA can supplement but not replace in-person engagement and works best in situations where a working relationship and trust has already been established.

Towards the end of the evaluation period, ICD was already devoting significant efforts to learn from the pandemic and develop good practices for hybrid training. Moreover, an ICD working group on blended learning was set up aiming at defining a framework for the development of blended learning, including for the RCDCs.

3.10 Conclusion - Entity level

Management and Steering Committees managed to steer the Centers through an unprecedented period of time, with continuation and producing commendable results. The pandemic introduced new ways of collaborating that with time could contribute to more regular and detailed engagement, including in reviewing CD programming through the RBM approach.

The complex governance structure of RCDCs do not necessarily inhibit strong coordination within the IMF. The internal governance reflects a very strong commitment of the IMF to quality assure the CD delivered. The evaluation does not therefore find the governance structure to be a significant concern warranting restructuring. Backstopping works well overall but coordination of efforts at country level could be further strengthened. While the Centers share information regularly some development partners see it as not having the sufficient depth to properly inform discussions about progress and challenges in the area of revenue administration and domestic revenue mobilization.

Overall, the Centers are sufficiently funded for the current phases with significant cost savings from the impact of Covid-19 but remains heavily dependent on fundraising. There is some potential for increasing membership contributions and for IMF 01 funding to be increased. The allocation of resources across the three RCDCs is broadly reasonable. An entirely even distribution of resources should not be the goal, especially not for ATI.

The mix of CD modalities used varied naturally across the three Centers and were not yet reported against the six main CD modalities recently adopted by the IMF, which will strengthen the ability to monitor and evaluate impact according to modalities. The evaluation found strong evidence of opportunities to strengthen synergies between training and TA in West Africa.

Overall, the Centers are diverse in terms of their staff composition but with room for improvement in terms of gender and linguistic profile. The requirement to recruit STX exclusively through the centrally managed roster currently poses some constraints in terms of further diversification and use of local experts. Considerable efforts have been made to deliver content that corresponds with new IMF and global priority and growth areas, especially on gender and governance, while also being conscious of the need for demand and ownership of the CD offer. While there has been a transition in the definition of fragile countries, AFW2 and ATI made great strides to tailor support to fragile countries, although further adjustments need to be considered in relation to hybrid modalities and absorption capacity.

While emphasized from the outset, the evaluation also appropriately closes by underlining the extraordinary context under which CD was delivered in the period under evaluation. It was an achievement in its own right to keep CD on track and the results achieved were more than reasonable. Lessons have been learned as well with numerous tweaks that can be made to improve virtual and hybrid delivery, and blended learning. The key lesson from this evaluation in this regard is that fragile countries are disproportionally affected by virtual delivery and a host of other reasons, including cultural, suggest that in Africa, the majority of engagement should continue to be in person. However, to reap the full benefits of improved cost-efficiency of virtual delivery at the same time as strengthening effectiveness and impact, this learning and reflection process needs to continue.

4 LESSONS AND RECOMMENDATIONS

The evaluation has produced 10 lessons with associated recommendations. This number of recommendations was condensed from a much longer preliminary list based on feedback. The 10 lessons are now associated with 11 key recommendations.

Lesson 1: As a diagnostic tool to inform CD to revenue administration, TADAT has great value which should be further exploited to inform TA. It is important that the Centers and authorities are able to draw on up to date TADAT diagnostics that could be prepared either as full repeat TADATs or as self-assessments that are supported by Center experts for example by contracting experienced TADAT assessors as STXs to support such processes. As TADAT does not cover customs the availability of good diagnostics to inform revenue administration can be skewed in spite of other diagnostics being used to fill this gap.

➢ **Recommendation**:

- RTACs, FAD and the ADs to support authorities in keeping TADATs and follow up plans up to date either through repeat TADATs or by supporting and quality assuring TADAT self-assessments.
- FAD to continue working with WCO on rolling out ISOCA as a diagnostic tool for customs at country level.

Lesson 2: Country CD strategy notes (CSNs) do not fulfil their intended purpose of guiding CD delivery strategically at country level as they are often outdated and sometimes fail to articulate how structured content training should complement TA at country level. Addressing this would be an important first step towards stronger synergies between training and TA which is currently lacking in AFW2 and ATI case countries.

➢ **Recommendation:**

- Area departments should ensure that CSNs are regularly updated and work with RCDCs to ensure that CSNs set out clear strategies for synergies between training and TA at country level. This should address how structured content courses delivered through ATI can be used to close critical capacity gaps in targeted agencies that also receive TA. This approach should be piloted for revenue administration CD in all countries that currently have outdated CSNs.

Lesson 3: Synergies between training and TA in revenue administration could be substantially strengthened by ensuring that TA recipients are selected for and follow relevant ATI courses. Close and systematic coordination between Res Reps, ATI course lecturers and LTXs is critical to make this happen. Emerging good practices from Sierra Leone and the Gambia should be replicated systematically in other countries to ensure that informed discussions take place prior to the planning of courses and TA projects.

➢ **Recommendation:**

- ATI and ICD to make it a requirement for lecturers to consult Res Reps and LTXs in relevant workstreams on selection of participants with the result of the consultation being referenced in end-of-course reports. AFD and RTACs to make it mandatory for resident representatives and LTXs to respond to consultations.
- ATI to grant Res Reps and LTXs access to the ATI participant database on a confidential basis.

Lesson 4: While all Centers invest significant effort and resources in sharing information about activities and results of CD in revenue administration, development partners (both financing partners, other CD providers and partners providing budget support) find that the depth and technical content of the information shared could be increased. In terms of TA, LTXs have the knowledge and expertise to brief development partners on key issues and ensure alignment with other activities in the field of domestic revenue mobilization. This would be most efficiently done by using (or supporting the setup of) thematically focused coordination forums on domestic revenue mobilization. In terms of structured content training by ATI the large membership makes it impractical for ATI lecturers to engage in consultations and this should the responsibility of Res Reps, which will be facilitated by the systematic consultations and access to information in the previous recommendation. Res reps should also support the mapping and coordination of capacity needs assessments at revenue administration agency level to help guide training and TA.

- **Recommendation:**

- RTAC LTXs in revenue administration should offer regular briefings of interested development partners on the progress of CD to revenue administration.

Lesson 5: Room for improvement in the use of the RBM approach to CD remains and need to be addressed urgently as CDMAP and the possibility of undertaking cost-efficiency and cost-effectiveness analysis to inform prioritization, planning and delivery of CD increases. Milestone and outcome indicator ratings (both for TA and training) need to be reported consistently to CDMAP and across other means of communication, monitored and ideally outcome indicator ratings should be subject to backstopping by ICD to ensure consistency. There is also a need to invest in strengthening the methodology to rate achievement of milestones and indicators ensuring that these are attributable to the activities supported by the project throughout the results chain (linking inputs, outputs, outcomes and objectives) as well as considering their relative importance to the broader objective. This will help demonstrate impact and value for money for external and IMF own funding of CD.

- **Recommendation:**

- ICD to develop a step-by-step methodology for ensuring clear and credible results chains to demonstrate attribution of achievements to CD delivery. This could be done by linking inputs (CD days, missions), outputs (milestones), outcomes (outcome indicators) and objectives in a convincing and consistent manner in the RBM, annual reports and TA reports. When possible, complement internal analysis with repeat TADAT assessments regarding results claimed at outcome indicator level.

Lesson 6: ATI and ICD collect substantial evidence on the performance of courses and individual participants through pre- and post-course test results and more integrated RCDCs like CAPTAC-DR providing substantial amounts of structured content training would greatly enhance the ability to monitor and evaluate training by adopting a similar approach. The robust approach was more consistently applied to ICD courses than, for example, FAD courses and it was challenged by the shift to virtual delivery. Moreover, almost no information is collected in terms of follow-up on the use of the knowledge and skills by participants in the workplace which complicates the RBM rating and reduces sustainability of the training. The impact of courses is also likely to be strengthened if participants had the opportunity to consult with lecturers on the application of techniques and tools.

- **Recommendation:**

- ATI and ICD should develop an approach to provide follow-up with course participants, partly to ensure collection of information on utilization of knowledge and skills, and partly to provide participants practical after training support in the application of knowledge and skills. To reduce the cost of follow-up, the approach could be semi-automated with a need for some direct engagement from course lecturers to be expected and factored into work planning.

Lesson 7: The expansion of CDMAP offers great opportunities for using cost-efficiency and cost-effectiveness analysis systematically to TA and training in revenue administration. The novel and experimental analysis carried out as part of this evaluation provided some initial findings indicating outliers in the cost of a CD day in some countries and demonstrated future potential for expanded and more robust use of such tools. However, the credibility of such analysis is inherently linked with assuring that classification of resources is done accurately (and in line with the new classification of CD modalities) as well as the required strengthening of the RBM approach (generally and particularly concerning the relative importance of indicators to the achievement of the objective) and should always be used with caution as a complementary tool to qualitative analysis.

➢ **Recommendation:**

- AFW2 and ICD should look into the costs of the tax project in two out of the three case study countries (in one of the countries, all costs need to be reviewed while in the other backstopping costs needs to be reviewed) to check for reporting and classification issues or other explanatory factors for the high cost per CD day identified. Based on this consider developing an application in CDMAP that automatically flags outliers for check and correction.

Lesson 8: Realizing the potential for impact of the identified revenue administration projects and courses can be strengthened if the RCDCs take action in relation to aspects such as considering political economy and governance issues in collaboration with LEG for example in relation to the support for customs integrity frameworks, engage in efforts to broaden ownership of reforms beyond the core revenue administration entity, ensure close alignment with programming and surveillance, carefully consider absorption capacity and map critical capacity gaps, carefully consider over-reliance on technology, costly data analysis software and IT skills , foster soft skills such as change management, communication skills and cultural competencies. The resulting impact should be carefully and credibly documented and communicated to development partners at the end of the current phases.

➢ **Recommendation:**

- RCDCs to pay close attention to the areas and countries identified as having stronger potential for impact, scrutinize assigned RBM ratings at the end of the current phases closely and identify metrics to inform value-for-money assessment of impact. In addition, they should share the resulting impact stories with current and potential development partners.

Lesson 9: The transition to virtual CD which gained huge impetus during the Covid-19 pandemic imposed significant constraints to effective CD delivery especially in fragile country contexts, but it also resulted in significant cost-efficiency gains. Hybrid and blended approaches are the new normal for CD delivery and it is important that this takes lessons from the pandemic and this evaluation into account. If used adequately these approaches hold potential for increased effectiveness, impact and value for money. Key lessons from this evaluation are that in Africa, and particularly concerning the fragile countries on the continent, virtual delivery should be used to complement in-person delivery and Res Rep offices should be resourced and equipped to function as country-level resource centers for reception of virtual training. Finally, virtual follow up on missions should be planned and structured to reduce the number of missions and efforts made to ensure that staff time is recorded adequately across different work modalities and assigned to the countries supported virtually.

➢ **Recommendation:**

- AFD should ensure that Res Rep offices in fragile country contexts can offer access to a training space for courses delivered virtually, which would overcome a range of identified issues associated with remote learning.

Lesson 10: RCDCs are generally well placed to deliver CD to revenue administration and are performing reasonably well across a range of areas that were assessed by this evaluation. Strong areas include facilitation of the oversight role of steering committees also during the pandemic, fundraising from development partners and members, with few exemptions a fair and efficient allocation of resources across the membership, varied mix of modalities considering the constraints imposed by the pandemic, engagement in the introduction of new thematic areas and, of course, the remarkable accommodation to the impact of the Covid-19 pandemic. Areas for further improvement concerns the need to involve the steering committee in more in-depth discussions about programming, budgets, results, cost-effectiveness, and impact, diversification of the expert base, including increased use of local experts which might require adjustments to the requirements of the use of the roster, increase the offer of structured content training on revenue administration (mainly the responsibility of FAD), secure funding for capital investments in necessary improvements to deliver state of the art training and TA, develop a realistic membership payment structure for participation in training, galvanize the ability to provide TA to Portuguese speaking countries, scrutinize the reasons for comparatively low participation of some Eastern African countries in ATI courses. As the evaluation was primarily focused on the project and course level and evidence concentrated on revenue administration issues, no specific recommendations are suggested for these lessons although addressing the areas with room for improvement is strongly encouraged.

ANNEX

1. Summary table of aggregate DAC ratings for the Revenue Administration workstream

Table 8: Aggregate ratings for the Revenue Administration workstream across DAC criteria

Relevance	**3,2**
Coherence	**2,3**
Effectiveness	**2,3**
Efficiency	**2,9**
Impact & sustainability	**2,2**

The ratings in this table are based on simple averages of the ratings presented in table 32 in Volume II