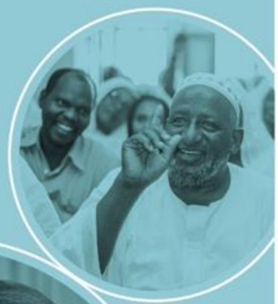
Corporate Unit Evaluation

Central Project Evaluation

KNOWING WHAT WORKS



Central Project Evaluation

Strengthening Good Financial Governance in Zambia

PN 2014.2075.1

Evaluation Report

On behalf of GIZ by Tobias Gerhard (Syspons GmbH) and
Raphael Kamoto (Ferrari & Associates)

Published version: 24 April 2019

Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

Publication details

GIZ is a federal enterprise and supports the Federal German Government in achieving its objectives in the fields of international education and international cooperation for sustainable development.

GIZ`s Evaluation Unit reports directly to the Management Board. It is separate from GIZ`s operational business. This organisational structure strengthens its independence. The Unit is mandated to generate evidence-based results and recommendations for decision-making, to provide plausible verification of results and to increase the transparency of findings.

The Evaluation Unit commissioned the evaluation to external independent evaluators. The evaluation report was written by these external evaluators. All opinions and assessments expressed in the report are those of the authors.

Evaluators:
Tobias Gerhard, Syspons GmbH
Raphael Kamoto, Ferrari & Associates

Authors of the evaluation report:
Tobias Gerhard, Syspons GmbH
Raphael Kamoto, Ferrari & Associates
Lennart Raetzell, Syspons GmbH

Consulting:
SYSPONS GmbH
Prinzenstraße 84
10969 Berlin / Germany
T +49 30 69815800
www.syspons.com

Concept, Coordination und Management
Claudia Kornahrens, Head of section
Lennart Bendfeldt-Huthmann, specialist
GIZ Corporate Unit Evaluation
Central project evaluations section

Responsible:
Dr. Ricardo Gomez, Director
GIZ Corporate Unit Evaluation

Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices:
Bonn and Eschborn

Friedrich-Ebert-Allee 36 + 40
53113 Bonn, Germany
T +49 228 4460-0
F +49 228 4460 - 1766

E evaluierung@giz.de
I www.giz.de/evaluierung
www.youtube.com/user/GIZonlineTV
www.facebook.com/gizprofile
https://twitter.com/giz_gmbh

Design/layout etc.:
DITHO Design GmbH, Cologne

Printing and distribution:
GIZ, Bonn

Printed on 100 % recycled paper, certified to FSC standards.

Bonn, April 2019

This publication can be downloaded as a pdf file from the GIZ-Website atwww.giz.de/evaluierung. For a printed report, please contact evaluierung@giz.de

Contents

List of abbreviations

AV	Officer responsible for the commission
BMZ	German Federal Ministry for Economic Cooperation and Development
CIM-expert	Expert Integrated into partner structures with BMZ-financing
DC	Development Cooperation
DFID	UK Department for International Development
EQM	Electronic Queue-Management (System)
GDP	Gross Domestic Product
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GVR	Joint Procedural Reform
IFMIS	Integrated Financial Management Information System
IMF	International Monetary Fund
KfW	Kreditanstalt für Wiederaufbau/German Development Bank
LNOB	Leave-No-One-Behind Approach
M&E	Monitoring and Evaluation
MDTF	Multi-Donor Trust Fund
MIS	Management Information System
MMD	Movement for Multi-Party Democracy
MNDP	Zambian Ministry of National Development Planning
MoF	Zambian Ministry of Finance
MPSA	Ministries, Provinces and Spending Agency
MSC	Most Significant Change (Approach)
NDP	National Development Plan
OBB	Output-based Budgeting
OECD/DAC	Organisation for Economic Co-operation and Development/Development Assistance Committee
PEFA	Public Expenditure and Financial Accountability
PF	Patriotic Front
PFM	Public Financial Management
PPP	Public-Private Partnership
SDG	Sustainable Development Goal
SMART	Specific Measurable Achievable Reasonable Time-bound (Criteria)
SME	Small and Medium Enterprises
TADAT	Tax Administration Diagnostic Assessment Tool
TC	Technical Cooperation
ToC	Theory of Change
TSA	Treasury Single Account
UPND	United Party for National Development
VAT	Value Added Tax
ZAMeT	Zambian Mobile eTaxation Platform
ZRA	Zambian Revenue Authority

Project number:	2014.2075.1
Creditor Reporting System Code (CRS code):	15111 Verwaltung der öffentlichen Finanzen
Project goal:	The transparency, credibility and effectiveness of public finances have increased
Project term:	01/2016 – 12/2018
Project volume:	EUR 8,550,000
Commissioning body:	BMZ
Lead executing agency/agencies:	GIZ
Implementing organisations (in the partner country):	Ministry of Finance (MoF), Ministry of National Development Planning (MNDP), Zambian Revenue Authority (ZRA)
Other development organisations involved:	DFID (co-financing of up to £3,550,000)

1 Summary

Background

The GIZ evaluation unit has commissioned the independent consultancy Syspons to carry out an evaluation of the GIZ project ‘Strengthening Good Financial Governance in Zambia’ (PN 2014.2075.1). The **purpose of the evaluation** is threefold: (1) to provide accountability; (2) to explain why and how different aspects of the intervention do or do not work to improve decision-making within the intervention; and (3) the study’s findings are expected to contribute to the planning process, since a mission for the planning of a next phase intervention has taken place in the follow-up to the field phase for this evaluation.

The **project runs** from January 2016 to December 2018 and builds on the predecessor project ‘Strengthening Good Financial Governance in Zambia’ (PN 2011.2114.4). The total estimated project value is EUR 8,550,000, including £3,550,000 co-financing from the UK Department for International Development (DFID) that was transferred from the previous project phase to the current one. Activities under the co-financing had started in 2015, using some of the budget. During 2016 and 2017, DFID had reassessed its support to the government-wide Management Information System, which resulted in a reduction of the original commitment. EUR 209,959 was used from the co-financing from the beginning of this project phase in January 2016 to December 2017. At the time of writing the report, it is expected that up to £855,000 will be spent by the end of the project phase. Including spending under the previous phase, the total co-funding commitment is estimated to be roughly half of the initially expected commitment.

The **project objective** is that ‘transparency, credibility and effectiveness of public finances have increased’ (intended outcome). This is to contribute to the achievement of the following intended programme-level impact: ‘Sustainable and efficient public revenue and expenditure systems have been established, and the respective accountability relations strengthened.’ To achieve these results, the project provides support in the four action areas described below. The DFID co-funding exclusively goes towards action area II.b.

- support to the Zambian Revenue Authority (ZRA) on domestic revenue generation (action area I),
- support to the Ministry of Finance (MoF) on budget preparation (action area II.a),
- support to the Ministry of National Development Planning (MNDP) on the implementation of a Management Information System (MIS) (action area II.b), and
- support to MoF on budget execution and internal audit (action area III).

The project was evaluated using a **theory-based evaluation design** that relied on the project’s theory of change (ToC) as a basis for analysis. Specifically, the evaluation team implemented a contribution analysis, which was complemented by the most significant change (MSC) approach for selected elements of the ToC. A contribution analysis consists of an analysis of the contribution of a project, and examines the extent to which observed (positive or negative) results can be related to the project.

The evaluation was carried out by a team of two consultants and relied on three main data sources: internal documentation provided by the project team; secondary data identified by the evaluation team; and first-hand interviews conducted by the evaluation team. The interviews were conducted with GIZ project staff, other GIZ staff, staff from the partner institutions (ZRA, MoF, MNDP), other development partners, the German embassy, and representatives of civil society organisations. Most interviews were conducted during a two-week field mission in Lusaka between 8 January and 22 January 2018. Preliminary findings were presented to and discussed with the project team, the appraisal mission for the follow-up project and the head of cooperation at the German embassy at the end of the field mission.

Assessment of OECD/DAC criteria

In terms of **relevance**, the project is rated successful (84 out of 100 points). It reflects and builds upon key strategic documents from the Zambian government and the BMZ, as well as the Sustainable Development Goals. In terms of the partner strategies, these include the *National Planning and Budgeting Policy*, which is the basis for the drafting of the *Public Financial Management (PFM) Reform Strategy*; the expected *National Planning and Budgeting Bill*; and the *7th National Development Plan (NDP),* each of which guides project activities. Regarding the German strategy, the project reflects the *Country Strategy Zambia* and the *BMZ Strategy on Good Financial Governance*. The project's alignment with the UN agenda is manifested in SDG 16.6 'develop effective, accountable and transparent institutions at all levels', which is closely related to the project's outcome objective.

Nevertheless, the project's successful technical approach, which focuses on the technical dimension of good financial governance rather than normative and political aspects, does have limitations. Despite the good formal match of strategies, a potential discrepancy is observed between formal and de facto strategic commitments by the Zambian government to PFM reform. More specifically, the reported slow pace of the PFM reform progress and the reported unsustainable debt situation put the government's commitment to transparent and effective PFM into perspective. In this context, the conceptualisation at the time of planning appears to have been plausible and reflects the needs and priorities of the partner. In view of the changing framework conditions, however, the focus on the technical dimension of good financial governance may not be sufficient to sustainably address PFM reform.

In terms of **effectiveness**, the project is rated successful (82 of 100 points). The stated outcome indicators are all expected to be achieved by the end of the project. One of three outcome indicators has already been achieved and the project is on track to achieving the other two. To provide a more differentiated analysis of the project's contribution to the outcome objective, the evaluation carried out a contribution analysis in each of the four action areas.

The contribution analysis revealed that the project successfully contributes to enhancing partner capacities in line with the outcome objective. However, it also showed that the indicator system does not adequately reflect the project's activities.

- In **action area I**, the project supported ZRA in developing the small and medium-sized enterprise (SME) taxation strategy, in enhancing the quality of SME tax audits and in improving taxpayer service quality. Activities on each of these aspects were reported to contribute to increasing the ZRA capacity to raise SME taxpayer compliance. It was, however, found that the proposal's indicator to measure the quantity of SME tax audits (indicator A2) was outside the project's sphere of influence and not appropriate to evaluate qualitative improvements of tax audits.
- In **action area II.a**, the project supported MoF by enhancing legal and administrative processes within the budget process. While the risks of external factors influencing the achievement of the relevant outcome indicator M2 were reported to be high, a plausible contribution of the project's legal and procedural advice to an increased capacity of MoF to prepare the budget in line with the future national planning and budgeting legislation could be carved out. One of the output indicators was found to be inappropriate, however. Indicator B1 measures the roll-out of the Integrated Financial Management Information System (IFMIS) budget module to ministries and spending agencies (MPSAs), an activity which MoF is already pursuing with support of the PFM reform programme financed by a Multi-Donor Trust Fund and managed by World Bank. It was reported that already 44 MPSAs are connected to the IFMIS, but functionality is nevertheless limited. The project therefore decided to adjust activities in this action area towards supporting the MoF in further piloting and initiating the roll-out of output-based budgeting. This is with the

view of enhancing cabinet ownership of the budget and restoring legality to supplementary expenditures, while aspects of IFMIS functionality were picked up in action area III.

- In **action area II.b**, the project supported MNDP in breaking down long-term plans to performance indicators and developing the necessary IT infrastructures to link national planning with budgeting, which is measured in output indicator B3. While the project was successful in terms of the former aspect of establishing the monitoring basis, activities are much delayed regarding the IT infrastructure. The delays were reported mainly to be a result of limited government capacity and resources but delays were also project related. On the one hand, challenges in coordination between the project, the co-funding partner DFID and the partner at management level were reported. On the other, activities in this area were not adequately reflected in the GIZ results model and indicator system. Despite these activities being planned to absorb more than half of the overall project budget (including co-funding from DFID) and despite the separation of the planning function from MoF, the results model in this area remained unchanged with only one output indicator and no module indicator.
- In **action area III**, the project supported MoF in setting up procedures and requirements to facilitate the execution of budget by MPSAs in accordance with plans and the budget act. This was achieved by technical advisory activities aiming to optimise administrative procedures for budget implementation, to enhance the internal audit function in line with international standards and to strengthen the risk management function. In terms of the first aspect, the project successfully conducted a first business process mapping to help the Accountant General's Office, in collaboration with Public Service Management Division, in developing measures to optimise administrative processes for budget execution. Regarding the second aspect, the project contributed to increased internal audit capacity by introducing quality assurance through training and technical advice. Only the last aspect of strengthening the risk management function has not achieved any results so far, because MoF's risk management unit has not yet been installed.

Finally, to some extent, the project has addressed the potential unintended positive or negative results. While crucial risks were identified in the initial results model and reflected in some of the interviews with project staff, at present, the project does not follow a systematic approach to prevent them from yielding unintended results.

In terms of **impact**, the project is rated rather successful (72 of 100 points). The overarching long-term results in terms of programme-level indicators and impact on the general population can only mean a limited success. In terms of the programme-level indicators, only one out of four is currently expected to be achieved by the end of the project term. External factors beyond the project's influence – such as macroeconomic developments (e.g. the volatile copper price) and the political situation (e.g. intransparent PFM practices, tightening of government control) – play a large role in explaining these shortcomings. The project's technical focus and three-year time frame, however, meant that it could not take measures to address the latter at a more strategic and political level. Looking at just the technical results, one can see a plausible contribution to expected long-term results in terms of increased domestic revenue generation and better alignment of budget planning and execution processes.

With regard to the third assessment criterion, the evaluation team assesses the achievement and monitoring of unintended results as partly met. No unintended results, whether positive or negative, were actually observed. However, a few risks such as the political instrumentalisation of ZRA were identified. The project had discussed but not yet addressed these by developing a mitigation strategy or adapting relevant activities. The concept note for the follow-up phase clarifies, however, that related risks are anticipated and plans to conduct results-oriented monitoring and an exit strategy, if risks get too high.

In terms of **efficiency**, the project is rated rather successful (70 of 100 points). Regarding production efficiency, the following issues arose despite an overall sensible allocation of resources. First, given the crucial role that action areas II.a and II.b play in terms of the module objective and the financial resources, it is surprising that there was only one advisor for each area. Action areas I and III, in contrast, work with three and two advisors respectively. In action area II.b, the limited human resources were reported to have contributed to the limitations in meeting the project's objectives. Second, inefficiencies were detected regarding the co-financing in action area II.b. On the one hand, the project's high overarching costs appear to be partly a consequence of related additional administrative processes. On the other hand, there were considerable deviations between the initially planned and the actually spent resources in action area II.b, which ultimately contributed to undermining the project's ability to maximise results in that area.

Regarding allocation efficiency, the evaluation team observed examples of synergies that could, however, be further exploited in part through better coordination within the donor landscape. In particular, while project results could be leveraged with activities of the PFM reform programme of the World Bank-managed Multi-Donor Trust Fund that this project related to through the common DC-measure with KfW, the direct coordination beyond the exchange in coordination meetings through the Cooperating Partners Group headed by German Cooperation may have been able to achieve even more.

In terms of **sustainability**, the project is rated successful (82 of 100 points). The prerequisites for ensuring long-term success in the form of tools, concepts and approaches being anchored in the partner structure are largely fulfilled. The project design is focused on finding technical solutions, and the evaluation team has assessed it to be successful at doing that. Given the important political dimension of the challenges in the PFM framework, however, a purely technical approach may not be enough. Here, the team believes that the political challenges also need to be addressed with regard to the political and normative dimensions of good financial governance. Moreover, the close embeddedness of the project's experienced technical advisors in specific units within the partner structure, in spite of its merits, creates a risk of dependency and of leaving a gap, once the project is over.

Finally, the evaluation team observed strong results regarding economic sustainability and an indirect contribution to social and environmental aspects of sustainability. The project's focus on economic sustainability is sensible given its overall objective. Social and environmental aspects (e.g. climate-proof and gender-sensitive budget proposals) appear to be a low priority on the partner's agenda, which may explain why there was no follow-up of some related activities from the previous phase.

The overall score for all criteria adds up to 390 / 5 = 78, which amounts to the rating 'rather successful'.

Criterion	**Score**	**Rating**
Relevance	*84 points*	*Successful*
Effectiveness	*82 points*	*Successful*
Impact	*72 points*	*Rather successful*
Efficiency	*70 points*	*Rather successful*
Sustainability	*82 points*	*Successful*
Overall score and rating for all criteria	*390 / 5 = 78 points*	*Rather successful*

100-point-scale	6-level-scale (rating)
92–100	Level 1 = very successful
81–91	Level 2 = successful
67–80	Level 3 = rather successful
50–66	Level 4 = rather unsatisfactory
30–49	Level 5 = unsatisfactory
0–29	Level 6 = very unsatisfactory

Key recommendations

Recommended changes to the indicator system:

Recommendation one: **Reformulate outcome indicator M2**

Recommendation two: **Revise output indicator B1**

Recommendation three: **Revise output indicator A2**

Recommendation four: **Develop separate output and outcome indicators for the future in action area II.b once the strategy is aligned with DFID**

Strategic short-term recommendations:

Recommendation one: **Consider ways to improve coordination with the government through an effective steering structure**

Recommendation two: **Consider ways to achieve even more synergies with the other key international partners**

Recommendation three: **Consider ways to manage risks and avoid unintended results more systematically**

Strategic medium-term recommendations:

Recommendation one: **The definition of the target group should be reassessed**

Recommendation two: **Consider cluster-based advisory approach to break the silo structure within the government**

Recommendation three: **Consider ways to strengthen transparency and accountability within the PFM system, for example by involving external control actors or through partnerships**

2 Evaluation objective and questions

2.1 Objectives of the evaluation

This evaluation of the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) intervention ‘Good Financial Governance in Zambia’ is one of the first central project evaluations to be implemented after GIZ’s fundamental reform of its evaluation system in 2016. The goals of this interim evaluation, which is taking place after two of the three-year project duration, are twofold: (1) It aims to explain why and how different aspects of the intervention do or do not work to improve decision-making within the intervention. (2) The study’s findings are expected to contribute to the planning process, since a mission for the planning of a next phase intervention has taken place in the follow-up to the field phase for this evaluation.

The evaluation’s central stakeholders are the GIZ evaluation unit; the intervention staff and the partner organisations; the Ministry of Finance (MoF); the Ministry of National Development Planning (MNDP); and the Zambian Revenue Authority (ZRA). The GIZ evaluation unit pointed out that it has the specific goal to achieve accountability and enhance the basis for decision-making within the project. Moreover, it is interested in receiving feedback on the implementation of the new central project evaluation and its tools (e.g. the efficiency measurement tool) under the new framework conditions (Agenda 2030, the new joint procedural reform (GVR)) (interview 1). The project team and its head (AV) seek to find out what and how it has worked in order to incorporate this knowledge into the next phase’s planning. The AV pointed out that the national partners are most interested in learning about the modalities of the cooperation with GIZ in terms of strategic planning and steering the intervention.

2.2 Evaluation questions

The intervention was assessed on the basis of Organisation for Economic Co-operation and Development/Development Assistance Committee (OECD/DAC) standardised evaluation criteria (relevance, efficiency, effectiveness, impact and sustainability). The evaluation dimensions and analysis questions derived from this were specified by GIZ and broken down into an evaluation matrix. In addition to these evaluation criteria, the contributions to the Agenda 2030 and its principles (universality, integrative approach, Leave-No-One-Behind (LNOB), multi-stakeholder partnerships) were taken into account. The evaluation questions also relate to cross-cutting issues such as gender, the environment and human rights.

Preparatory discussions were conducted between the evaluation team and the AV as well as representatives from the three output areas of the intervention to find out, among other things, whether specific evaluation questions need to be added from the perspective of the intervention team. No additional evaluation questions were brought forward in these discussions. It was merely stressed that the intervention is very interested in input for the planning mission of the follow-up intervention (interview 5). GIZ’s sectoral unit also emphasised their interest in learning from this evaluation for the preparation of the follow-up measure. In particular, it was interested in finding out which areas of cooperation are most relevant to the partner; how much influence the political tensions and changes actually had on the project’s results at partner level; and what the project could have done better to mitigate the risks resulting from it (interview 9). The entire list of evaluation questions with respective indicators and data sources can be found in the evaluation matrix in Annex 1.

3 Object of the evaluation

3.1 Definition of the subject matter

The mid-term evaluation involves the technical cooperation (TC) measure, Strengthening Good Financial Governance in Zambia (PN 2014.2075.1). The TC measure supports selected partners of the national government, namely MoF, MNDP and ZRA, and addresses meso- and macro-level reform of the Public Financial Management (PFM) system by supporting the shaping of institutional framework conditions for more effective and transparent public financial management.

The intervention lasted for three years from January 2016 to December 2018. The total estimated project value was EUR 8,550,000 including £3,550,000 co-financing from the UK Department for International Development (DFID). This co-financing amount stems from residual funds of the predecessor programme Good Financial Governance in Zambia (PN 2011.2114.4). Activities under the co-financing had already started and consumed budget in the previous phase in 2015. During 2016 and 2017, DFID had reassessed its support to the government-wide Management Information System (MIS), which resulted in a reduction of the original commitment. From the beginning of this project phase from January 2016 to December 2017, EUR 209,959 had been used from the co-financing. At the time of writing the report, it is expected that up to £855,000 will still be spent by the end of the project phase. Including the spendings under the previous phase, the total co-funding commitment is estimated to be roughly half of the initially expected commitment.

The intervention takes place in Zambia and works with selected Zambian national government institutions, namely MoF, ZRA and MNDP, which was newly established in 2016. For certain reform initiatives, such as the introduction of output-based budgeting (OBB) or the government-wide MIS for national planning and budgeting, additional line ministries are involved in specific activities. The project is part of the BMZ-funded common DC-measure with KfW ‘Good Financial Management’ of EUR 12,000,000 running from March 2016 to December 2020. The KfW contribution is channelled through the Multi-Donor Trust Fund into the PFM reform programme, which is managed by World Bank and combines financing from Germany, the UK and Finland.

In the political context, the MoF was at the time of the evaluation under the leadership of Felix Mutati, who is a member of the Movement for Multi-Party Democracy (MMD) in contrast to President Edgar Lungu from the Patriotic Front (PF). In February 2018, he was replaced as Minister of Finance by Margaret Mwanakatwe. In ZRA, Kingsley Chanda was appointed as new Commissioner General after the start of the project in 2016. Questions were raised regarding the political independence of ZRA on account of authorities lawfully seizing the contents of the premises of the opposition newspaper *The Post*. The seizure was based on a court order, due to tax fraud in the run up to the 2016 elections (interviews 15, 33). Finally, the separation of MNDP from the MoF meant that new processes needed to be established, new responsibilities needed to be defined and positions needed to be filled. Until now, many positions in the new ministry have not yet been filled. The minister was replaced during the time of this evaluation’s field mission, which marks the third minister since the beginning of the project (interviews 23, 33).

Regarding the political-economic situation, the partner’s strategic reference is the *National Planning and Budgeting Policy*, which forms the basis for the drafting of the *PFM Reform Strategy*, the expected *National Planning and Budgeting Bill,* and the *7th National Development Plan (NDP),* each of which guides project activities. Independent analysis shows the lack of transparent and effective processes within the government (see Beardsworth 2017). Over the past two years, several corruption scandals were reported (interviews 10, 28, 32, 33) and the International Monetary Fund (IMF) had temporarily put negotiations with the Zambian government on hold due to unclarities over its fiscal policies and borrowing plans.[1] Regarding the macroeconomic situation, latest IMF figures for 2016 report public debt to be 60.5% of gross domestic product (GDP) and the fiscal deficit to be 5.8% on a cash basis (see IMF 2017). Political tensions between the ruling political party Patriotic Front (PF) and the main opposition party United Party for National Development (UPND) are reported to pose risks to the stability of the country (see IMF 2017). Moreover, the long-expected National Planning and Budget Bill, which was initially expected to be passed by parliament in early 2016, has yet to be adopted.

3.2 Results model including hypotheses

Based on the analysis and the associated problem tree, the intervention has identified three separate results models for each of the three intervention output areas. These are embedded into the outcome and impact-level results of the TC measure and development cooperation (DC)-programme respectively (see Annex 2). The results models were developed in preparation of the intervention and were partly updated (in the case of action area III) in early 2017 to reflect the adapted environment. They build upon the results matrix from the intervention's proposal (see GIZ 2016) and specify activities, results, assumptions and risks.

In addition to the results models, a more detailed theory of change (ToC) was developed for each action area that specifies relevant problems, strategies, tactics, targets, and short and long-term outcomes. On improving budget planning, output area II had developed two ToC. This was because of the different focus areas within the output area (strengthening budget preparation on the one hand and strengthening national monitoring and evaluation (M&E) systems on the other hand). The results models and the ToCs for all four action areas can be found in Annex 1. All combined, these results models contain one outcome with three outcome indicators and three output areas; each being measured against two to three output indicators. As will be shown, output area I operates largely independently of the other areas on domestic resource mobilisation, while output areas II and III are interlinked by addressing different aspects of the budget cycle within the PFM system. We now describe each of the output areas separately in terms of their objectives, indicators and underlying hypotheses.

Output area I on strengthening tax administration aims to increase the capacity of ZRA to raise tax compliance of small and medium-sized taxpayers (output A). Here, the underlying hypothesis is that an increase in the quality of SME taxpayer services, in combination with an increase in the number and the quality[2] of SME company audits, will strengthen ZRA's capacity in SME tax administration. In turn, it is expected that more SMEs will comply with the relevant tax regulations. Project activities in this area include the provision of training and technical advice on conducting SME tax audits and process advice on developing and implementing the SME taxation strategy. ZRA contributes to achieving objectives by providing staff time, among other things, for improving the quality of tax audits and working on the SME taxation strategy. If successful, results in this area contribute to reaching the following indicator (outcome indicator 1): 'The share of middle-income taxpayers complying with value added tax legislation under the purview of the Medium Taxpayer Offices (MTOs) North and South has risen to 60%'. This output area is assessed by the following two indicators:

- Two measures of the strategy for the taxation of SMEs have been implemented (A1)
- The number of audits of SMEs has increased by 20% (A2)

A number of potential risks (e.g. that qualified ZRA auditors might be hired by the private sector and possibly bypass legal provisions or that elections might affect reform dynamics) were considered in the project design. Additional risks and context factors that arose during project implementation were the allegations of political

[1] See http://www.imf.org/en/News/Articles/2017/11/30/tr113017-transcript-of-imf-press-briefing (retrieved from the internet on 21 March 2018).

[2] As was suggested in the inception report and confirmed in the field study (see Section 5.2 on Effectiveness), not only the quantity, but more importantly the quality of SME tax audits is a crucial aspect determining SME taxpayer compliance. While this aspect was not incorporated in the indicator system it was implicitly reflected in the project´s activities, therefore forming part of the underlying hypothesis.

exertion of influence surrounding ZRA's involvement in closing down an opposition newspaper prior to the 2016 election due to tax fraud (see Section 3.1) and the organisational restructuring of ZRA after the appointment of a new Commissioner General.

Output area II on improving budget planning works on enhancing budget preparation and strengthening a national MIS. For the work on MIS, the project receives co-funding from DFID. While the activities on budget preparation and MIS were initially planned complementarily in one output area, it was decided to divide them into two sub-areas after the separation of planning activities from MoF to the newly established MNDP and the request of DFID to have a separate ToC for the common intervention (interview 3). Now, *action area II.a* deals with budget preparation and seeks to move the negotiation and decision-making on resource allocation away from budget execution to the budget planning phase. *Action area II.b* works on strengthening MIS systems and seeks to match the ministerial budgets to Zambia's long-term strategic goals. Within the GIZ reporting system, however, they both remained under output B, striving to increase MoF capacity to prepare the budget in accordance with the future national planning and budgeting legislation (output B). This is to be measured on the module level through a 25% reduction in the number of budget heads deviating more than 5% between approved and actual expenditures (outcome indicator M2). Due to the potentially misleading formulation, the evaluation team suggested in its inception report to rephrase the indicator M2 as follows:

> *The share of individual budget heads that differ more than 5% between approved and actual expenditure has diminished by a quarter*.

  - Baseline value: 90% (47 of 52) of heads deviated more than 5% from their appropriation in 2015
  - Target value: 68% (= 90% x 75%)

In light of the operational separation of activities between the two sub-areas, the underlying hypotheses and indicators are now presented separately for each of the action areas.

For **action area II.a (budget preparation)**, the underlying hypothesis according to ToC is that enhancing legal, administrative and political processes of budget and supplementary budget preparation on the one hand and supporting the piloting and roll-out of OBB on the other will lead to more accurate planning of public expenditures. Project activities include the provision of legal, technical and procedural advice on the drafting of relevant reform bills and on measures to improve the analysis and preparation of budgets. The action area is closely linked to indicator M2, which seeks to reduce deviation between planned and actual expenditures. It is measured by the following output indicators according to the project proposal:

- 6 of a total of 54 MPSAs prepare their budgets in the IFMIS budget module (B1)
- 5 specific measures (e.g. improvement in the communication between MoF and MPSAs, submission of realistic budget proposals) to improve budget preparation have been implemented in 2 MPSAs (B2)

The results model stated the project's generic risks as being key (e.g. effects of the 2016 election on reform dynamics, a culture of circumventing legal provisions). Additionally, emerging risks that affect all areas, but particularly this action area, are the worsening public debt situation. Also, that after years of debate the budget and planning bill has still not passed parliament at the time of writing this report (see progress report (GIZ 2017b); interviews 12, 15).

For **action area II.b (implementing MIS)**, the underlying hypothesis according to the ToC is that by developing national and sectoral performance indicators, establishing the underlying processes and based on this developing suitable IT systems, MNDP will have an increased capacity to inform and guide the budget preparation process in MoF. Project activities include the provision of technical advice on developing the M&E strategy, establishing performance indicators and the analysis of existing IT infrastructures within the government, to advice on a suitable future MIS infrastructure. Within the GIZ reporting system, this action area is monitored with the following output indicator (B3):

- In the three sectors piloting the MIS Programme as well as in MoF and Cabinet Office, relevant data (e.g. programme results according to the National Development Plan) have been generated in a harmonised format for budget preparation purposes as well as for national reporting.

Here, the main change of contextual factors was that the planning function was separated from the MoF and integrated into the newly established MNDP, which was described as a slow and unfinished process (interviews 13, 23, 30).

Output area III on improving budget execution including internal audit aims to raise the ability of the MoF to facilitate effective implementation in accordance with plans and the enacted budget (output C). Here, the hypothesis is that optimising administrative processes of budget implementation, improving risk management and enhancing the quality of internal audits will lead to more regulated budget implementation. Key activities of the project in this area include the mapping of key business process in order to provide recommendations on improving administrative processes and the provision of training and advice on quality assurance in line with international standards. The output indicators in this area are:

- Under the guidance of the Accountant General and in consultation with other relevant actors (such as the Budget Office and selected institutions responsible for managing resources), action plans on optimising administrative processes in at least three areas of budget implementation (such as revenue, human resources (HR) expenditure and public investments) are coordinated and put into operation (C1, adapted with the 2016 progress report)
- 50% of participants trained in the risk management system prove with examples that they apply the risk management principles in their work (C2)
- Three specialised audits have been conducted by internal auditors (C3)

At the module level, the objective in this area is expected to be met once a rule-based procedure for the allocation of funds (e.g. an authorisation procedure based on expenditure in accordance with budget) is applied for at least one ministerial budget (outcome indicator M3). Due to the interconnectedness of budget preparation and budget execution, indicator M3 is also influenced by the activities in action area II.a. However, a major risk identified in project planning was the continuation of politically motivated funding reallocations. In terms of additional contextual factors, the Accountant General, as key partner in this area, was promoted to Permanent Secretary, while the position of internal auditors in the line ministries was described as not high enough to effectively fulfil their auditing function (interviews 14, 25, 26).

The project's **partner system** includes selected units within the partner institutions most relevant to the PFM system: MoF, MNDP and ZRA. For some activities, such as support to the internal audit function or the facilitation of cluster hearings with MPSAs, selected additional government agencies may be targeted and involved. The **capacity development strategy** focuses on the following areas. (1) On the individual level, different human capacity development measures (coaching, mentoring and advice) support partner staff in performing their tasks more effectively. (2) On the organisational level, the measure helps improve selected core processes through organisational and process advice, e.g. the piloting of a warrant system to replace cash rationing. (3) On the societal level, the measure seeks to engage with the wider population by working on the inclusion of tax issues in school curricula and by supporting the publication of the citizen's budget for example.

Regarding the three **dimensions of sustainability** (economic, social and environmental), economic sustainability is presently largely incorporated into the project design, since all project outputs contribute to the overall programme goal of achieving sustainable and efficient PFM systems. Social aspects of sustainability were considered when developing a guideline for gender-sensitive budgeting through training. In addition, all three dimensions are indirectly relevant at the impact level, in the sense that the programme is expected to contribute to making more efficient use of all public expenditures, including in areas of social and environmental policy.

Regarding the **quality of the results chain**, it was pointed out in the inception report that no single results model was documented for the overall project beyond the results matrix. Rather, separate results models were developed for each output area at the phase of project planning. Each results model represented clear causal chains and provided a profound analysis of the sphere of influence as well as potential risk factors. During implementation, these results models were translated into tabular ToCs for each action area based on the problem tree analysis to provide additional information on the project's approach, problems and expected outcomes.

3.3 Target group analysis

The project proposal defines the entire population of Zambia as the target group, since enhanced national-level PFM structures contribute to improved public service provision and therefore reach all segments of the population. Due to the higher dependency and disproportionate benefit of disadvantaged groups from public spending such as welfare, it is argued that the poor, in particular, will benefit from more effective public expenditures (see GIZ 2015b).

In the programme logic, however, the impact level, which aims to have 'sustainable & efficient governmental revenue & expenditure systems […] established, and the respective accountability relations strengthened' is measured through macroeconomic data and does not build a direct link to the stated target group (GIZ 2016: 9). Equally, at the outputs and outcomes level, the partner institutions (MoF, MNDP, ZRA) rather than the general population act as the project target group.

In conclusion, the project has two target groups that need to be differentiated. (1) The immediate target group is the partner institutions, which is reflected in the entire indicator system of the project including at programme level. (2) The ultimate target group, however, is the general population, which can benefit from the project through enhanced provision of public services in all policy sectors but is not reflected in the target system.

4 Evaluability and evaluation design – data sources, data quality and evaluation methods used

4.1 Data sources, data quality

The evaluation relied on three main data sources: internal documentation provided by the project team, secondary data identified by the evaluation team, and first-hand interviews conducted by the evaluation team. The internal documentation included the project proposal and annual reports, the proposal and evaluation report of the preceding project, the programme proposal and progress report as well as the concept note for this follow-up measure. In addition, it included a range of internal documents such as the operational plan, steering structure, capacity development strategy, stakeholder maps, results models and the theory of change. Moreover, the evaluation team consulted, the Delegated Cooperation Agreement between DFID and the German Ministry for Economic Cooperation and Development (BMZ) for the co-financing as well as DFID reports on project progress in action area II.b. Finally, the team investigated internal cost data for the analysis of efficiency. Annex 4 sets out a full list of these documents.

Secondary data identified by the project team included studies on the political and socio-economic situation in Zambia to better understand the context conditions and put the expected impact into perspective. A full overview of these documents can also be found in the bibliography in Annex 4. The interviews were conducted with project staff, other GIZ projects, staff of the partner institutions, other development partners, civil society organisations and the German embassy. Interview partners were selected on the basis of their ability to provide relevant information about the project and the sectors in which the project is operating. To come to this selection, the evaluation team first established a list of internal and external stakeholders with whom the project is cooperating on the basis of the project documentation (project proposal, progress reports, and stakeholder maps). Then we consulted the GIZ sector expert who was involved in planning the intervention, to add relevant interview partners from the planning stage. Finally, the evaluation team asked the project team to comment on and add to this list. This approach allowed the evaluation team to conduct interviews with interview partners covering diverse perspectives from within and outside.

Because of the high number of stakeholders identified, the evaluation team decided to conduct some of the interviews with partners as group interviews to take into account all relevant perspectives in the given time frame. Finally, some interviews conducted during the evaluation mission led to the identification of additional interview partners on a specific subject matter. The evaluation team thus added some interviews to the schedule on the spot during the evaluation mission in Lusaka. A full list of interviews conducted can be found in Annex 3.

The gathered data was documented through interview notes and the evaluation matrix attached to this report. The evaluation team held an internal synthesis meeting every evening of the field study, to reflect on the findings and triangulate data (see methods used below).

Overall, the evaluation team considers that the data on which this evaluation is based is of good quality, in the sense that there was ample documentation on processes supported by the evaluation. Regarding the interviews, the team managed to talk to almost all relevant interview partners in spite of the high number of partners with whom the project works in the different partner institutions (MoF, MNDP, ZRA). Only two of the planned interview partners were unavailable for interview. The evaluation team considers that the availability of many high-level interview partners including two Permanent Secretaries, the Accountant General and ZRA’s Commissioner General is in itself testimony to the high standing of the project within the partner structure. It was found, however, that in some cases the conduct of group interviews limited the ability to reflect and triangulate data. In particular in action areas II.a and III, the group interviews reduced the number of data sources, and the evaluation team observed a tendency among group interview participants to agree with each other. While the evaluation team attempted to follow up with selected interviews, this was not possible due to both the interview partners' and the evaluation team's limited time.

Moreover, the evaluation team used the project's monitoring data, which tracks progress against the indicators from the project proposal in the GIZ online tool 'results monitor'. This monitoring data is usually collected by the advisors from their respective partner institutions. Where possible, partner data was used, such as in the case of the data on value added tax (VAT) compliance by ZRA. As previously indicated by the project staff, however, the quality of data provided by the partner is partly insufficient. The data on VAT compliance in action area I for example was provided as a table in a Word document, which increases the risk of manipulation or mistakes. Other data, such as the deviation of actual from planned budget (outcome indicator M2), was calculated by the project itself based on the budget and expenditure data provided by MoF. Therefore, much of the data for the evaluation was drawn from the project team and then triangulated in the interviews with partners and other stakeholders. Apart from the aforementioned aspect, the evaluation team found the information from the project's monitoring, including the baseline data, to be reliable; and it is in line with information provided by partners and external stakeholders during the interviews conducted in the evaluation mission. However, the indicators formulated in the project proposal do not fully do justice to the changes to which the project aims to contribute. To address this, the evaluation team's assessment of effectiveness and impact not only takes into account progress against these indicators, but also a qualitative assessment of the project's contribution to the intended outcomes and overarching development results (see Sections 5.2 & 5.3 on Effectiveness and Impact).

4.2 Evaluation design, basis for assessing OECD/DAC criteria and methods used

Evaluation design

In line with the Terms of Reference, the evaluation team adopted a theory-based approach that relied on the project's theory of change as a basis for analysis. Specifically, it implemented a contribution analysis related to the OECD/DAC criteria effectiveness and impact. A contribution analysis consists of an analysis of the contribution of a project and analyses the extent to which observed (positive or negative) results can be related to the project (Mayne 2001). Contribution analysis differs from other forms of theory-based evaluation in so far as it not only analyses the hypotheses of the ToC, but also seeks to identify alternative explanations that may explain observed impacts. Contribution analysis does not seek to prove that one factor 'caused' the intended impact, but analyses the extent to which the project has contributed to the observed impacts. Data from various sources is collected to analyse the causal hypotheses between inputs, outputs, outcomes and impacts formulated in the ToC. Contribution analysis thus seeks to construct a credible 'performance story' to show whether the programme was relevant, possibly together with other factors, to lead to change (ibid). Context factors that play a role in achieving (or not achieving) the project's objectives are explicitly taken into account in contribution analysis.

Contribution analysis falls into the category of the **generative/mechanisms approach** to causal inference. This approach relies on identifying the 'causal mechanisms' that generate the desirable effects. In order to use this approach, the existence of one case with good quality data sources is sufficient. The approach is based on an existing theory for the project in question which allows the evaluator to understand the factors that cause the observed effect. As a result, this approach permits an in-depth understanding of the case and its context, proving a detailed explanation of both of them (Stern et al. 2012). The evaluation team chose this approach over other approaches to causal inference, such as the experiment/counterfactual approach, the regulatory approach or the multiple causation approach. An experiment/counterfactual approach was not deemed feasible because it is not possible to meet the methodological requirements for this approach in terms of finding a comparable control or comparison group. A regulatory approach was deemed largely unsuitable because for most of the ToC causal hypotheses, the number of direct beneficiaries (e.g. staff of partner institutions) is limited. It would therefore have hardly been possible to draw conclusions on cause-effect-relationships on the basis of a high number of observed cases for these causal hypotheses, over time. Lastly, a multiple causation approach to causal inference was not deemed feasible due to the complexity of the project.

In order to strengthen the generative/mechanisms approach to causal inference, the evaluation team combined the contribution analysis with the MSC approach. This qualitative evaluation method makes use of storytelling and assumes that certain less visible impacts and unintended results can best be identified through key events or changes experienced by the main stakeholders. Particularly in complex multi-stakeholder environments with limited transparency, this approach can offer valuable insights that are otherwise difficult to capture. The MSC approach was used to gain a deeper insight into the interview partners' perspectives on the evaluation dimensions of effectiveness, impact and sustainability, because in these areas MSC has most potential to reveal more hidden results and relationships.

Presentation of the basis for assessing the OECD/DAC criteria

As a systematic basis for assessing the intervention based on OECD/DAC criteria, the evaluation team adapted the evaluation matrix that was provided by GIZ to the present context. For each of the criteria, the guiding questions were specified into concrete evaluation questions, indicators and the respective data sources, and evaluation methods were identified. The full evaluation matrix can be found in Annex 1.

The **relevance** criterion analyses the extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, regional needs, global priorities and partner' and donors' policies. The underlying question is whether the intervention sets the right priorities, both in how it was planned initially, and in how it is implemented in practice. In this regard, the evaluation team analysed how far the objectives of the intervention are aligned with regional strategies of German development cooperation partners. Where applicable, it was also assessed in how far evolving needs were taken into account in programme implementation.

The **effectiveness** criterion analyses progress towards the intervention's objective. The focus of this analysis lies on an assessment of progress towards the module objective indicators (outcome level), and given the approach of a contribution analysis towards the output indicators. As described above in Section 3.2, not all indicators fulfilled the SMART quality criteria and therefore some required adaptation. A detailed assessment of the indicators and the new indicators formulated by the evaluation team can be found in Annex 2. In addition to the indicator-based analysis, the assessment of effectiveness included an analysis of the quality of key processes supported by the intervention and their contribution to results, as well as an analysis of any potential unintended results.

The **impact** criterion measures the extent to which the intervention contributes to the achievement of overarching development results. In this regard, the evaluation questions relate to the contribution to the programme-level objectives, to the implementation of a national development strategy, and to the implementation of the Sustainable Development Goals (SDG) agenda. In particular, the evaluation team assessed to what extent and how much the intervention has affected the general population as its ultimate target group. Moreover, similar to the analysis on effectiveness, potential overarching unintended results were analysed.

The **efficiency** criterion measures the extent to which objectives of an intervention have been cost-effectively achieved. An intervention is thus efficient when the maximum amount of results are achieved with the available financial resources. Given the mid-term nature of this evaluation after two of the three project years, the analysis focused on production efficiency: the transformation of inputs to outputs. To implement this analysis, the evaluation team tested an Excel tool introduced by GIZ's evaluation unit to carry out a 'follow-the-money' analysis. In this analysis, resources allocated to outputs are analysed against progress towards achieving these outputs. In addition, other relevant aspects, raised by the intervention team in terms of efficiency, were also documented.

The **sustainability** criterion examines the extent to which positive results of the intervention can be expected to last once the intervention has ended. In this regard, the evaluation team analysed the efforts of the intervention towards sustainability, e.g. the extent to which approaches and tools are elaborated jointly with the partner to foster ownership. The evaluation also analysed the extent to which partners have already taken up or are using approaches and tools developed within the intervention.

Given the different focus of the previous programme, it was decided during the inception phase not to focus on that previous measure. Nevertheless, the evaluation team looked into the results of the intervention's previous phase to put any potential long-terms results into perspective.

Methods used

The methods used in the evaluation included analyses of internal documentation, secondary data and the interviews conducted by the evaluation team. A systematic approach was used for the document analysis. In the inception phase, the project proposal and the results models were used to understand what the project intends to achieve and how. The progress reports were used to understand progress towards the project's objective as well as context factors that affect progress. In addition, stakeholder maps were also consulted to understand the role of the different actors involved. The evaluation team further reviewed the capacity development strategy, which proved useful for better understanding the context in which the project is operating. The internal documentation was continuously revisited during the evaluation mission and, in the analysis phase, used to triangulate and complement it with information from other sources. This was especially important because the project's proposal and reporting do not fully capture the project's complexity.

The strength of the internal documentation lies in the fact that it provides information that can be directly related to the project's results model and the quality of the implementation process. Internal documentation, however, comes with a potential bias, since most of the internal documentation is prepared for the commissioning party (BMZ), and there may be an incentive to focus on successes rather than weaknesses. This bias was compensated for by also consulting external evaluations, reports and studies on the Zambian PFM system and the political-economic framework. In the context of this evaluation, the strength of interviews was that they provided detailed qualitative information on the quality of processes but also on political context factors highly relevant for achieving results. Depending on the type of stakeholders, some interview partners may also have a certain bias against addressing weaknesses in the project. The evaluation team found the project team itself open in talking about both strengths and weaknesses, although some of the partners were rather prone to focus on strengths only. It was useful to take into account the perception of other development partners and external actors in order to get a balanced perspective.

To arrive at valid and reliable information, whenever possible the evaluation team aimed for systematic data triangulation (taking into account the perspectives of different stakeholders on the same aspect) and/or method triangulation (using various methods of data collection to collect information on the same aspect). It was not always possible to do both for every aspect. Possibilities for data triangulation were limited for some evaluation aspects because only the project team and the respective partner with whom the project is cooperating in a given area knew specifics about the project. For this reason, external actors (other GIZ projects, sector experts, other cooperation agencies or donors) could only triangulate part of the evaluation aspects, such as general developments in a sector, the quality of coordination between different development partners, or an overall impression of the project's contribution to changes in a given area. Possibilities for method triangulation were limited for some evaluation aspects because not all aspects of the project are covered in internal documentation, and secondary data does not cover project specifics. However, either method or data triangulation proved possible for most aspects analysed by the evaluation team. The evaluation matrix in Annex 1 and Section 6 on evaluation findings gives the sources and methods of data collection for each finding to clearly show how the evaluation team came to its conclusions.

In addition to data and method triangulation, the evaluation team carried out researcher triangulation. The local and the international evaluators regularly reflected and analysed their findings during the evaluation mission. Additionally, the observation and analysis of critical points was reflected upon with the international back stopper at the end of each field study week. The analysis of evaluation results was systematically carried out along the lines of the evaluation matrix in Annex 1. This evaluation matrix was developed in the inception phase of the evaluation and details evaluation dimensions, analysis questions and indicators for each evaluation criterion. During the evaluation mission, the evaluation team documented results in interview minutes. At the end of the evaluation mission, the international evaluator elaborated the presentation for the debriefing of the mission, which was commented on by the local evaluator and revised accordingly before the debriefing. A triangulation of results with involved stakeholders was carried out by means of an internal debriefing session with the project team, an external debriefing with the head of cooperation at the German embassy and a debriefing session with the appraisal mission for the follow-up project. The comments made by participants in the debriefing sessions were taken into account in the preparation of the final report, which was drafted by the international evaluator. Finally, the international back stopper and the local evaluator reviewed the draft report, further consolidating the researcher triangulation.

For the assessment of the efficiency criterion, the evaluation team made use of the 'follow-the-money' tool introduced by the GIZ evaluation unit to better trace spendings. This tool places spendings in relation to the intervention objectives hence providing the basis for an informed discussion about resource allocation within the project. For this purpose, the AV was requested to provide the latest cost-obligations report, an overview of staff involved relative to the different outputs, and the latest monitoring results prior to the field study. All this information was then transferred into the 'follow-the-money' tool by the evaluation team. At the beginning of the field mission, the evaluation team additionally requested access to all financing and consulting contracts in order to map all project costs in relation to the different outputs. Given the number of contracts and the various underlying procurement procedures (e.g. via the project office or the country office), this exercise posed a burden on the project team in terms of time and logistics. Moreover, neither OBB nor target-performance planning ('*Soll-ist-Planung*') were mandatory in terms of GIZ planning requirements before the common procedural reform in 2017, which further complicated the task.

4.3 Evaluation process

The evaluation included an inception phase, a data collection phase, and an analysis and reporting phase. The inception phase took place in November 2017 and included the clarification of roles in the evaluation team, informational interviews with the GIZ evaluation unit, the project team and BMZ, a desk study and the elaboration of the inception report. The data collection phase mainly revolved around the field mission in Lusaka, which was carried out 8–22 January 2018. The analysis and reporting phase started at the end of the field mission. The final report was submitted to GIZ at the end of March 2018.

Overall, the evaluation team considers that the evaluation process went smoothly. As described in the previous sections, relevant documentation was available for analysis, and the evaluation team managed to talk to almost all relevant interview partners. The evaluation process was however characterised by two minor challenges that are described in the following:

- The evaluation could not live up to the initial aspirations of GIZ to involve the partners in the inception phase. The project team decided against sharing the inception report with the partners, as it was deemed too technical. The evaluation team shares this view. The format for the inception report requires the evaluation team to address a number of technical aspects, ranging from an assessment of a project's adherence to GIZ internal standards, to the discussion of different evaluation theories. This makes the inception report potentially difficult to digest for a lay audience. A further aspect that made it difficult to share the inception report is that no one in the partner institutions, neither on the political nor technical level, is fully aware of the entire range of the project. The project works with different units in three institutions (MoF, MNDP, ZRA). Although all partners were made aware of the entire results model and associated indicators by the project at the beginning of the project phase, each partner is typically only aware of the parts of the project that concern him or her. While the evaluation team deems this sensible given the partners' scarce resources, this makes a participatory approach to evaluation all the more difficult.
- Furthermore, no participation of the general population as the ultimate target group was expected in the evaluation, as the gap between advisory services at the level of partner institutions and results at the level of the population is substantial. To compensate this shortcoming, civil society organisations and secondary literature were consulted to put the potential long-term impact of the project into perspective.

5 Assessment of the project's results (OECD/DAC criteria)

5.1 Relevance

The relevance criterion examines the extent to which a development intervention addresses relevant needs of the partner and the target group. Therefore, this section first presents the relevant framework conditions and strategic documents including UN strategies and international-level principles, BMZ strategy documents, strategic documents of the Zambian government and the perspective of the target group – both the general population as the ultimate target group and the partner institutions as immediate target group. We then assess for each of these levels to what extent the intervention fits into the framework conditions in both the planning and in implementation of the project.

The main need of the ultimate target group in this respect is the effective and efficient provision of public services to the benefit of the general population. In relation to this, the partner institutions as the immediate target group, have a need in establishing an effectively, efficiently and transparently functioning PFM system in order to deliver these public services.

The key strategic documents and principles at the international level are the UN Agenda 2030 SDGs and its underlying principles such as LNOB, do-no-harm and gender equality. On the side of the German ministry, the *BMZ Strategy on Good Financial Governance* (see BMZ 2014), the BMZ *Country Strategy for Zambia* (see BMZ 2016) and the *Digital Agenda* are considered as reference points for assessing its relevance. With regards to the strategic framework within the Zambian government, key documents are the *National Planning and Budgeting Policy*, which forms the basis for the government's *Public Financial Management (PFM) Reform Strategy* (see GRZ 2016), the expected *National Planning and Budgeting Bill,* the *7th National Development Plan (NDP)* (see MNDP 2017), the long-term *Vision 2030* (see GRZ 2006) that the NDPs are based upon, the *Economic Stabilization and Growth Program – Zambia Plus* (see MoF 2017a) and the *National Performance Framework* (see GRZ 2016). The perspectives of the target groups were assessed by civil society stakeholders and through secondary sources (see ZIPAR 2017; Beardsworth 2017). While some of the key strategic documents including the 7th NDP and the Economic Stabilization and Growth Program were only passed or published after beginning of the project, they offer important reference points.

International-level intervention is aligned with the UN framework conditions, most importantly manifested in SDG 16.6 ('develop effective, accountable and transparent institutions at all levels'). The intervention's module objective is making clear reference to this objective ('transparency, credibility and effectiveness of public finances have increased') and all outputs seek to contribute to one or several aspects stated in this objective: action area I working towards enhanced credibility and effectiveness through domestic resource mobilisation; action areas II.a and III working towards budget credibility, effectiveness and, to a lesser extent, transparency through better budget planning, improved coordination and enhanced processes; and action area II.b working towards budget credibility and transparency by supporting the linking of national plans to the budgeting process.

The intervention thereby seeks to contribute to enhancing economic sustainability of Zambia's PFM system. Analysing the application of the LNOB principle in conceptualising the intervention, the specific needs of parts of the population including women or marginalised groups were partly prioritised in the conception of the intervention. While the development of gender-sensitive budgeting guidelines was foreseen in planning the project (interview 12, 15, 21) and considerations of potential negative impacts on specific groups were reported (interviews 11, 14), a systematic analysis of risks regarding the LNOB principle and the subsequent translation into mitigation strategies or other operational steps had not taken place. This appears to be a consequence of the evaluation of the previous phase (see GIZ 2015a, interview 6), which recommended a shift away from the development orientation towards more technical aspects of enhancing budget credibility.

In terms of the embedding into the German government's strategic framework, the intervention addresses core needs identified in the BMZ strategic documents such as the BMZ *Strategy on Good Financial Governance*, the BMZ *Country Strategy for Zambia* and the BMZ *Digital Agenda*. The *Strategy on Good Financial Governance* is reflected in both planning (see the project proposal, GIZ 2015a, p. 16) and implementation of the intervention, since the project addresses both the budget planning and execution process as well as domestic resource mobilisation. In particular, it focuses on technical aspects of the PFM framework through procedural and legal advice, while it puts less emphasis on the normative and political dimensions of the strategy. This approach is in line with the BMZ strategy's approach to select an appropriate set of measures depending on the context (BMZ 2014, p. 10). This strategic focus also reflects the partner government's request for technical advice at the time of the intervention's planning and takes into account the evaluation report of the previous measure, which recommends more focus on addressing the technical aspect of budget credibility (see GIZ 2015a, p. 11).

The BMZ *Country Strategy* did not exist at the time of conceptionalising the intervention; nevertheless, it matches well with activities and objectives of the intervention (strengthening accountability and good governance) (see BMZ 2016; interview 8). In addition, the intervention has been applying relevant aspects of the BMZ's newly adopted *Digital Agenda*. Particularly goals 1.1 on enhancing IT infrastructure and 1.4 on modernising public institutions by integrating digital solutions, such as an electronic queue-management (EQM) system to improve taxpayer services in action area I or the support of IT solutions, as in the planned MIS for M&E in action areas II and III (see BMZ 2017, pp. 12–13).

The intervention is also in line with the partner government's strategic documents including the *PFM Reform Strategy*, the medium and long-term development strategies and the relevant government programmes and laws. First, the *PFM Reform Strategy*, which was developed in the framework of the World Bank-managed PFM reform programme under the Multi-Donor Trust Fund (MDTF), forms the basis for the government's efforts to increase efficiency and accountability of public resources. The project design of this intervention picked up some aspects of five of the ten PFM reform programme components (integrated planning and budgeting, domestic revenues, IFMIS and cash management, enhanced internal audit and control, and monitoring and evaluation) (see GRZ 2013, p. 6). It works closely with the PFM reform coordinator under the Accountant General's Office in the MoF (interviews 14, 21).

Second, both the long-term (*Vision 2030*) and the medium-term (*7th NDP*) key strategic documents of the government are reflected in the activities that the intervention is pursuing. In action area II.b, the intervention is supporting MNDP in breaking down the strategic plans into key performance indicators and data catalogues in line with the national and sector performance frameworks in order to make the budget planning process ultimately more transparent and measurable (Interviews 13, 24).

Third, the *Economic Stabilization and Growth Program* is a medium-term action plan with the goal of 'restoring fiscal fitness for sustained inclusive growth and development', which is reflected in crucial activities that the intervention is supporting its partner in (see MoF 2017a). For example, the project is addressing the issue of budget credibility, i.e. reducing deviations of actual budget from parliamentary appropriations (ibid., p. 3). Moreover, the project is supporting domestic resource mobilisation by facilitating an increase in taxpayer compliance through automation and enhanced taxpayer services (see GRZ 2017, p. 6), which the intervention is working on in action area I. Also, the programme seeks 'raising the levels of accountability and transparency in the allocation and use of public finances' (ibid. p. 8), which are key points reflected in the module and programme objectives of the intervention. Lastly, the project is building its activities in action areas II and III on the expected *National Planning and Budgeting Bill*, which is seen as a crucial reference point for the PFM reform process (interviews 3, 29, 37).

In spite of the clear anchoring of the project in the Zambian government's strategic framework documents, the actual political will for advancing the PFM reform agenda might be lower than the formal commitments described above suggest. Evidence for this can be found in the slow progress regarding PFM reform, which put the written commitments into perspective. For example, the National Planning and Budgeting Bill, which was expected to be passed in 2015 by the project proposal document (see GIZ 2015b, p. 6), has still not been adopted. Moreover, the slow process of establishing and empowering the MNDP (interviews 35, 38) and the governments increasing external debt assessed unsustainably high by IMF (see IMF 2017, p. 12) indicate a discrepancy between documents and reality.

In terms of the important role that a functioning PFM framework plays for all policy sectors, the intervention has relevance for all policy sectors beyond the immediate partner institutions. In practice, a number of sector ministries are already involved in the piloting of more effective planning and monitoring tools. Regarding the former, OBB is currently being rolled out in three ministries (two Ministries of Education, and the Ministry of Community Development in 2018) after training had been conducted in five ministries in 2017 (interviews 12, 22). Regarding the latter, MIS implementation is ongoing, albeit with delays, in five ministries together with DFID and EU support (interviews 24, 30, 35).

Considering both the immediate (partner institutions) and ultimate target group, the intervention's objective is addressing the target groups' needs as it is working towards the sustainable management of public finances. Issues of transparency, accountability and efficiency and the public debt are reported to pose a key challenge for public service provision (interviews 12, 33, 35). Hence, given the challenging PFM framework (see Chaponda et al. 2017), the objective of the project is meeting the partner institution's need to ensure effective and efficient provision of public services. Even more so, the need of the general population as ultimate target group is addressed by the objective, as it depends on the provision of public services and suffers from unclear or ineffective PFM practices. In this context, several external stakeholders viewed the approach of supporting domestic revenue generation as well as budget planning and execution processes coupled with enhanced monitoring of national plans and their translation into the budget as relevant (interviews 28, 29, 32, 33). Likewise, these views could be confirmed by the literature review covering Zambian research institutions (ZIPAR 2018 budget analysis (see ZIPAR 2017)), international organisations (IMF Article IV Consultations (see IMF 2017), and the German research organisation, GIGA Political-Economic Short Analysis (see Beardsworth 2017).

In terms of adapting to the changes that have occurred in the framework conditions, the project was partly successful. Key contextual changes – most notably organisational changes in the partner structure, a growing public debt and changes within the donor landscape – were addressed by the project on the technical, but not on the strategic level. These three different dimensions of changing framework conditions will be discussed separately, as follows.

- First, relating to organisational changes, planning activities were separated from MoF to the newly established MNDP. This considerable reorganisation of ZRA took place after the appointment of a new Commissioner General, and the Accountant General was promoted to the level of Permanent Secretary. The formation of the new MNDP is an ongoing process, which slowed down processes and which was reported to have weakened the political weight of the planning function (interviews 30, 38). While the project addressed this change by adapting the inherent ToC, it was not translated into an updated results model; and it did not succeed in working towards the formation of a new steering structure (interview 24, 30). The ZRA restructuring and the Accountant General's promotion were organisational changes that the project adapted to in its activities and that helped leverage the project's results (interviews 11, 16, 14, 21).
- Second, the public debt, which was expected to be reduced over the course of the intervention, had remained high with the tendency to increase in light of new external loans to finance infrastructure

projects.[3] The IMF has identified a sharp increase in external debt since 2011 (see IMF 2017b p. 2). The intervention in its initial project design as well as the tactical adjustments has focused on seeking a technical solution through sophisticated procedural and administrative advice and support. This deliberate strategy, which in turn entailed a neglect of the political and normative dimensions of good governance might have contributed to the high level of appreciation the intervention experiences among the partner. Yet, it also became evident that seeking a merely technical solution will not alone be sufficient to solve the political problem of unsustainable and intransparent PFM. While technical aspects are important to improve processes, their impact remains limited without sufficient political will (interviews 15, 33, 35).

- Third, political issues and changes within the donor landscape occurred, which had an effect on the modalities of cooperation and potential for impact. On the one side, Western cooperation partners appear to be losing influence within the Zambian government compared to other players, who provide financing opportunities with less conditionality (interviews 28, 36, 38). This shift in the donor landscape away from the traditional Western cooperation approach is a topic high on the international actors' agenda (interviews 10, 15, 28, 36). On the other side, although formally aligned on objectives, there are still challenges in the partners' coordination with other international actors. Some interviewed stakeholders voiced concerns that the alignment of interventions may result in a competition among donors for influence in supporting future PFM reform in Zambia (interviews 13, 15). Further challenges were observed in terms of the legal status of the cooperation agreement with DFID in action area II.b. The co-funding from DFID was agreed upon in a Delegated Cooperation Agreement between DFID and BMZ. It was discovered that such an agreement is non-binding in international law, leaving room for interpretation of the cooperation modalities and funding commitments. This significantly complicated implementation as every activity of the project needed to be agreed upon between DFID, GIZ and MNDP on an ad hoc basis in comparison to co-funding agreements with other donors where implementation is fully delegated to GIZ. The project brought the implementation risks surrounding any cooperation agreement with DFID to the attention of the GIZ head office. It is planned to agree on a solution between the heads of DFID and GIZ in the near future. At the same time, the project was not fully successful at solving related issues such as facilitating the establishment of a functioning steering committee with MNDP (interviews 21, 30, 34).

The intervention staff managed to address evolving partner needs by analysing the technical aspects of changes within the partner organisations beyond the structural changes discussed above. In action area I, it was decided to support the clean-up of the taxpayer database, once it became evident that the outdated database poses a fundamental challenge to effective tax auditing (interviews 11, 17). In action area II.a, the programme was flexibly providing technical and legal advice on drafting the excess expenditure (limitation and conditions) bill that was then fused into the planning and budget bill, while the National Planning and Budgeting Bill was on hold (interviews 12, 22). And in action area III, a peer-learning visit was organised with the Republic of South Africa National Treasury to work on cash management issues at the request of the MoF (interviews 14, 21). The intervention's progress reports reflect the organisational (e.g. separation of planning function from MoF, restructuring of ZRA) and contextual changes (tightening public debt situation, delayed National Planning and Budgeting Bill), the resulting evolving partner needs and the technical adaptations mentioned above (see 2016b and 2017b).

Finally, at the time of planning, the design of the project was adequately adapted to the chosen goal. But in view of the changes of the political framework conditions, its technical nature may not be sufficient to address the target group's need of a transparent, credible and effective PFM system over the long term. The intervention's focus on the technical dimension of good financial governance was as a result of a partner

[3] While exact figures are not known to the public, different reports highlight the grown external debt. See, for example, these sources retrieved from the internet on 1 February 2018: https://www.bloomberg.com/news/articles/2017-10-12/debt-management-should-be-priority-for-zambia-kalyalya-says and https://www.lusakatimes.com/2017/10/11/no-imf-bail-zambia-fund-says-country-high-risk-debt-distress/

request and the previous phase's evaluation report, which recommended focusing on the critical aspect of credibility in the budgeting process (interviews 15, 37; see GIZ 2014b). While the BMZ Strategy on Good Financial Governance suggests that the focus on specific aspects (such as the technical dimension in this case) can be sensible in certain contexts, it was reported that addressing the normative and political dimensions may need more attention to maximise impact (interviews 5, 36, 38). Within the limits of the technical nature of the project, the project has adapted to this by working towards increased political ownership over the budget through pre-cabinet meetings (interviews 12, 22) and plans to further address these issues in the coming project phase.

Overall assessment of relevance

The evaluation team concludes that the project fits well into the relevant strategic reference framework. It reflects and builds upon key strategic documents from the Zambian government, the BMZ as well as the SDGs. What needs to be pointed out is the potential discrepancy between formal and de facto strategic commitments by the Zambian government to PFM reform. More specifically, the reported slow pace of the PFM reform progress and the debt the government is accumulating put the government's commitment to transparent and effective PFM into perspective. Therefore, the fit of the project into the relevant strategic reference frameworks is rated high, but not at a maximum, with 37 of 40 points.

Regarding the suitability of the strategy to match core needs of the target group, the intervention is considered highly relevant in terms of working towards the sustainable management of public finances, which address both core needs of the government and the general population. The latter, however, is not addressed in a differentiated way by considering the needs of marginalised groups. Given that questions regarding safeguards are not yet considered for accountability aspects at this stage,[4] and unintended results will be further discussed below in Section 5.3 on impact, the intervention's overall relevance to the international framework conditions is rated high with 25 of 30 points.

The design of the project is assessed as mostly adapted to the chosen goal. The conceptualisation at the time of planning was plausible and reflected the needs and priorities of the partner. In view of the changing framework conditions, the focus on the technical dimension of good financial governance may however not be sufficient to address PFM reform sustainably. Given that the project has initiated addressing such issues within the limits of the technical dimension of the project, this aspect is awarded 15 of 20 points.

Finally, the adaptation of the conceptual design to changes is assessed as mostly successful. The project team adapted to evolving partner needs and changing framework conditions by adjusting activities and leveraging results. However, it did not manage to fully adapt to these through more strategic measures such as facilitating the establishment of a functioning steering structure or fully updating the results model, resulting in a rating of 7 of 10 points.

The overall score for the assessment criterion relevance adds up to 84 out of 100 points: successful.

[4] The GIZ 'safeguards & gender management system' referred to was developed after conceptualization of this intervention. Therefore, the GIZ evaluation unit had pointed out that the respective questions have informative rather than evaluative character in this evaluation (interview 1).

Criterion	Assessment dimension	Score
Relevance	The project fits into the relevant strategic reference frameworks	*37/40 points*
	Suitability of the strategy? the conception? to match core problems/needs of the target groups	*25/30 points*
	The design of the project is adequately adapted to the chosen goal	*15/20 points*
	The conceptual design of the project was adapted to changes in line with requirements and re-adapted where applicable	*7/10 points*
Overall rating relevance		*84 of 100 points*

5.2 Effectiveness

The effectiveness criterion measures progress against the project's objective. This includes an analysis of the degree to which the outcome indicators are achieved, as well as an analysis of the extent to which the project successfully contributes to the achievement of the objective. In addition, the occurrence of additional, not formally agreed, results is examined under effectiveness.

A necessary condition for using indicators as a basis for assessment is that they fulfil the SMART quality criteria. Of the three outcome indicators (M1 and M3) formulated in the project's proposal, two were considered fully SMART by the evaluation team, and one (M2) partly SMART. Therefore, indicator M2 was rephrased during the inception phase and agreed upon for the assessment of effectiveness during the inception phase of the evaluation:

> *The proportion of individual budget heads that differ more than 5% between the originally approved budget by parliament and actual expenditure incurred by MPSAs has diminished by a quarter.*

- Baseline value: 90% (47 of 52) of heads deviated more than 5% from their appropriation in 2015
- Target value: 68% (= 90% x 75%)

Moreover, the inception phase had found four output indicators not to be fully SMART, which had partly been updated in the meantime. Indicator A1 has been modified, specifying that two measures of the SME taxation strategy should be implemented. Indicator B2 was clarified in the briefing session, where it was better explained what is in the scope of the suggested measures. Hence these two indicators can now be considered SMART. Indicator C2 is still containing a missing value on the number of training participants. While the project team explained that training could not yet take place, because the necessary risk management unit had not yet been established by the government, the target value should nevertheless be stated for effective monitoring. Finally, the inception report had raised the question whether output indicator

A2 is sufficient to measure actual results of the project in terms of tax audit improvements, which will be further discussed below in the analysis of action area I.

With regard to achieving the module indicators, the project is on track. Indicator M1 on increasing medium taxpayer compliance for VAT to 60% was already achieved in 2017 with 63% for the northern region and 59% for the southern region.[5] Relating to indicator M2, which seeks to diminish the share of individual budget heads that differ more than 5% between approved and actual expenditure by a quarter to 68%, the share was 82% for 2016 compared to 90% in 2015. The figures for 2017 were not available at the time of the field study, but the partner could confirm the positive development and the project expects that the target will be met by the end of the project (interviews 12, 22, 37). It was, however, conceded that there is uncertainty involved, for example, because emergency situations may require additional resources (interviews 12, 37). Also, indicator M3, seeking to apply a rule-based procedure for one ministerial budget, is expected to be achieved by the end of the project. A rule-based warrant system, which has been developed for the two Ministries of Education in parallel to rolling out OBB, has been piloted to replace the current cash rationing system. It is expected to be reviewed in May 2018 by an external consultant with MDTF funding. If successfully operating, this should allow the budget office to fulfil its core function of facilitating the formulation and oversight of a credible budget rather than engaging in short-term budget allocation (interviews 12, 22; MoF 2016).

Progress towards the project's stated and, where applicable, adapted indicators will be analysed below for the different action areas separately. Due to the operational separation of activities under output area II described above (see Section 3.2), the analysis will also distinguish between action areas II.a and II.b. In each case, the degree to which alternative factors play a role in explaining the results achieved will be examined on both the output and outcome levels, in order to come to a qualified assessment of the contribution that the intervention is making towards reaching the module objective of '[increased] transparency, credibility and effectiveness of public finances'.

In **action area I (domestic revenue generation)**, the objective is to increase ZRA's capacity to raise compliance of SME taxpayers (output A). This is to be achieved by increasing the quantity and quality of tax audits, by improving taxpayer services and by enhancing ZRA's strategy for the taxation of SME. At the outcome level, the success of this action area is measured by indicator M1, which seeks to increase medium taxpayer compliance regarding VAT to 60%. The project supports ZRA in achieving the objective through the provision of technical advisory services (e.g. on the development of the SME taxation strategy) and by financing specific additional activities or procurements deemed crucial for improving taxpayer compliance (e.g. recruitment of temporary ZRA staff to clean-up the taxpayer database; procurement and setup of an EQM system). In terms of modes of delivery, the services are provided mainly by the project's three advisors, two of them international and one of them national.

Based on the Tax Administration Diagnostic Assessment Tool (TADAT) prepared in early 2016 with support from different international partners, the intervention provided recommendations to ZRA on operationalising the findings made in the assessment (Rojas et al. 2016). These recommendations included cleaning up the taxpayer database to have a more accurate foundation for tax audits, simplifying VAT collection through a reverse charge to reduce delays and increase compliance, and supporting voluntary compliance of taxpayers through different measures. In line with the recommendations provided based on the TADAT report, findings of short-term consultants from the Bavarian tax authorities, who were invited by the project to conduct training on indirect tax audit methods, identified a need to enhance the taxpayer database (see Neumann 2017). Based on these recommendations, the project supported the clean-up of parts of ZRA's taxpayer database through the recruitment of temporary ZRA staff. More concretely, workstations and computers were provided for 10 interns, which were selected by ZRA and financed by the project through a financial agreement (interviews 17, 27). Technical advice to the senior management of ZRA raised awareness on the benefits and

[5] This information from the GIZ monitoring results was triangulated with reports provided by ZRA. The figures from ZRA were stated in a word document without specifying sources or data collection methods, which puts the reliability of the data into question. Several interviews could confirm the increase, however (interviews 11, 17, 27).

technical aspects of introducing the reverse charge for VAT collection (interviews 11, 16). In order to address the last aspect of voluntary compliance, the project supported ZRA in developing tax education materials with the education ministry, but these have still not been integrated into the school curricula (interviews 2,18).

Besides the database clean-up, the project worked towards enhancing SME tax audits by conducting training on indirect tax auditing methods and revising the existing audit training modules and materials (interviews 17, 27). This capacity development was reported to be instrumental for ZRA, because it provided tax auditors with new methods and tools (e.g. a template for indirect tax audit methods and sector notes with audit recommendations for specific sectors), which were deemed very useful for improving the quality of tax audits (interviews 17, 18, 19). With regard to output indicator A2 on the quantity of SME tax audits, the field study confirmed the inception report's suspicion that a quantitative measure does not accurately reflect project progress. While the number of tax audits does influence taxpayer compliance (module indicator M1), it is a function of ZRA's resources and management decisions that are beyond the scope of the project (interview 16). The qualitative improvements of tax audits in terms of more accurate taxpayer data and tax audit quality were nevertheless plausibly explained (interviews 11, 16, 17, 18, 20).

Another aspect of the project's approach to increase SME taxpayer compliance is its effort to improve service quality and to make compliance easier. In this respect, the procurement and introduction of the EQM system was a visible tool to reduce disorder and waiting times at taxpayer offices, and it helped ZRA to better monitor and manage its service offices (interviews 11, 16, 18, 19, 20). For example, average waiting times were reportedly reduced from 2:30 hours to 20 minutes (ZRA data generated by EQM system, see ZRA 2017). Moreover, the project had worked on two public-private partnerships (PPP) projects together with German IT companies aiming to make tax registration and returns more accessible via technology. One of these projects will be continued beyond this project phase with additional funding (interview 11).

Finally, the strategic advice provided by the project supported the development of the SME taxation strategy that formulated targets and measures for further increasing SME taxpayer compliance in the medium term (ZRA 2017). In particular, the project's technical advice helped identify key areas of improvement, for example, to simplify tax regulation for small businesses and to facilitate their compliances via service points on markets or via mobile applications (interviews 17, 27). The strategy was developed in a workshop facilitated by the project and adopted by ZRA at the end of 2017. ZRA requested support from GIZ to implement two of the 50 measures set out in the strategy, which forms the basis for project activities in the remaining year of the project (interview 11).

To better assess the contribution that the above-described activities and results have made to the objective of increasing SME taxpayer compliance, they were juxtaposed with a number of alternative explanations and discussed with the interview partners as well as internally among the evaluation team. Looking at the outcome indicator M1, we see that the following alternative explanations were carved out as being relevant: (1) The organisational restructuring initiated by the new Commissioner General in 2016 had initially resulted in a drop in the number of tax audits in 2016, but subsequently both quantity and quality of audits as well as VAT compliance for medium taxpayers had increased due to the more effective organisation of ZRA units under a direct and an indirect tax division (interviews 11, 16). (2) The introduction of online tax registration in 2016 and a tax amnesty introduced in 2017 had created pressure and incentives respectively on SME to comply with the tax legislation (interviews 16, 17). (3) New reverse VAT collection practices were introduced, which the project had influence on but without being a driving force. Through the MSC approach it was revealed that some interview partners viewed the project's activities, in particular the introduction of the EQM system, as the most significant positive development of the past two years (interviews 19, 20). Others considered external factors most significant, in particular the organisational restructuring (interview 27) and the reverse VAT collection practice (interview 16). Nevertheless, the majority of interview partners, including the latter two, reported that without the project's contribution to the quality of tax audits and taxpayer services, the achievements would not have been possible (interviews 16, 18, 19, 20, 27).

In **action area II.a (budget preparation)**, the central hypothesis is that the enhancement of legal and administrative processes as well as supporting the piloting and roll-out of OBB help increase the capacity of MoF to prepare the budget in accordance with the future national planning and budgeting legislation (output B). On the outcome level, the action area aims to contribute mainly to indicator M2 ('the share of individual budget heads that differ more than 5% between approved and actual expenditure has diminished by a quarter') and partly also to indicator M3 ('a rule-based procedure for the allocation of funds […] is applied for one ministerial budget'). In terms of activities, strong focus was put on providing technical advice to the MoF's budget office and other relevant stakeholders. These advisory services were implemented by one international senior advisor in this area, which will be complemented in 2018 by external consultants, which the government has procured with funding from the MDTF, to work on the reconfiguration of the IFMIS module.

Several measures to improve budget planning were introduced with project support, which form part of output indicator B2. The senior advisor's legal and technical advice regarding the legislative framework helped shape the drafting of the excess expenditure (limitation and conditions) bill that was then fused into the planning and budget bill, while the National Budgeting and Planning Bill was on hold (interviews 12, 22). The project supported the MoF budget office in using provisions of the new constitution to increase legislative oversight. As a result, the advisory services helped reduce the size of supplementary budgets and excess expenditures, which marks an important contribution towards module indicator M2 on reducing the deviation between approved and actual expenditures (interviews 12, 22).

In order to help MPSAs formulate more credible and effective budgets, the project supports the piloting and roll-out of OBB, which links the budget back to national and sectoral plans. It had initially been piloted in the education ministries outside the framework of this project and independently evaluated, with financing in early 2016 (see Bird et al. 2016). Based on this evaluation, the project facilitated the development of OBB developers' and user manuals, financed through the MDTF, by providing inputs to the assignment and conducting training with additional ministries to roll it out. For the time being, in addition to optimising implementation with the education ministries, OBB has been introduced in the Ministry of Community Development and Social Welfare for the 2018 budget and more MPSAs are planned for the 2019 budget (interview 12). Based on OBB, the project further aims to increase MoF capacities through the budget challenge function. For this purpose, a budget challenge manual is planned to be developed in 2018 that supports the analysis of ministerial budgets. In a second step, cluster hearings were facilitated that are jointly chaired by the MoF budget office and MNDP with the goal of analysing and challenging the estimated budgets of MPSAs around NDP clusters (interviews 12, 22).

Beyond the direct technical partner level, the project also reached out at the political and the citizen's levels. In order to enhance the political ownership of the budget, the project supported budget office in re-designing the supplementary budget and excess expenditure process to deepen the cabinet's understanding of the consequences of in-year political decisions. In addition, pre-cabinet meetings were initiated with ministers to discuss budgetary matters separately and interactively before formal cabinet meetings, which the partner reported to be very effective (interview 22). At the citizen's level, the project provided technical advice to MoF to produce and disseminate the citizen's budget in a timely manner. This activity was complemented by the GIZ civil society's programme, which supported civil society organisations in submitting proposals to the national budget (interview 31).

Indicator B1 (6 of a total of 54 MPSAs prepare their budgets in the IFMIS budget module), however, was found to be unsuitable during the field mission, as its achievement lies outside the sphere of influence of the project and no longer reflects its approach. The IFMIS budget module has been implemented in 44 MPSAs already, but it is not yet fully functional due to various shortcomings (interviews 12, 21). Given that the PFM reform programme under the World Bank-managed MDTF is pursuing the government-wide roll-out, this project has chosen to support the partner in identifying shortcomings to provide recommendations to the MDTF-funded consultants working on the IFMIS reconfiguration (interviews 11, 21, 22).

The issue of a more effective release of expenditure, which is connected to the achievement of module indicator M3, was addressed by supporting the budget office in moving from a system of cash rationing towards a warrant system. This warrant system was so far piloted in two MPSAs with the goal of helping the budget office to re-focus their work on budget analysis rather than the authorisation of expenditures (interview 12, MoF 2016). Difficulties emerged during piloting, because the effective use of a warrant system requires coordination between the relevant stakeholders including budget office, the cash management unit and the Accountant General's Office, which appears insufficient (interviews 12, 14). Besides, the concern was voiced that the PFM coordinator under the Accountant General's Office is not fully informed about reform activities in the budget office, indicating that limited coordination is a critical factor (interviews 3, 21).

Ultimately, other contributing factors were considered to put the project's contribution to the module objective into perspective. It was reported that the constitutional amendment from 2016, which required all in-year changes to the budget to be endorsed by parliament in supplementary budgets, also contributed to the reduced deviation of planned and actual expenditures (module indicator M2) (interview 29). Moreover, the results of the constitutional amendment may reflect more formal than actual improvements in budget credibility, because de facto excess expenditures may be formally approved through appropriation bills at the last minute (interviews 12, 22). However, beyond the formal effect it had, the new constitution could be used as an argumentative tool to further advance the reform agenda regarding supplementary and excess expenditures, hence making an effective contribution in itself (interviews 12, 37).

According to the project proposal and GIZ results model, **action area II.b (implementation of MIS)** equally contributes to output B (increase MoF capacity to prepare the budget in accordance with future national planning and budgeting legislation) and the respective module indicator M2. Output-level activities focus on supporting MNDP in breaking down long-term plans to performance indicators and developing the necessary IT infrastructures to link national planning with budgeting, which is measured in output indicator B3. Advisory services are conducted by one international advisor, who is complemented at times by external consultants. The DFID co-financing to the project is exclusively channelled into this action area.

The action area's inherent ToC had changed twice throughout the course of the project; however, this was not reflected in changes of the GIZ inherent indicator system. The first change was in 2014 with the plan to newly establish a comprehensive IT structure at the apex level (MNDP, MoF, cabinet) and five key ministries with support of the EU, DFID and GIZ. During the inception phase in 2015, it was found that this plan was too ambitious and so it was narrowed down to a four-pillar approach focusing on M&E, IT, capacity development and the governance structure. After another review at the request of DFID, a third ToC was finally formulated in July 2017, which focused on the apex level while limiting the efforts on establishing IT systems in the sector ministries. This approach was chosen with the goal of having a more realistic workplan in line with reduced budgetary commitments expected from DFID. While formally agreeing to this approach, however, DFID and GIZ were not aligned regarding the obligations under the ongoing delegated coordination agreement due to its non-legally binding status, which created uncertainty on behalf of the partner on how to proceed (interviews 15, 24, 30). Moreover, the results model and indicator system for this action area remained unchanged in spite of the aforementioned strategic changes and the fact that since the project started, the national planning function had moved from MoF to the newly established MNDP. This means that the project no longer contributes to increasing MoF capacity in this area (output B) but does to MNDP capacity. The implicit underlying hypothesis hence is that by developing national and sectoral performance indicators, by supporting the underlying processes and by eventually developing suitable IT systems, the capacity of MNDP to inform and guide the budget preparation process is increased.

Considerable progress was made regarding the first aspect of developing performance indicators. A common M&E strategy was developed, and subsequently MNDP was assisted in breaking down the long-term development plans of Vision 2030 into national and sector performance frameworks. Key performance indicators were defined and data catalogues were established, all of which contributed to strengthening MNDP monitoring capacity. This progress is also reflected in the gradual reduction of time lag for the publication of annual progress reports from almost 1.5 years for the 2014 to 8 months for the 2016 report, and the 2017 report is expected to come on time (interviews 13, 24).

With regard to the underlying processes, however, efforts to establish a suitable governance structure for MIS implementation with MNDP in lead and the accompanying capacity development activities could not be implemented as planned. In consequence, this resulted in delays regarding the implementation of IT systems (interview 13). While the various existing IT tools in the ministries were analysed in terms of the possibility to build an integrated IT system, the project did not get to the point of advancing on the MIS implementation. At the partner's request, the project then focused on supporting the enhancement of the Monthly Monitoring System, which was developed by the predecessor project as an interim solution (interviews 24, 25).

Investigating causes of the observed delays, the following core challenges to advancing with MIS implementation on the government side were identified: (1) MNDP, in particularly the M&E unit, was reported to be heavily understaffed. (2) A weak management capacity in MNDP combined with the lack of a crucially needed project management unit was identified. (3) It was observed that numerous IT systems and initiatives exist in the government without proper coordination (interviews 13, 30, 35). Moreover, it was elaborated that the rather new MNDP is not yet as functional as other ministries and that the coordination issues between DFID and GIZ on their respective obligations and the approach to take towards further implementation (see also relevance Section 5.1) due to legal uncertainty additionally impeded the project progress (interviews 13, 15, 23). (4) The lack of an effective steering committee further added to the partner finding it difficult to move forward with its reform agenda. In this context, it was reported that the partner was left unclear about the type and degree of support to be expected from the project (interview 24, 30), which could have been prevented by effective use of a steering committee or possibly earlier escalation to the higher management level. Ultimately, this combination of factors led to the understanding that intervention in this action area could not deliver what was initially expected.

In **action area III on budget execution including internal audit**, the objective is to raise the ability of the MoF to facilitate effective budget implementation in accordance with plans and the enacted budget (output C). This is to be achieved by optimising administrative procedures for budget implementation, by enhancing the internal audit function in line with international standards and by strengthening the risk management function. If successful, the results are expected to contribute to an increase in the effectiveness of public finances, which is indirectly reflected in module indicator M3. In terms of modes of delivery, advisory activities are implemented by one international advisor working with the Accountant General's Office and one national advisor having the Controller of Internal Audit Department as counterpart.

In spite of the delayed start of activities due to the project's initial lack of human resources in this area, the project has already contributed to more effective budget execution in several ways. First, a business process mapping for payroll activities was conducted by external consultants. It is currently under review by the working group of key stakeholders (Accountant General's Office, PSSD, Budget Office, Controller of Internal Audit, MoF HR and Smart Zambia) and two more business process mappings are planned on debt and investments (subject to approval by the IDM director) and on revenue. These business process mappings are assessed positively and are expected to help the Accountant General's Office to guide activities within the PFM reform programme on further advancing the IFMIS roadmap. Based on these business process maps, the roles and responsibilities of actors involved in budget implementation can be clarified (interviews 14, 21).

Second, the project supported the strengthening of the audit function. After completion of a first comprehensive IFMIS audit under the previous project phase, a second IFMIS audit was conducted by the Internal Audit Department with project support to analyse transactions and identify weaknesses still inherent in the current IFMIS. While the results of the audit were kept confidential by the Accountant General due to the sensitivity of the contents, the audit was considered helpful in identifying and addressing the IFMIS module shortcomings (interviews 21, 26). Additionally, the Internal Audit Department was supported in adhering to quality assurance in line with international standards through training and technical advice.

Building on the activities initiated in the previous phase of the intervention, a performance audit is planned of an agricultural subsidy support programme under the Ministry of Agriculture to test and further enhance the capacity of the Internal Audit Department (interviews 25, 26).

Third, to facilitate issues of cash management, a peer-learning visit of the South African Treasury was organised to enhance functionality of the newly established Treasury Single Account (TSA) with a key focus on cash flow forecasting. The partner reported the visit to be highly successful as it provided a first-hand learning experience and identified key areas for improvement regarding cash management (interviews 14, 21, 37). The visit was also important, as it brought together actors from the Accountant General's Office (key partner in action area III) the budget office (key partner in action area II.a) and Investment and Debt Management department, both of which had described experiencing a lack of mutual communication and coordination (interviews 21, 22). Enhanced coordination between these actors is considered crucial for the budget office to resume more of its core function of budget analysis and planning and allow the cash management unit under the Accountant General to deal with cash management (interview 14).

Last, the government-wide risk management framework that was first drafted in 2015 (see MoF 2015) had been passed by cabinet in 2017, but still needs to be operationalised. The process was spearheaded by the Internal Audit Department with the technical advice of the project as well as its predecessor phase (interviews 24, 25). The necessary risk management unit is, however, not yet in place; therefore no risk training has been conducted so far. While this unit was expected to be established in early 2018, the exact time of establishment remains unclear (interviews 14, 26). Once realised, the risk management system is expected to contribute to an increased capacity to effectively plan and implement the budget by being able to identify and systematically address risks (interviews 14, 25).

To put the project's contribution to improved budget execution into perspective, potential alternative explanations for improved budget execution were analysed and discussed with the interview partners. Two key changes outside the project's influence were identified over the last two years, which appeared to have also played a role. (1),The promotion of the Accountant General to Permanent Secretary level helped him and his office to fulfil their roles in budget execution more effectively, because it strengthened their position. (2) The establishment of the TSA was revealed as a crucial measure to enhance budget execution by improving appropriation control, by making cash management more efficient, and by allowing more complete and timely information on government resources. The key role of the TSA was confirmed through the MSC approach in interview 21. While both of these aspects were reported to be important factors for improved budget execution by different interview partners, it also meant that the project could leverage its activities through these changes (e.g. South African Treasury visit to enhance TSA, address shortcomings in IFMIS after audit) (interviews 14, 21, 34). Hence, the project did make a difference towards enhancing partner capacities on budget execution.

Analysing further success factors relevant to all **project levels**, both strengthening and inhibiting aspects were observed. Regarding the former, both partners and external stakeholders reported that the high technical expertise of the advisors in all areas contributed to project success (interviews 17, 21, 22, 27, 37). Moreover, it was positively noted that the project's approach flexibly follows functionality rather than form; this contributes to meeting the partners' requirements and helping them in setting the right priorities (interviews 14, 21, 22). At the process level, weekly internal team meetings as well as weekly partner meetings with the technical experts help coordinate processes effectively. Nevertheless, it was also found that no coherent steering structure exists for the project (interviews 15, 22), which bears a risk of critical challenges not being addressed adequately and on time, as appeared to be the case in action area II.b.

Relating to unintended results, additional results – whether positive or negative – were not reported, but also not specifically considered in the project's monitoring system. In terms of the former, no interview partner had observed any additional results except the risk for potential negative results discussed below (interviews 11, 18, 20). In terms of the latter, it was reported that the results monitoring was used as a reporting tool for progress towards milestones and the indicator goals, rather than as a management instrument (interview 6). While the project uses an Excel-based tool to monitor project progress on a weekly basis, no specific monitoring for unintended results was reported (interview 15).

It was found that risks were stated in the results model and brought up in the discussions with project team-members (e.g. political abuse of power, misappropriation of funds, decreasing weight of Western cooperation partners); however, they were not systematically addressed in project implementation (interviews 11, 14, 15). Especially action area I – which deals directly with domestic revenue generation and auditors of SME taxpayers – faces potential risks of implicitly supporting harmful or inappropriate practices, because ZRA is perceived to be exploited for political gains in some instances by parts of the wider public. Similarly, the other areas deal with sensitive areas of PFM, which carries a high risk of corruption. Although these issues are discussed at team meetings, no strategy exists to deal with such risks effectively in terms of an internal risk management, exit strategies or safeguard mechanisms (interview 15). The project team reported that it plans to address such issues more systematically in the next programme phase from 2019.

Another potential negative unintended result is the risk of limiting the partner's ownership by driving the PFM reform agenda too much from the outside, rather than letting the partner government set the right priorities by themselves (interview 14). This risk applies to this project as much as to others, particularly the PFM reform programme, and it is also reflected in the observation that the actual political will may be lower than the formal commitment (see Section 5.1 Relevance). Given the divisions of labour within the German development cooperation with BMZ funding and negotiating projects and GIZ implementing them, this risk can be mitigated only at the political level of the intergovernmental consultation process. In that regard, however, the project has limited influence, therefore it will not be considered for the assessment of the project's effectiveness.

Overall assessment of effectiveness

Balancing all aspects, the project is assessed to mostly achieve the objective on time and in accordance with the TC-measures' goal indicators. The stated outcome indicators are all expected to be achieved by the end of the project. Outcome indicator M1 has already been achieved and the project is on track towards achieving indicators M2 and M3.

Despite the good progress towards meeting the outcome indicators, however, the design of the intervention is not sufficiently adapted to the chosen goals. In particular, a number of output indicators do not fully reflect the results of the project. In action area I, the project successfully contributes to increasing the quality of tax audits, the improvement of taxpayer services and a more strategic approach to SME taxation, which will lead to an increase in taxpayer compliance. The stated output indicators, however, merely measure the quantity of SME tax audits and aspects of the SME taxation strategy. They neglect quality aspects of tax audits as well as taxpayer service improvements entirely, both of which were carved out as key success factors in the field study (interviews 11, 16, 17, 18, 19, 20) and reflected in framework documents (Article IV Consultations, IMF 2017, p. 13). In action area II.a, output indicator B2 offered sufficient flexibility to ensure the relevance of all measures taken for meeting the module objective and indicator, but output indicator B1 is considered inappropriate by the evaluation team to reflect progress towards the module objective, as its achievement lies outside the sphere of the intervention influence. The roll-out of the IFMIS budget module to the MPSAs is an activity pursued by the government with support of the World Bank-managed MDTF, which is beyond the scope of this project. The advisors in this intervention, therefore, adapted their strategy towards supporting the MDTF-funded consultants to reconfigure it accordingly at the request of the partner (interviews 12, 21, 22). In action area II.b, the discrepancy between the project proposal and the actual activities is most obvious. Despite the fact that these activities were initially planned to absorb more than half of the overall project budget (including co-funding from DFID designated for this action area) and despite the separation of the planning function from MoF, the results model in this area remained unchanged with only one output indicator and no module indicator. Hence, the project design no longer fully matches the stated objective. On a tactical level, the intervention has adapted to this mismatch, however. This was achieved partly by operating outside the proposal's output indicators but within the module objective's framework (as in action area I and II.a), and partly also without the necessary adaptation of the strategic framework, i.e. clarifying an independent results model with output and outcome indicators (as in action area II.b).

In light of the described limitations, the evaluation team awards 35 of 40 points to the first aspect of goal achievement despite the good progress towards the outcome objectives.

The services implemented by the project are assessed to successfully contribute to increasing the partner capacities in almost all areas. In every action area, the contribution analysis could carve out how the project's technical advisory services help increase partner capacity in line with the module objective. In action area II.b, the project fell short of expectations, however, owing to factors beyond (e.g. limited resources, management capacity and political weight of MNDP) and within the sphere of influence of the project (e.g. overly ambitious planning, limited coordination with DFID as co-funding partner). In this context, the lack of an effective steering committee aggravated the delays in progress, because critical issues were not appropriately addressed. Keeping these shortcomings in mind, the evaluation team awards 27 of 30 points for the successful contribution to the achievement of goals.

Finally, the occurrence of unintended positive or negative results is considered addressed to only a certain extent. While crucial risks were identified in the initial results model and reflected in some of the interviews with project staff, the project does not yet follow a systematic approach to deal with them. In conclusion, the evaluation team awards 20 of 30 points to this aspect.

The overall score for the assessment criterion effectiveness adds up to 82 out of 100 points: successful.

Criterion	Assessment dimension	Score
Effectiveness	The project achieves the goal on time in accordance with the TC-measures' goal indicators agreed upon in the contract.	*35/40 points*
	The services implemented by the project successfully contribute to the achievement of the goal agreed upon in the contract.	*27/30 points*
	The occurrence of additional (not formally agreed) positive results has been monitored and additional opportunities for further positive results have been seized. No project-related negative results have occurred – and if any negative results occurred, the project responded adequately.	*20/30 points*
Overall rating effectiveness		*82 of 100 points*

5.3 Impact

The impact criterion measures the extent to which the intervention contributes to the achievement of overarching development results. In this regard, the evaluation questions relate to the contribution to the programme-level impact indicators, and to the implementation of national (7th NDP) and international development agendas (SDGs). It is also assessed to what extent and how the general population as the ultimate target group of the intervention is affected by it and in how far unintended results at the superordinate level can be observed. With regard to the ultimate target group, however, there is an attribution gap between module objective (increased transparency, credibility and effectiveness of public finances) and the general population, which is not considered with any indicators in the indicator system. Moreover, population-level primary data could not be collected in the framework of this study, which draws on secondary sources and interview partners from civil society. Therefore, the explanatory power of analysis regarding impact for the population is limited.

To assess the programme-level impact, the evaluation team analysed the four impact indicators of the DC-measure, which combines the intervention subject to this evaluation and the KfW contribution to the World Bank-managed MDTF:

1. Taxes collected as a percentage of GDP have increased from 15.8% to 18%.

Baseline value: 15.8% (preliminary estimate Article IV Consultations of the International Monetary Fund, IMF, June 2015)
Target value: 18%

2. Zambia's fiscal deficit has decreased from 8.14% of the GDP to less than 5% of the GDP.
Baseline value: 8,14% (2015)
Target value: < 5%

3. Zambia's sovereign debt has stabilised at the level of approximately 40% of the GDP.
Baseline value: 60%
Target value: 40%

4. The Open Budget Index has improved from 39 in 2015 to 50 in 2018.
Baseline value: 39
Target value: 50

When looking at the macroeconomic impact indicators from the programme proposal, the announced superordinate long-term results are only partly expected to be achieved. An analysis of the latest IMF Article IV Consultations (see IMF 2017), which serves as source of verification for three of the four programme-level indicators, reveals the following results. Taxes collected as a percentage of GDP (indicator 1) had decreased from 15.8% to 12.9% in 2016 (preliminary data) (ibid. p. 31) while the target value is 18%. For 2017, however, ZRA reported that tax collection had overperformed resulting in a tax rate of 16.8% (interview 16). Also, fiscal data from the MoF budget office shows that domestic revenue (tax and non-tax) was 18.24% in 2016 depicts a better performance than the preliminary IMF data (see MoF 2017d). Hence, while achievement of the indicator against IMF data remains unclear, there seems to be a positive trend. Regarding indicator 3, the objective is unlikely to be achieved. The sovereign debt (indicator 3) remains high at 60.5% (ibid. p. 31) while the aim was to reduce it from 60% to 40% in 2018. The fiscal deficit (indicator 2) has been reduced from 8.14% to 5.8% on a cash basis for 2016 with the target being to reduce it to under 5% (ibid. p. 8). For 2017, it has been announced by the Minister of Finance to be 6.1% (MoF 2017). The fourth indicator, drawing on the international benchmark of the Open Budget Index on budget transparency dropped from 39 to 8 between 2015 and 2017,[6] falling considerably short of the objective of reaching 50 by the end of the project term. There has been a methodological shift in calculating the index between 2015 and 2017 though, resulting in a majority of countries falling in their rating.[7] The analysis below will have investigated to what extent the project made a contribution to these developments and what were other factors influencing them.

In light of the political factors that work against the establishment of a sustainable PFM system, the project can only make a limited contribution to wider macroeconomic impact. These influencing factors as well as the limits of the current project design to address the challenges will be analysed in the contribution analysis below.

A plausible link can be identified by analysing the contribution of the intervention's outputs and outcome objectives to the objectives at the impact level. Activities in action area I contribute to improvements on the tax collection ratio (programme indicator 1) as well as the public deficit and overall debt (indicators 2 and 3). Activities in action areas II and III contribute to developments regarding programme indicators 2, 3 and 4 on the Open Budget Index. Compared to external influence factors, however, most notably global economic developments such as the volatile copper price (generally a downward trend, but increased in 2017), increased external loans from countries outside the established donor coordination circles, and a political climate within Zambia characterised by intransparency and tightening government control (see Beardsworth

[6] Data from the internet accessed on 08.02.2017 at https://www.internationalbudget.org/open-budget-survey/results-by-country/country-info/?country=zm

[7] Retrieved from the internet on 29.03.2018 at https://www.internationalbudget.org/open-budget-survey/methodology/

2017), the project's activities were reported to play a rather small role in contributing to achieving these indicators (interviews 15, 28, 36).

In this context, the quantified assessment of PFM reform progress that all interview partners were asked to provide on a scale from one (no progress) to five (very good progress)[8] revealed both strengths and weaknesses of the project. On average respondents gave an assessment of little to medium progress (2.75), but the assessment considerably differed between the partner's perspective (almost good progress of 3.9), whereas all other stakeholders on average rated the progress as little (2.1). Given the low number of responses and the potentially biased view that the involved public servants may have on the reform progress, the figures should not be interpreted independently.

On the one hand, the discussions revealed that there are individual improvements that contribute to the programme-level impacts. For example, improved financial management through the TSA and IFMIS, which helps make PFM processes more efficient, make a contribution to impact indicators 2 and 3 (interviews 14, 21). It was further plausibly explained that the project did make a meaningful contribution to these by helping enhance the TSA through the peer-learning process with South Africa (interview 14, 21) and by providing advice on improving IFMIS functionality (interviews 12, 21).

On the other hand, it was reported by several external stakeholders that in spite of incremental improvements at the technical level, larger political challenges remained untouched. These include the limited political will for reform, a continuation of intransparent procurement practices, and an observed gap between written law and its practical application (see Beardsworth 2017). While these aspects mostly remain outside the scope of this project, they were reported to reduce the impact that the project's technical cooperation could have had (interviews 8, 28, 29, 33).

When looking at further factors that positively or negatively influenced impact, the interaction with other international players played a mixed role. The World Bank-managed programme under the USD 17 million MDTF had been the main driving force of the PFM reform process until now. Overall, this programme was reported to have made a contribution to PFM reform, for example with regard to the TSA or the roll-out of OBB and IFMIS (see World Bank 2017, PEFA 2017; interviews 21, 34). While the current project is formally aligned with the PFM reform programme and has tactically adapted to leveraging its results in many instances (see Section 5.2: Effectiveness), a closer cooperation could have contributed to better address impacts in a coordinated way (interview 34). For example, it was reported that parts of MoF that are involved in the PFM reform process under the MDTF are not sufficiently informed about the interactions between the MDTF-funded activities and the GIZ project (interview 21). Furthermore, the initial formulation of an output indicator (B1) that relates to activities in the scope of MDTF-funded PFM reform programme only indirectly related to GIZ project activities raises questions.

Moreover, the project came under scrutiny from co-funding partner DFID in action area II.b. While the DFID review in January 2017 had still reported good coordination and endorsed the advancement of the work plan (see DFID 2017, p. 20), more critical perspectives were voiced during the evaluation mission. Delayed processes were reported from within the government as well as insufficient or delayed coordination at a technical level, due to the character of the delegation agreement on behalf of DFID and GIZ (interview 24, 30). Conversely, it was also reported that DFID had reduced the budget due to domestic pressures, which had resulted in a lack of support to the IT roadmap and contributed to the delays (interviews 13, 15). While it appears that a number of factors have contributed to the delays, ultimately less than half of the planned co-funding budget was used in this action area. This partly remains outside the project's sphere of influence. Nevertheless it did limit the contribution this action area could have made to wider social impacts.

Looking beyond the coordination with other cooperation projects described above, external loans from third

[8] The full question provided in the interview guides was: How would you assess progress of the Zambian government in terms of establishing a sustainable, transparent and efficient public financial management system over the past two years? (1 – no progress, 2 – little progress, 3 – medium progress, 4 – good progress, 5 – very good progress). Only 14 of 27 interview partners provided a quantitative statement in additional to their explanations, while the others preferred to give a narrative assessment only.

countries for large infrastructure projects outside the regular budget framework were reported to increase stress on public finances as well as indicate a change in the political dynamics of international cooperation in Zambia. On the one hand, this resulted in raising questions of debt sustainability, which indicated a lack of political will for reform (interviews 28, 33). On the other, the government's confidence in taking these loans in spite of IMF's inability to enter into a programme with Zambia as a result of this additional borrowing was assessed as a sign of the decreasing weight of larger, Western-style cooperation (interviews 10, 15, 36). Put together, these points indicate that the project can only make a limited contribution to wider macroeconomic impact in light of the political forces that neglect the need of a sustainable PFM system.

In terms of implementing the national and international development agenda, the project supported the initiation of necessary steps. In action area II.b, the project supports the implementation of the SDGs through operationalisation of the 7th NDP by creating the National Performance Framework as a basis and by establishing key performance indicators and seeking to integrate national planning with budgeting. The process of defining the indicators was still ongoing at the time of the field study, and the next step is to plan their translation into the budget (interview 12). While delays in the implementation of M&E systems and necessary support structures in the government slow down the process in this area (interviews 24, 30), an analysis in 2017 has nevertheless revealed the link between the National Performance Framework and 7th NDP and the SDGs (interview 12). Hence, the project made a contribution to Zambia's advancement of the international development agenda.

The project's impact in terms of scaling up was limited by the reported government silo structures. While certain aspects of the project were successful in reaching out further than the immediate partners at the centre (e.g. EQM and new tax audit methods in the regions in action area I; involvement of other government agencies in piloting OBB in action area II.a; gradual roll-out of enhanced internal audit techniques to internal auditors in the line ministries in action area III), the project falls short of scaling up further to regions, government agencies or the wider population. This stems partly from the challenging political environment as described above, but partly from the silo structure within the government, which was brought up at multiple times during the field study (5,15, 23). In this context, it was reported that communication and coordination between different ministries, and even units within ministries, is often insufficient. For example, the PFM reform coordinator under the Accountant General's Office was said to lack information on activities of the budget office (interview 21). Given that both the Accountant General's Office and the budget office are supported by different action areas of the project, to some extent this silo structure was mirrored by the project, which limited the transfer of knowledge and processes to further levels (interview 5). Nevertheless, the project also advanced initiatives to bridge this silo structure; for example, the South African Treasury peer-learning visit brought together different stakeholders from within the MoF (interviews 14, 21).

Moreover, when looking at the target groups, the field study found strong links to the economic dimension of sustainability and limited links to the social and environmental dimensions. For this project, the economic dimension of sustainability is most relevant given the focus on good financial governance. Regarding the immediate target group (partner institutions), the project contributes to increasing the economic sustainability of public finances by enhancing legal and administrative processes of budget planning and execution (interviews 21, 24, 26). Social and ecological dimensions of sustainability are assumed to be addressed indirectly through the more efficient implementation of social and environmental policies (see BMZ 2014). For cross-cutting issues, which are important for addressing social and environmental aspects more directly, the project did not manage to achieve any results beyond the activity phase. And although gender-sensitive budgeting guidelines were developed, they were never implemented by the partner, presumably due to low prioritisation from the government (12, 15, 37).

Unintended results were not observed at the impact level beyond the two aspects of potential negative results discussed in the Section 5.2 on effectiveness that also influence the impact level. One such potential unintended impact is the risk of ZRA as a semi-autonomous authority being manipulated for political purposes. The project team reflected this risk in internal discussions and no evidence for actual negative effects was found; but to date it has not reported any measures to address and mitigate the risk. The concept note for the follow-up phase clarifies, however, that related risks are anticipated and plans close results-oriented monitoring and an exit strategy, if risks get too high. Another potential negative impact is that the government may be hindered in establishing a sustainable PFM system itself due to perceived push factors from interventions such as this one (interview 14). While government representatives spoke very positively of the PFM reform progress (interviews 21, 25,37), other stakeholders viewed it much more critically, arguing that the reform process is too slow and that the political system is holding back reform (interviews 32, 33). Given the tendency to demand tangible results and make support conditional on quick progress, it may therefore be a negative unintended result where the project is contributing to unrealistic prioritisations on behalf of the partner, which ultimately results in a lack of PFM reform (interview 14). To mitigate this risk of any cooperation project, however, the project's advisory approach deliberately seeks to manage expectations of the partner.

The assessment of the project's risk analysis and handling provides a mixed picture. The proposal includes a number of key risks (e.g. politically motivated reallocation of funds, insufficient rule of law, more profound limitations of IFMIS than expected) and suggests mitigation strategies, though it assessed the influencing ability of the project as low (GIZ 2015a, p. 18). The risks stated in the proposal were also brought up by interview partners (interviews 15, 32, 33, 36). Nevertheless, the project did not appear to follow a coherent strategy to address these risks. It rather sought to apply its technical advisory capacity flexibly, to mitigate the risks as much as possible on an ad hoc basis (interview 15). For the upcoming phase, however, the project is planning to include a number of aspects in the project design to better address existing risks in a structured way (interview 15, 36).[9]

Overall assessment of impact

With regard to the first assessment dimension, the table below shows that the impact indicators defined in the project proposal have only been achieved to a limited extent. However, external macroeconomic factors and the political situation play a large role in explaining these shortcomings. Moreover, the project was bound in its conceptual design to focus on the technical level, which limited the possibilities for addressing the political level. Overall, the evaluation team awards 25 of 40 points to the first question of meeting the foreseen superordinate long-term results.

With regard to the second assessment dimension, the evaluation finds that the project nevertheless contributed to the intended, long-term development results. In spite of the challenging macroeconomic and political environment, the project's technical advice managed to support crucial improvements to the PFM framework such as the TSA or enhancing legislative oversight over supplementary expenditures. The replication of the government's silo structures, however, led to limitations in the way impact could be achieved within the partner structure. Therefore, 27 of 30 points are awarded to this assessment dimension.

[9] The follow-up project seeks to address risks in the following way: Firstly, an internal risk management strategy is planned for the project that analyses key risk areas, develops mitigation options as well as an exit strategy. Secondly, it is sought to address transparency and accountability more proactively by including external actors such as civil society organizations to also address the political challenges. Lastly, the project aims to follow a more cluster-based approach in the next phase to break or at least better bridge the silo structure within the partner government (interviews 15, 36).

Indicator	Degree of achievement
1. Taxes collected as a percentage of GDP have increased from 15.8% to 18%. *Baseline value*: 15.8% (preliminary estimate Article IV Consultations of the International Monetary Fund, IMF, June 2015) *Target value*: 18% *Current value: 12.9% (preliminary data for 2016)*	Indicator may be achieved. While ZRA and MoF reported an increase in domestic revenue generation, achievement of the indicator against the target value as defined by IMF data remains unclear.
2. Zambia's fiscal deficit has decreased from 8.14% of the GDP to less than 5% of the GDP. *Baseline value*: 8.14% (2015) *Target value:* <5% *Current value: 5.8% (2016)*	Indicator expected to be achieved. The fiscal deficit has decreased to 5.8% on a cash basis for 2016.
3. Zambia's sovereign debt has stabilised at the level of approximately 40% of the GDP (preliminary estimate Article IV Consultations of the International Monetary Fund, IMF, June 2015) *Baseline value*: 60% (2015) *Target value*: 40%	Indicator not expected to be achieved. The sovereign debt remains with 60,5% (2016) roughly at the level of the baseline value.
4. The Open Budget Index has improved from 39 in 2015 to 50 in 2018. *Baseline value*: 39 (2015) *Target value*: 50% *Current value: Data on the 2017 survey are not available yet.*	Indicator not expected to be achieved. The international benchmark of the Open Budget Index on budget transparency dropped from 39 to 8 between 2015 and 2017, making the achievement of a ranking of 50 unlikely by the end of the project term.

With regard to the third assessment dimension, the evaluation team assesses the achievement and monitoring of unintended results as partly met. No unintended results, whether positive or negative, were actually observed. However, a few risks were identified that the project did reflect in discussions and is planning to address more strategically in the follow-up phase, but did not yet address by developing a mitigation strategy or adapting activities accordingly. Therefore, the evaluation team awards 20 of 30 points to this dimension.

The overall score for the assessment criterion impact adds up to 72 out of 100 points: rather successful.

Criterion	**Assessment dimension**	**Score**
Impact	The announced superordinate long-term results have occurred or are foreseen (should be plausibly explained).	25/40 points
	The project contributed to the intended superordinate long-term results.	27/30 points

	The occurrence of additional (not formally agreed) positive results has been monitored and additional opportunities for further positive results have been seized. No project-related negative results have occurred – and if any negative results occurred the project responded adequately.	20/30 points
Overall rating impact		72 of 100 points

5.4 Efficiency

The efficiency criterion measures the extent to which objectives of a project are cost-effectively achieved. Two aspects form the basis for assessment of efficiency in GIZ central project evaluations, both of which will be evaluated in production efficiency, which describes the transformation of inputs to outputs; and allocation efficiency, which examines the transformation of inputs to outcomes. In addition, aspects of implementation efficiency (structures and processes of project implementation) are analysed and considered in the assessment of production efficiency, since they contribute to maximise outputs.

To perform the analysis of production efficiency, the GIZ evaluation unit has introduced an Excel tool to carry out a 'follow-the-money' analysis (see Annex 9). The tool is intended to show how many resources have been spent on the respective outputs, and to analyse this against progress on the indicators associated to each output. In the case of this project, however, the following challenges to using this tool occurred, which limit the explanatory power of the approach. First, the AV had changed several times since the beginning of the intervention in 2016 with the current AV only joining the project in spring 2017, which contributed to the difficulty of distributing spending among outputs (see Section 4.2 on methodological approach). Second, the indicator system was found not fully suited to conducting the analysis. Some of the output indicators are outside the project's sphere of influence (output indicators A2 and B1) and action area II.b has not been translated into a separate output area, meaning that resources cannot be tracked in relation to a separate output (see Section 3.2.). In order to conduct the assessment, the evaluation team has used the Excel tool to examine how much money has been disbursed or committed so far for each of the three output areas from the project proposal, and has then undertaken a qualitative assessment of progress on the different action areas.

In contrast to production efficiency, allocation efficiency puts inputs in relation to outcomes. The traditional approach for such an analysis is to monetise the added value of outcomes. In the case of this project, however, the evaluation basis for such an approach was limited, since it is not possible to monetise the added value of increased transparency, credibility and effectiveness of the Zambian PFM system with this evaluation's resources. The evaluation team therefore undertook a qualitative assessment of allocation efficiency. In this regard, the evaluation team examined to what extent the outcome has been maximised given the available resources, e.g. by leveraging resources of other organisations.

In order to assess the intervention's production efficiency, the distribution of costs among outputs was discussed with the AV, the finance manager and the intervention's M&E specialist based on the 'follow-the-money' tool described above. The illustration below displays all costs including obligations and total expected co-financing. Output B absorbs the largest share with 41% of all financial resources, which can be explained by the fact that it entails two action areas, in particular the action area II.b with a considerable co-funding share and resource-intensive IT-related consultancies on the implementation of MIS. Excluding the expected total sum of co-financing, 32% of the BMZ resources are used in this output area. Output A absorbs 25% of resources, mainly consisting of human resources, but also of procurements as described below. Output C makes up the smallest share with 15%, because advisors could only be recruited at the end of year one of the intervention. The overhead costs make up 19% of project resources. In the following, the resource use for each action area (hence differentiating between action area II.a and II.b under output B) and the overarching costs are analysed separately, considering possibilities to maximise yield and the deviations between planned and actual costs.

Modulziel	Transparency, credibility and effectiveness of public finances have increased.
Kosten inkl. Obligo	4.278.015,63 €
Ko-Finanzierungen	649.959,00 €
Partnerbeiträge	50.000,00 €
Gesamtkosten	**4.977.974,63 €**

Modulziel Indikatoren	The share of middle-income taxpayers complying with value added tax legislation under the purview of the Medium Taxpayer Offices (MTOs) North and South has risen to 60%.	The share of individual budget heads that differ more than 5% between approved and actual expenditure has diminished by a quarter.	A rule-based procedure for the allocation of funds (e.g. an authorisation procedure based on expenditure in accordance with budget) is applied for one ministerial budget.
Zielerreichung	103%	56%	50%

	Output A	Output B	Output C	
Outputs	The capacity of the ZRA to raise tax compliance of small and medium-sized taxpayers has increased.	The capacity of MoF to prepare the budget in accordance with the future national planning and budgeting legislation has increased.	The ability of the Ministry of Finance to effectively implement the budget in accordance with plans and the budget act has been raised.	**Übergreifende Kosten**
Kosten inkl. Obligo	1.228.044,31 €	1.362.341,37 €	741.534,03 €	946.095,92 €
Ko-Finanzierungen	0,00 €	649.959,00 €	0,00 €	0,00 €
Partnerbeiträge	16.666,67 €	16.666,67 €	16.666,67 €	0,00 €
Gesamtkosten	**1.244.710,98 €**	**2.028.967,04 €**	**758.200,70 €**	**946.095,92 €**
Gesamtkosten in %	**25%**	**41%**	**15%**	**19%**

Illustration 1: Overview of the ‘follow-the-money’ tool with resources distributed among outputs

In action area I, the use of 25% of project resources – more than in any other action area alone – contributed to the achievement of the relevant module indicator before the end of the project term. Main spending areas had been human resources with two international and one national advisor on the one hand, and procurements as well as consultancies on the other. With regard to the latter aspect, around EUR 150,000 was spent on procuring the EQM system for the three largest taxpayer service stations, on contracting consultants and trainers for improving tax audit techniques, and on hiring temporary staff for ZRA to clean-up the taxpayer database. Having already met module indicator M1 and output indicator A2, the focus for the final year lies on supporting implementation of the SME taxation strategy to achieve output indicator A1 and to foster the results beyond the project term (interviews 11, 16).

When looking at the value for money in action area I, most interview partners did not identify any inefficiencies or alternative measures that could have achieved more results (interviews 11, 16, 17, 18). Two relevant concerns were brought up, which could subsequently be justified, however. First, not everyone considered the data clean-up exercise conducted by the project to be efficient. One interview partner stated that it may be a duplication of an activity routinely conducted by ZRA itself (interview 19). Another interview partner regarded it very useful, but stated that it covers all taxpayers, but large taxpayers are not in the scope of this project (interview 20). Confronted with these observations, the project staff explained that the guided clean-up exercise was necessary because of the poor quality of previous activities by ZRA alone and that SME made up a considerable part of the taxpayers in the database. Second, the procurement of the EQM system initially raised questions in terms of efficiency. While it was assessed very positively and it provided a highly visible improvement of service quality (interviews 11, 16, 18, 19, 20), the lack of an output indicator assessing the crucial aspect of improving taxpayer service quality means that an evaluation of cost-effectiveness is only possible to a limited extent. Nevertheless, an analysis of EQM data reveals a considerable reduction of waiting time in the taxpayer offices since introduction of the system, indicating an increase in service quality and making the purchase plausible (ZRA data generated by EQM system, see ZRA 2017).

In action area II.a, much has already been achieved considering that few financial resources were so far spent beyond the human resources of the senior advisor in this area. Given the high technicality of this area, which requires strong advisory skills, the partner and other stakeholders assessed this approach as positive (interviews 22, 36, 37). Remaining financial resources for the final intervention year will be used to deepen the integration of planning and budgeting through ‘cluster’ policy and budget hearings, and support the further roll-out of the OBB to additional MPSAs (interview 12). In light of the crucial role that this action area plays for the overall success of the project, being key to address the critical issue of budget credibility and working towards the central module indicators M2 and M3, it is surprising that the human resource concept only considered one advisor for it compared to three and two in action areas I and III respectively.

In action area II.b, the resources (potentially) available for the MIS implementation could not be used for maximum impact. More precisely, until December 2017, less than 30% of the initially earmarked co-funding sum had been used due to a number of external and internal factors. While the reduction in spending was also a result of the UK government’s reassessment of support to the MIS, which can be seen as an efficient step to prevent the waste of resources when objectives cannot be achieved, it fails to fulfil the maximum principle of achieving the maximum possible results with the given resources. The field study has shown that a number of measures could have contributed to preventing the limitations in project progress, as follows.

- First, the human resources in the form of one advisor being in charge of both technical advisory and management activities were reported to be insufficient to meet the objectives initially outlined in the planning (interviews 24, 30).
- Second, the inadequate reflection of action area II.b in the GIZ results model and the measure’s indicator system may have limited the capacity of for the project’s effective management and monitoring. Several interview partners commented on the overly ambitious initial planning of the project (interviews 5, 13, 30, 35). While the ToC had then been revised twice in 2016 and 2017, these changes were not translated into the GIZ indicator system.
- Third, the lack of an effective steering committee meant that critical issues could not be addressed in a timely manner and the partner was left unclear about what support to expect (interviews 23, 24).
- Fourth, coordination issues at both the technical level (i.e. what is the common strategy to achieve the agreed objectives?) and the political level (what are the modalities of the Delegated Cooperation Agreement?) between GIZ and DFID led to misunderstandings and appear to have contributed to funding delays/reductions.

For the remaining year, the amount of resources still allocated in the framework of the co-funding agreement was still in discussion with DFID at the time of the field study, but in light of the delays and the limited absorption capacity of MNDP, the stated objective was merely to foster the interim achievements (such as enhancing the interim solution of the Monthly Monitoring System) and underlying M&E processes (interviews 8, 13, 15).

In action area III, the project is making efficient use of its resources in light of the delays that have occurred by both GIZ and the partner. Expenses occurred later in this area, because the technical advisors could only be recruited at the end of year one, resulting overall in less spending up to the time of evaluation. Resources other than human resources were mainly spent on consultancies for developing business process mappings and training sessions with the Internal Audit Department. Additionally, a peer-learning visit to the South African Treasury was organised, resulting in the identification of key improvements for cash management with minimal outlay (interviews 14, 21, 37). Comparing this activity with alternative options, such as an exchange visit of relevant MoF staff in South Africa or conducting a series of training, the peer-learning visit was assessed as highly cost-effective, both in terms of the few resources spent (mainly the travel cost of the two visitors) and the results achieved (15 key MoF personnel gaining first-hand experience through peer exchange) (ibid.). Given the current gap between achievements to date and expected output indicators, comparatively more resources have been allocated for achieving the remaining output indicators in this action area until the end of the intervention. In this respect, the training on risk management was purposefully postponed until the risk management unit in MoF is established, which was expected for the first quarter of 2018.

Finally, the overarching costs of 19% are high by GIZ standards, resulting in an administrative burden that could potentially be used to maximise results. Looking into the figures in more detail, the tool reveals that staff costs in particular are high in this regard: 50% of national and 16% of international human resources as well as 59% of local consultant personal are booked as overarching costs. A likely explanation for these comparatively high costs is the additional administrative effort required to manage and implement co-financing from DFID, which was confirmed by the project staff (interviews 3, 15). The evaluation team deems it important to shed light on such costs necessary to leverage co-financing, which needs to be taken into account for future potential co-financing plans.

After analysing aspects of implementation efficiency, the project team provided several examples of following the efficiency principle, but also observed limitations in the way the project navigates the partner structure. In terms of positive examples, it was reported that allowances were not paid for training activities in order to save resources and to avoid setting false incentives (interview 14). It was also stated that the project team attempts to make use of their own work time as efficiently as possible and that all decisions on expenses were prudently taken, though the methods or approaches used could not be specified (interview 11). Funding for GIZ projects is agreed between the two governments involved up to a specified amount but not on a year-to-year basis; therefore, in-year allocation decisions for each specific region/project remain at the discretion of BMZ and GIZ head office. Funds received in 2016 were lower than originally planned for project operations so that some activities had to be postponed. This practice carries the risk of potentially inefficient spending decisions based on the availability of funds per given year. While this circumstance lies outside the sphere of the project's influence and is therefore not considered in the assessment, it should be mentioned as a factor inhibiting efficient resource allocation. As reported, however, it also led to the project leveraging support from AFRITAC South and the World Bank-managed MDTF for piloting OBB and developing the OBB manuals (interview 12).

Regarding limitations, it was reported that the PFM reform coordinator in the MoF has limited knowledge of what the GIZ project is doing in the other action areas, creating a risk for inefficiencies (interview 21). In line with this observation, several other stakeholders characterised the partners as operating in 'silo structures' (interviews 14, 15, 17, 23), which the project did not manage to break on a more strategic level beyond specific activities such as the South African Treasury visit. Furthermore, the project's lack of an effective formal steering structure to coordinate activities, address problems and to ensure subsequent scale-up with the partners is seen as a shortcoming (interviews 15, 21, 30). In particular, in action area II.b it became clear that the insufficient communication on various levels resulted in a situation that made an efficient and effective resource use impossible.

In order to assess the allocation efficiency, the evaluation team first analysed the use of resources in view of the outcome objectives and indicators; then considered alternative options; and last, examined the opportunities for synergies. In terms of the progress towards the outcome indicators, M1 regarding domestic revenue generation has already been met whereas indicators M2 and M3 are expected to be achieved within the project term. Objectives in action area II.b, which are not directly reflected in any of the outcome indicators but play a crucial role for improving transparency and credibility of public finances, are not expected to be met, however. While explanatory factors in part lay outside the sphere of influence, the inability to make full use of the co-funding potential in the latter area needs to be considered a shortcoming regarding allocation efficiency as well, since the resources could not be used to full potential at the outcome level.

Rethinking the human resource concept is an alternative option for the project to increase effectiveness as well as efficiency. Currently, each action area is covered by at least one international advisor, who is embedded into the partner structure almost like an integrated CIM-expert (interview 5). In addition, action area I is currently covered by one more international and one more national advisor, and action area III by one more national advisor. While the close integration of advisors into the partner structures was successful at establishing good working relationships with specific stakeholders within the partner structure, this approach limits possibilities to cover a wider area within the partner structure. Moreover, the critical action areas II.a and II.b appear understaffed in comparison to the other action areas. While the partner's absorption capacity of additional advisors would require careful consideration, particularly in MNDP which still falls short of meeting its own human resource commitments, a relative increase of human resources in these areas could help facilitate results. For the future, the project is considering a more cluster-based approach with the project's advisors covering larger areas within the partner structure, which can be complemented with specific consultancies if needed. In light of GIZ efforts to raise capacities of regional experts in the framework of the Regional Support to Good Financial Governance project in Southern Africa (see Raetzell 2015), regional experts from the GIZ network could provide an additional pool of candidates for both short and long-term assignments.

Regarding synergies with other funding sources, leverages could be achieved with the World Bank-managed PFM reform programme under the MDTF and by taking part in two German PPP projects in action area I. Each of these aspects will be examined separately in the following.

First, the ongoing PFM reform programme managed by World Bank offered considerable room for synergies, which were mostly but not fully exploited. It was reported that initial activities were planned together with the German Development Bank (KfW), which manages the German contribution to the MDTF, (interviews 15, 38). The project documents stated that evaluated project objectives as well as action areas and the MDTF should be clearly linked (e.g. the GIZ project working on five out of ten components of the PFM reform programme). Also, examples for leveraging each other's activities and results were provided. For example, the MDTF financed follow-up training after GIZ technical advice to the Internal Audit Department (interview 25), and consultancies contributing to the achievement of project results were also financed by MDTF (interview 12). Moreover, the project regularly shared experiences in the PFM reform programmes coordination meetings.

Nevertheless, the evaluation team observed that a difference in priorities between World Bank as manager of the PFM reform programme and GIZ objectives emerged in the course of the implementation due to non-aligned partner requests. While the timely roll-out of IFMIS to all MPSAs is a key objective of World Bank-managed activities, which is also reflected in output indicator B1 of this project proposal, the project adapted its strategy to first address underlying issues, because it was noted by MoF that meeting the stated quantitative indicators alone will not bring about the expected results. Therefore, the project started to first analyse and map underlying business processes in key areas of budget execution (interview 14). It supported the reconfiguration of IFMIS by advising the partner in communicating their requirements to better meet specific needs such as OBB functionality (interview 12). This adaptation was successful at better leveraging each other's results and coordination took place in the framework of the steering structure around the PFM reform program; nevertheless, more strategic coordination might have been better (interviews 15, 21, 34). In terms of conception, for example, it is unclear why indicator B1 on rolling out IFMIS was chosen for the project in the first place, if this aspect is covered by the PFM reform programme. Moreover, it was reported that parts of MoF involved in the PFM reform process under the MDTF are not sufficiently informed about the interactions between the MDTF-funded activities and the GIZ project (interview 21).

Second, the involvement in two PPP projects in action area I helped leverage the resources towards increasing the revenue side of public finances (interviews 11, 16. 28). In particular, the successful launch in 2015 of a PPP project ZAMeT, which aims to enable SME to register as tax payers and file returns using mobile phones to simplify tax compliance, is planned to be rolled out further from 2019 with additional external PPP funding.

Overall assessment of efficiency

Overall, the evaluation team observed a mostly efficient allocation of resources to the outputs, but some limitations require mentioning. (1) It is striking that while the two action areas under output B – which (for action area II.a) works towards achieving the module's two core outcome indicators (M2 and M3) and which was planned to absorb (in case of action area II.b) about half of the project's overall resources (including co-financing) – were only covered by one advisor each, while action areas I and III were covered by three and two advisors respectively. Particularly in action area II.b, the limited human resources were reported to have contributed to the failure of the project to meet its objectives. (2) Inefficiencies were detected regarding the co-financing in action area II.b. The project's high overarching costs appear to be partly a consequence of related additional administrative processes. However, there were considerable deviations between the initially planned and the actually spent resources in action area II.b for various reasons, which ultimately undermined the project's ability to maximise results in that area. (3) In light of the good performance of action area I, a shift of unallocated resources for the final term could achieve more results in other areas. Balancing the aforementioned aspects, the evaluation team awards 50 of 70 to the aspect of production efficiency.

In terms of allocation efficiency, the evaluation team observed examples of synergies, which could be further exploited in part through better coordination within the donor landscape. In particular, project results could be leveraged with activities by the World Bank-managed PFM reform programme under the MDTF, but a closer coordination beyond the exchange in coordination meetings under the MDTF might have achieved even more. In conclusion, the evaluation team awards 20 of 30 points to the aspect of allocation efficiency.

The overall score for the assessment criterion efficiency adds up to 70 of 100 points: rather successful.

Criterion	Assessment dimension	Score
Efficiency	The project’s use of resources is appropriate with regard to the outputs achieved [Production efficiency: Resources/Services in accordance with the BMZ]	*50 / 70 points*
	The project’s use of resources is appropriate with regard to achieving the TC-measures’ goal (outcome). [Allocation efficiency: Resources/Services in accordance with the BMZ]	*20 / 30 points*
Overall rating efficiency		*70 of 100 points*

5.5 Sustainability

The assessment of sustainability at this point is limited by the fact that this is a midtermevaluation. Therefore, the focus of the analysis lies on the prospect of sustainability. To assess this, the evaluation team took into account the following three aspects. First, it examined the extent to which results are currently anchored in partner structures. Closely intertwined with this is the second aspect of sustainability, a forecast of the durability of results. To analyse this aspect, the evaluation team differentiated between the extent to which results are owned at the operational level and the way and extent to which the policy level influences the sustainability. Third, the evaluation team analysed the extent to which project results are balanced between economic, social and ecological dimensions of sustainability.

In terms of the first assessment dimension, the project provides successful examples of anchoring results in the partner structures in all action areas. In each partner institution, it was reported that the project is consolidating achievements in the partners’ tools and structures, which was facilitated by the close embeddedness of the project’s advisors into the partner organisations (interviews 16, 21, 22, 24). These efforts will be described separately for each of the action areas.

In action area I, the project anchored its results in the SME taxation strategy, in the audit training manual and in the physical infrastructure of the EQM system. Strategic advice was institutionalised in the SME taxation strategy, which forms the basis for the future work of ZRA to increase SME taxpayer compliance (interviews 17, 18, 20). Also, the support of tax audit functions through training and accompanying manuals as well as the development of sector notes for specific business sectors provided a basis for ZRA tax inspectors to do their work more effectively. The results of the training were used to update ZRA’s internal training manuals, making their long-term application likely (interviews 18, 20). The EQM system, which was procured for three taxpayer offices through the project, is expected to help ZRA to effectively manage their service points for many years beyond the project. Partner staff was trained on its use and the analysis of data generated by the system, hence making its long-term use likely (interview 19). Finally, an important observation was made by the project in the beginning, which was considered crucial for both effectiveness and sustainable impact of the project. It was found that the taxpayer database as the foundation for effective tax collection was inaccurate, providing false information on some businesses, containing various organisations no longer operating and omitting a large part of active taxpayers. Without an accurate taxpayer database, it was reported that it would not be possible to institutionalise higher quality tax audits (interview 11). In consequence, the project financed the recruitment of 10 short-term staff for ZRA with the goal of cleaning and updating the database. While this is a one-time activity that needs to be followed up and continuously improved to have an impact (interview 19), it can be seen as a plausible first step to showcase the potential of a solid taxpayer database for more effective tax revenue generation in the future.

In action area II.a, the senior advisor made sure to consolidate achievements of technical advisory services through legislative or procedural measures. On the legal side, the project provided expert advice to the partner on various different reform bills including the draft excess expenditure (limitation and conditions) bill and the long-awaited budget and planning bill (interviews 12, 22). This is expected to result in the institutionalisation of critical reform initiatives in the expected law. While there is always the risk of a discrepancy between formalised law and its application, these legal measures are nevertheless a necessary starting and reference point for reform towards more transparent and credible budget planning in the long run (interview 12). With regard to procedural measures to anchor project activities in the partner structures, the project focused on institutionalising suitable cooperation formats within the government to foster integrated budgeting processes involving all relevant stakeholders. In this respect, cluster hearings with groups of MPSAs that are jointly chaired by MoF with MNDP were introduced with the objective to institutionalise aligned planning and budgeting processes. The budget's challenge function was strengthened as a supporting tool. This included the expected establishment of a 21-person budget formulation unit from Q2 2018 and the ongoing drafting of a budget challenge manual for budget analysts in this new unit office to help scrutinise budget proposals coming from the spending agencies (interviews 3, 22). Moreover, the supplementary budget and excess expenditure process was re-engineered to deepen the cabinet's understanding of the consequences of in-year political decisions, and pre-cabinet meetings were initiated to discuss budgetary matters separately before full cabinet meetings. These initiatives sought to institutionalise the political ownership of the budget (interview 12). Finally, the piloting of OBB in three ministries and of a warrant system to replace cash rationing are measures that aim to foster more effective, efficient and transparent approaches to budget planning in the future. Again, the advisory work was accompanied by the formulation of user manuals and the provision of training by leveraging support from AFRITAC South, which was reported to be useful (ibid.). Particularly in light of the challenging political situation, there are risks that even adopted laws are not followed as envisioned and that the established procedures are not applied in the long run (ibid.). Nevertheless, the operational partner generally assesses the sustainability of results as high (interview 22).

In action area II.b, the project was successful at providing the newly established MNDP with tools and methods to operationalise strategic plans, though it fell short of initial expectations to implement MIS. With support of the advisor, MNDP managed to break down the long-term Vision 2030 into the National Performance Framework. Sectoral and ministerial strategies, key performance indicators for the strategies' implementation were developed, and data catalogues with rules and specifications were created (interviews 13, 23). Given the slow progress in terms of developing the MIS as the targeted IT system, however, MNDP decided to rely on the Monthly Monitoring System, which was developed with support of the previous project phase as a predecessor to the MIS (interview 24). The goal is to enhance the Monthly Monitoring System as much as possible until the end of 2018. Particularly in light of the support from this project running out at the end of 2018, there is a risk of the initiated processes not being able to continue. This risk is exacerbated by the fact that human and organisational resources are insufficient at present to absorb advisory results (interviews 13, 23). The lack of a project management unit, which was approved by the government but never installed, is a case in point. While interview partners within the government partly see the lack of automation as the main challenge and therefore ask for IT solutions (interview 24), the project and external stakeholders consider the limited capacity and resources to be more critical (interviews 13, 30, 35).

Finally, in action area III, results are mostly anchored in the partner structure, for example through training manuals and in a peer-learning process initiated with the South African Treasury. The Internal Audit Department capacity is sustainably enhanced by incorporating a quality assurance system into the daily work and by fostering both internal peer review as well as an external review after five years (interviews 25, 26). Internal audit manuals are gradually being disseminated to internal auditors in the MPSAs through training. The work on risk management has not started, since the relevant risk management unit has not yet been established. Yet, by planning to anchor the training work in the future risk management unit, the project seeks to ensure the institutionalisation of risk management beyond the project term (interviews 14, 26). The peer-learning visit organised with the South African Treasury to share and gain experiences can be assessed very positively in this regard, as it created ownership of the reform process and allowed the partner define its own priorities in improving the functionality of the TSA (interviews 14, 21, 37).

In addition to the good prerequisites for long-term success, there are many examples of partner structures owning the tools and approaches developed with the project. In the following, these will be discussed for each action area separately.

Action area I:

- ZRA was reported to have a sophisticated project management unit that coordinates all international projects and follows up on common milestones in a proactive way (interviews 11, 18).
- It is planned to continue the training on SME tax audit without project support, which highlights the interest ZRA has to follow up on the results of the project (interviews 19, 27).

Action area II.a:

- The project is attempting to increase political ownership of the budget by supporting the budget office in re-designing the supplementary budget and excess expenditure process and briefing cabinet before their meetings on budgetary implications of in-year changes (interviews 12, 22).

Action area II.b:

- The cluster hearings, chaired jointly by the Permanent Secretaries in the MNDP and the Secretary to the Treasury and Permanent Secretary (Budget and Economic Affairs) in the MoF, to challenge the budget proposals of MPSAs institutionalise common budget planning procedures (interviews 23, 30).

Action area III:

- The peer-learning process with the Republic of South Africa that was initiated with a first National Treasury visit in 2017 and is planned to continue helped put the Accountant General’s Office into the position of identifying reform challenges and learning from a regional government’s experience (interviews 14, 21).

Nevertheless, the forecast of durability is limited by inhibiting factors on the operational and political levels. These challenges are discussed separately, as follows.

On the operational level, crucial factors inhibiting the durability of results related to both the partner and the project. First, the available resources of the government are often insufficient to implement reform processes (interviews 23, 25, 35). The human resource situation is particularly difficult in MNDP, which is still reported to be understaffed two years after its establishment (interviews 13, 24, 30). In MoF high staff turnover was observed. Here, the six internal auditors interviewed were currently in acting positions, and the risk management unit, which was expected in 2017, was not yet in place at the time of the case study (14, 25). Second, the long-standing closeness of advisors to the partner, despite its merits, creates a personal dependency. Hence, there is a risk that the advisors’ role will be hard to fill once the project is over (interview

36). This question was also brought up in the debriefing session that the close embeddedness of the advisors has many benefits although there is a in terms of sustainability. Third, several interview partners stated that the Zambian government is requesting IT systems that the structures are not ready for, which results in a high risk of losing gains after the end of the project. This is particularly the case in action area II.b with the delays in MIS implementation. To a lesser extent, action areas II.a and III are also affected in view of the IFMIS, which after years of international support is still not effectively functioning (interviews 12, 14, 33).

On the political level, risks are perceived as high for durable and resilient results in light of the overall difficult PFM framework (interviews 28, 33, 36). This can be explained by the following factors. First, academic research describes the political situation as worsening, with increasing corruption and pressure on civil society, media and the opposition (Beardsworth 2017). Second, the public debt remains high (IMF 2017). While the Zambian government has again taken up talks with IMF, there are no signs that it plans to reverse its overall debt strategy given the announcement of further loans and the continued unclarity over the total debt (interviews 34, 38). Third, the political will for a more profound reform process is perceived to be limited. As was elaborated in Section 5.1 on Relevance, there is certainly a formal commitment to political reform, and various technical-level actors are also pushing for reform. At the same time, a gap between written law and actual practice is sometimes observed thus possibly limiting long-term sustainability (interviews 8, 28, 29, 33). Last, the weight of Western cooperation partners is reported to be decreasing. More concretely, it was reported that a general shift of Zambia’s bilateral cooperation is taking place towards emerging players, while Western cooperation partners and their conditionality approach are gradually playing less of a role (interviews 8, 28, 36).

Analysing the political dimension further, the influence the project has on these challenges is limited. Neither the domestic political debates nor the geopolitical trends are developments that the project with its technical advisory services can address. Although the project cannot be expected to offset the aforementioned risks, the question came up if it could offer more of a normative counterbalance, for example through civil society or other external accountability actors, which forms an additional component of the BMZ Good Financial Governance Strategy (BMZ 2014; interviews 28, 33, 36). These considerations have also been brought forward in the concept note for the follow-up measure (see GIZ 2017c). At the planning stage, this technical focus was considered an appropriate follow-up to the predecessor phase (see Section 5.1 Relevance). Yet, in light of the increasingly challenging political situation, the mere technical focus of the project may not be enough to sustainably address PFM reform (interviews 33, 36).

Finally, with regard to the three dimensions of sustainability (economic, social and ecological), the project successfully works towards economic sustainability, but is only partly addressing the other two. While the limited role of social and particularly ecological aspects in a good financial governance project is obvious, there are some areas of interactions. The interactions regarding each dimension are discussed below.

Most prominently, the project is fostering economic sustainability, since the it works at the governance level with central government institutions on PFM. The economic dimension is clearly addressed by supporting the government in increasing domestic revenue generation and in bringing budget execution more in line with national development plans and the planned budget, hence making public finances more sustainable as stated in the overall programme objective of the DC-measure (see GIZ 2016). This aspect is reflected in the expected results in all action areas (see Section 5.2 on Effectiveness).

Intended positive results regarding ecological and social aspects of sustainability could not be detected in the field mission given the long causal chain between the project and the ultimate target group. Several interview partners pointed out, however, that the project’s results of more effective and credible PFM capacities by the central government institutions can ultimately lead to more effective and efficient public service provision in the spheres of social and ecological policy (interviews 15, 16, 24).

Relating to certain cross-cutting issues, it was reported that the project had in the previous phase developed guidelines for climate-proof and gender-sensitive budget proposals (interview 12), but neither was followed up to the stage of implementation in the current phase. Ultimately, no negative interactions between the different sustainability dimensions were observed.

Overall assessment of sustainability

With regard to the first assessment dimension of sustainability, the evaluation team found that results are well anchored in the partner organisations. In each action area, the project is successfully consolidating achievements in the partners' tools, methods and structures. Only in action area II.b, the project was not able to sufficiently institutionalise results in the partner structure due to its limited absorption capacity and the slow project progress. Therefore, the evaluation team awards 37 of 40 points to the first dimension.

Considering further factors influencing durability, however, the prospect for sustainability is lower. The larger political environment, which includes the continuation of intransparent financial management, a reduction of freedom for civil society, and a decreasing weight of Western-style development cooperation approach of good governance, indicates that the political will for sustainable PFM reform is limited. Also, at the operational level, the potential for sustainability is reduced by the partners' limited resources and reliance on the project's embedded advisors on the one hand, and by the technical focus of the project that neglects the political dimension on the other hand. Nevertheless, the partner shows high ownership of the approaches, methods and tools developed together in the project. Balancing these aspects and considering the limited influence the project has on the political factors, the evaluation team awards 20 of 30 points to this dimension.

Finally, the evaluation team observed strong results regarding economic sustainability and an indirect contribution to social and environmental aspects of sustainability. The project's focus on economic sustainability is sensible given its overall objective. Social and environmental aspects (e.g. climate-proof and gender-sensitive budget proposals) appear to be a low priority on the partner's agenda, which may explain why certain related activities from the previous phase were not followed up. Ultimately, the evaluation team awarded 25 of 30 points to the aspect of balancing the sustainability dimensions.

The overall score for the assessment criterion sustainability adds up to 82 out of 100 points: successful.

Criterion	Assessment dimension	Score
Sustainability	Prerequisite for ensuring the long-term success of the project: Results are anchored in (partner) structures	*37/40 points*
	Forecast of durability: Results of the project are permanent, stable and long-term resilient	*20/30 points*
	Are the results of the project ecologically, socially and economically balanced?	*25/30 points*
Overall rating sustainability		*82 of 100 points*

5.6 Long-term results of predecessor

The predecessor intervention, Good Financial Governance in Zambia (PN 2011.2114.4), sought to increase transparency, accountability and development orientation of public incomes and expenditures. In spite of a similar overall structure, this predecessor fell short of addressing the core problem regarding budget credibility according to the project evaluation conducted in 2015 (see GIZ 2015a). In this context, the current intervention subject to this evaluation was designed with the objective of achieving conformity between planned and executed public expenditures and creating the necessary legal framework for this. Furthermore, DFID's co-financing in the area of MIS implementation (action area II.b) that continues today was agreed upon in the framework of this predecessor in 2014 (see GIZ 2014a). Due to the different focus and the existence of a previous project evaluation of the predecessor, it was agreed with the AV that a broader investigation of the predecessor would not be necessary. The predecessor was merely planned to be considered in terms of the continuation of activities in action area II.b.

Nevertheless, the field study has revealed that many of the achievements in the current project phase cannot be viewed in isolation from the previous phase. For example, the Internal Audit Department had been supported by the predecessor project. Audit manuals were developed as a foundation for current activities of introducing quality assurance in line with international standards (interviews 14, 25, 26). In action area II.b, however, an analysis of long-term impacts could not be conducted as planned. Here, the work on MIS implementation was initiated in 2015, but during the inception phase it was realised that the initial plans needed to be changed, which resulted in the development of two new ToCs in 2016 and 2017 respectively. Therefore, it was not possible to evaluate the results under the previous phase in this area.

6 Overall rating

Considering all criteria together, the intervention is rather successful at meeting its objectives. On the technical level, the project team is providing highly relevant and effective advisory services to the partner. Due to the challenging political environment and the (so far) limited ways in which the project design has adapted to this, however, the project is not realising its full potential. In the following, each of the five evaluation criteria will be briefly discussed separately.

The **relevance** of the project in general is assessed as high with minor limitations. The project reflects and builds upon key strategic documents from the Zambian government (7th NDP, NPF, Medium-term Growth and Stabilisation Plan), the BMZ (Good Financial Governance Strategy, Country Strategy Zambia, Digital Agenda) as well as the Sustainable Development Goals (SDG 16.6). While there may be a discrepancy between the Zambian government's formal commitment and the actual political will for reform, the continuation of non-transparent procurement and budget processes showcase the need of a project on strengthening good financial governance, to which the project with its technical approach makes an important contribution. When looking at the project in detail though, its design is only partly adapted to the chosen goal in light of the challenges that exist. Even though the project's technical orientation was assessed to be sensible at the time of conception, this approach does not sufficiently address normative and political aspects of good financial governance. In light of the political situation, however, these appear to be of critical importance for lasting success.

With regards to its **effectiveness**, the project is assessed to be successful. This is limited, however, by the partly inappropriate indicator system and the handling of potential unintended results. The stated outcome indicators and the majority of output indicators are expected to be achieved by the end of the project. Only in action area II.b, the results fell short of the initial expectations. But while all outputs make a meaningful contribution to the module objective, not all indicators reflect the actual project progress, and action area II.b had not been appropriately conceptualised, which may have contributed to challenges in implementation. Moreover, the occurrence of unintended positive or negative results was only addressed to a certain extent. Even though no evidence was found of negative unintended results having materialised, crucial risks were identified in the initial results model and reflected in some of the interviews with project staff that may turn into unintended results. While plans exist to address this more systematically in the next phase, the project at the time of evaluation does not follow a systematic approach to deal with them.

With regard to the **impact** criterion, superordinate long-term results in terms of programme-level indicators and impact on the general population can only be expected to be achieved to limited extent. External macroeconomic factors and the political situation play a large role in explaining these shortcomings. Looking at the results on only the technical dimension, one can observe a plausible contribution to expected long-term results in terms of increased domestic revenue generation and better alignment of budget planning and execution processes. Yet this technical focus also limited the possibilities of the project to address the reform on a more political level. Though no negative results were observed, a systematic approach to monitor and address unintended results is not in place.

The **efficiency** of the intervention is assessed as rather successful both in terms of production efficiency and allocation efficiency. Regarding the former, the following issues arose despite an overall sensible allocation of resources. First, given the crucial role that action areas II.a and II.b play in terms of the module objective and for financial resources, it is surprising that there is only one advisor in each of these areas. Action areas I and III, in contrast, work with three and two advisors respectively. Particularly in action area II.b, the limited human resources appear to have contributed to the failure of the project to meet its objectives. Second, questions arose regarding the co-financing in action area II.b. On the one hand, the project's high overarching costs appear to be partly a consequence of the additional administrative effort required to manage and implement co-financing from DFID. On the other hand, there were considerable deviations between the initially planned and the actually spent resources in action area II.b, which ultimately undermined the project's ability to maximise results in that area. Last, in light of the good performance of action area I, a shift of available free resources for the remaining term towards other areas could have been considered to achieve more results elsewhere. In terms of allocation efficiency, the evaluation team observed examples of synergies, which could, however, be further exploited in part through better coordination within the donor landscape. In particular, while project results could be leveraged with activities of the PFM reform programme, a closer coordination beyond the exchange in coordination meetings under the MDTF may have been able to achieve even more.

Considering the criterion of **sustainability**, the project is assessed to be successful. The prerequisites for ensuring long-term success in the form of tools, concepts and approaches being anchored in the partner structure are largely fulfilled. The project design is focused on finding technical solutions and is assessed to be successful at that. Given the important political dimension of the challenges in the PFM framework, however, a purely technical approach may not be enough. Here, the evaluation team believes that the political challenges need to be addressed more in relation to the political and normative dimensions of good financial governance; for example, by involving external accountability actors through the GIZ civil society programme. Moreover, the close embeddedness of the project's experienced technical advisors in the partner structure, in spite of its merits, creates a risk of dependency and of leaving a gap, once the project is over.

Criterion	Score	Rating
Relevance	*84 points*	*Successful*
Effectiveness	*82 points*	*Successful*
Impact	*72 points*	*Rather successful*
Efficiency	*70 points*	*Rather successful*
Sustainability	*82 points*	*Successful*
Overall score and rating for all criteria	*390 / 5 = 78 points*	*Rather successful*

100-point scale	6-level scale (rating)
92–100	Level 1 = very successful
81–91	Level 2 = successful
67–80	Level 3 = rather successful
50–66	Level 4 = rather unsatisfactory
30–49	Level 5 = unsatisfactory
0–29	Level 6 = very unsatisfactory

7 Key recommendations

Considering that there is less than a year remaining in the project and that some recommendations require major changes of the project design, many of them may not be implemented within the frame of the current project phase. Nevertheless, they can be considered for the planning of the follow-up project. The following recommendations are hence divided into three sections. (1),A few changes to the indicator system are suggested. These may partly be implemented directly and partly may be useful to consider for the follow-up phase. Based on the discussion that took place at the debriefing workshop with the appraisal mission for the next phase, these might still be further specified or adapted by the GIZ planning specialists. (2) A number of short-term recommendations are given, which could already be addressed in the remaining time of the current phase. (3) Medium-term recommendations are given that are likely to only take effect in a follow-up measure. Nevertheless, the latter two can be seen on a continuum, meaning that some short-term recommendations might be too ambitious to implement now and that preparations for some medium-term recommendations can be already initiated in the current phase.

Recommended changes to the indicator system

Recommendation one: Reformulate outcome indicator M2

As discussed in the inception report, indicator M2 was found not to be SMART because it was phrased ambiguously. Therefore, the evaluation team recommended reformulating indicator M2 as follows. The baseline and target can remain unchanged.

> *The proportion of individual budget heads that differ more than 5% between the originally approved budget by Parliament and actual expenditure incurred by MPSAs has diminished by a quarter.*

Recommendation two: Revise output indicator B1

The evaluation has revealed that output indicator B1 is inappropriate to reflect project progress. The roll-out of the IFMIS budget module to the MPSAs is an activity pursued by the government with support of the World Bank-managed PFM reform programme, which is beyond the scope of this project. The advisors in this intervention, therefore, adapted their strategy towards helping the MoF in identifying shortcomings in the IFMIS module and to advise the partner on communicating their requirements for the reconfiguration of IFMIS. This adaptation has, however, not been reflected in the indicator system. Therefore, the evaluation team suggests using the following replacement indicator, which can better reflect the achievements of activities in this area relating to budget preparation and could already be applied to the current project.

> *Supplementary expenditures to the 2018 budget are incurred in accordance with the Constitution.*

Recommendation three: Revise output indicator A2

The field study has shown that the fulfilment of indicator A2 is currently outside the project's sphere of influence as the quantity of tax audits depends on internal ZRA resources. Moreover, its achievement does not reflect the qualitative improvements resulting from project activities. Therefore, the evaluation team recommends taking a more qualitative approach. While the measurement against alternative indicators may no longer be possible at this stage, the following indicator replacement could be considered for any future measures. Given the lack of baseline data, this indicator is recommended only to be used for a potential follow-up measure. This indicator may be specified or adapted by the GIZ planning unit, if necessary.

> *The number of pre-defined issues reported in an internal review process of tax audits has decreased by x% compared to a control group.*

Recommendation four: Develop separate output and outcome indicators for the future in action area II.b once the strategy is aligned with DFID

The evaluation revealed that action area II.b requires a results model including outputs and outcomes on its own, as it seeks to enhance capacities of an independent partner institution, the newly established MNDP. Hence, it is recommended to define an additional output area including appropriate indicators for this action area. First, it is recommended to add an additional output objective B.b. Second, it is recommended to reformulate and elevate the current output indicator B3 into an outcome indicator. And third, the outcome indicator should be matched by a set of appropriate output indicators on the level below.

Suggested output objective B.b:

> *The capacity of MNDP to inform budget preparation through M&E is increased*.

Suggested outcome indicator M2b:

> *The ministries piloting the Management Information System (MIS) as well as MoF and Cabinet Office use harmonised data for budget preparation purposes as well as for national reporting.*

However, as was explained in the sections on effectiveness and efficiency, activities were delayed in this area and they are not expected to be continued. More important, the strategy and expectations from DFID (as the co-funding partner in this area) and GIZ are currently not aligned, making a suitable (re)formulation of indicators at this point impossible. Once aligned, and if activities are planned to continue beyond this project phase, the output and outcome indicator recommended above can be used as a starting point for a separate results logic.

Strategic short-term recommendations

Recommendation one: Consider ways to improve coordination with the government through an effective steering structure

The evaluation has shown that particularly in action area II.b the lack of an effective steering structure resulted in challenges not being addressed on time and processes being delayed. Therefore, coordination with government structures should be improved. A key step to improve this coordination with government structures is the institutionalisation of a steering committee to better manage expectations both ways and to address critical issues on time. One option is to install an effective PFM reform steering structure that covers the entire government-owned PFM reform progress including, but not limited to this project. In action area II.b it could, for example, be investigated if and how the embassy could help facilitate the formation of a stronger steering structure with MNDP.

Recommendation two: Consider ways to achieve even more synergies with the other key international partners and projects

The evaluation team has shown that synergies with other international partners are not yet exploited to the fullest. With regard to cooperation with DFID in action area II.b, it is recommended to first address the legal challenges at a higher level and then agree on a realistic plan for the remaining time of the project. With regard to the World Bank-managed PFM reform programme under the MDTF, it was found that the project has adapted to leverage results from the PFM reform programme. Yet, the adaptation was not conducted in a fully coordinated way. Closer coordination at the planning stage (e.g. on questions such as: What are the requirements for reconfiguring IFMIS and how can strategies be aligned in this regard?) could therefore contribute to ensuring better alignment and ultimately more impact in the future.

Recommendation three: Consider ways to manage risks and avoid unintended results more systematically

The evaluation has shown that while critical risks such as the political abuse of ZRA are reflected internally by the project team, no systematic approach exists to deal with them. Similarly, potential unintended results were reflected in meetings, but not specifically monitored against or addressed more strategically. Therefore, it is recommended to address risks and unintended results more systematically by, for example, an internal risk management, safeguard mechanisms or exit strategies. This approach is aligned with the concept note for the follow-up measure with which the project team is planning to address existing challenges.

Strategic medium-term recommendations

Recommendation one: The definition of the target group should be reassessed

The evaluation has shown that there is a substantial gap to bridge between advisory services at the level of partner institutions and results at the level of the target group as currently defined: the whole population of Zambia. In addition, the analysis showed that the project's monitoring does not capture changes at the level of the target population. The evaluation therefore recommends re-examining the definition of the target group. In this regard, the project proposal could still make reference to the whole population, but it should also be made clear that the immediate target group, at the level of which changes can be observed, is not the whole population. If the project proposal makes reference to the whole population, it should include indicators that capture changes at that level.

Recommendation two: Consider cluster-based advisory approach to break the silo structure within the government

The evaluation team has shown that limited coordination within the government is an obstacle to effective public financial management. To avoid replicating the reported 'silo structure' of the government in the project, it is suggested to locate advisors less at specific positions within the partner government, but rather engage them on thematical bases that cover different counterparts on the government side. Thereby, cooperation within the government may be enhanced and sustainability risks can be reduced. This approach is aligned with the concept note for the follow-up measure with which the project team is planning to address existing challenges.

Recommendation three: Consider ways to strengthen transparency and accountability within the PFM system, for example by involving external control actors or through partnerships

The evaluation has shown that the political challenges are a key risk to achieving lasting change of the PFM system. While the project's technical advisory service is largely successful, its impact is limited in view of the existing challenges. It is therefore recommended to consider how the normative and political dimensions of good financial governance can be addressed in addition to the technical dimension. In order to do so, it appears sensible to involve external control actors to the PFM system. Potential partners could be civil society organisations, the anti-corruption agency or the Auditor General, each of which could contribute to creating pressure on the government to increase transparency and accountability of the PFM system. Given the supply-side approach of the project and the risk that direct engagement with these actors could pose to the relationship with key technical partners, such external actors may be best engaged indirectly through partnerships such as with the GIZ civil society programme. This approach is aligned with the concept note for the follow-up measure with which the project team is planning to address existing challenges.

Image 1: Official launch of the National Performance Framework attended by Dr Frank Hofmann, Head of Cooperation at the German Embassy

Image 2: Sign in front of the ZRA headquarters

Image 3: Publications made within the framework of the project

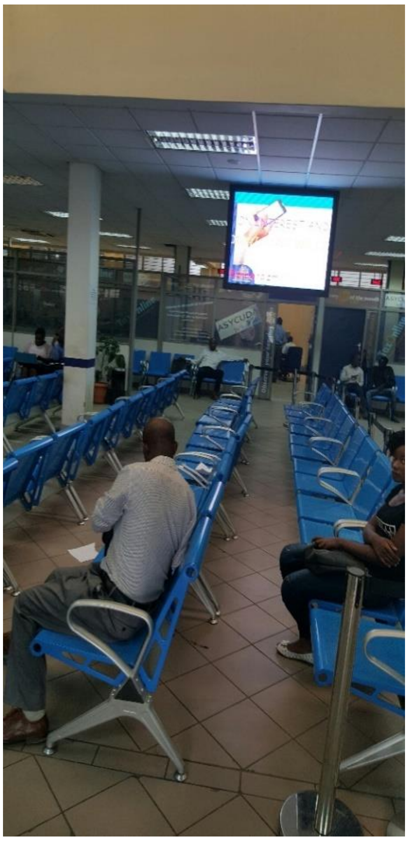


Image 4: Waiting area at the ZRA taxpayer service office with EQM system

Annex

Annex 1: Evaluation matrix

Relevance								
Evaluation dimension	**Analysis questions**	**Evaluation indicators / descriptors**	**Available data sources**	**Other planned data collection projects**	**Evaluation strategy (evaluation design, method, procedure)**	**Expected evidence strength (narrative)**	**assessment by evaluation team**	**comments by evaluation team**
The project fits into the relevant strategic reference frameworks.	1. Which framework conditions or guidelines exist for the project?	1. The intervention's proposal and progress reports refer to overarching principles / frameworks of development cooperation. A) SDGs b.) BMZ country strategy Zambia	Intervention's proposal and annual progress reports	/	Document analysis	Fair	With regards to the international level, the intervention is aligned with the UN framework conditions, most importantly manifested in SDG 16.6. (“develop effective, accountable and transparent institutions at all lev-els“). The intervention´s module objective is making clear reference to this objective (“transparency, credibil-ity and effectiveness of public finances have increased”) and all outputs seek to contribute to one or sever-al aspects stated in this objective (action area I working towards enhanced credibility and effectiveness through domestic resource mobilization; action areas II.a and III towards budget credibility, effectiveness and to a lesser extent transparency through better budget planning, improved coordination and enhanced processes; and action area II.b. towards budget credibility and transparency by supporting the linking of national plans to the budgeting process). The intervention thereby seeks to contribute to enhancing eco-nomic sustainability of Zambia`s PFM system. Analysing the application of the LNOB principle in conceptu-alizing the intervention, the specific needs of parts of the population including women or marginalized groups were partly prioritized in the conception of the intervention. While the development of gender-sensitive budgeting guidelines was foreseen in planning the project (interview 12, 15, 21) and considera-tions of potential negative impacts on specific groups were reported (interviews 11, 14), a systematic analy-sis of risks regarding the LNOB principle and the subsequent translation into mitigation strategies or other operational steps had not taken place. This appears to be a consequence of the evaluation of the previous phase (see GIZ 2015a, interview 6), which recommended a shift away from the development orientation towards more technical aspects of enhancing budget credibility, and reflects the perceived low priorization of these aspects by the partner (interviews 15, 37). In terms of the embedding into the German government’s strategic framework, the intervention addresses core needs identified in the BMZ strategic documents such as the BMZ Strategy on Good Financial Gov-ernance, the BMZ country strategy for Zambia and the BMZ Digital Agenda. The Strategy on Good Finan-cial Governance is reflected in both planning (see the project proposal, GIZ 2015a, p.16) and implementa-tion of the intervention, since the project addresses both the budget planning and execution process as well as domestic resource mobilization. In particular, it focuses on technical aspects of the PFM framework through procedural and legal advice, while it puts less emphasis on the normative and political dimensions of the strategy. This approach is in line with the BMZ strategy´s	
		2. Relevant framework conditions and risks for the program cited by interview partners are refelcted in the intervention's proposal and / or in annual progress reports.	Intervention's proposal and annual progress reports	Interviews with partners and external stakeholders	Comparison between framework conditions and risks described in project documentation and by interview partners	Fair		

	2. To what extent does the project contribute to the implementation of the underlying strategies?	1. The intervention's proposal and porogress reports build upon relevant national and international strategies: a) National Development Plans b) Economic Stabilization and Growth Programm - Zambia Plus c) BMZ Strategy on Good Financial Governance	Intervention's proposal and annual progress reports, NDP, ESGP, BMZ strategy on GFG	Interviews with intervention staff and partners	Contribution analysis, Comparison between partner strategy and results model, document analysis	Fair	approach to select an appropriate set of measures depending on the context (BMZ 2014, p.10). This strategic focus also reflects the partner gov-ernment`s request for technical advice at the time of the intervention's planning and takes into account the evaluation report of the previous measure, which recommended to focus more on addressing the technical aspect of budget credibility (see GIZ 2015a, p.11). The BMZ country strategy did not yet exist at the time of conceptionalizing the intervention, nevertheless it matches well with activities and objectives of the intervention (strengthening accountability and good gov-ernance) (see BMZ 2016; interview 8). In addition, the intervention has been applying relevant aspects of the BMZ´s newly adopted Digital Agenda in particular goals 1.1. on enhancing IT-infrastructure and 1.4. on modernizing public institutions by integrating digital solutions such as an electronic queue-management system (EQM) to improve taxpayer services in action area I or the support of IT-solutions as in the Integrat-ed Financial Management Information System (IFMIS) for financial planning and the planned MIS for M&E in action areas II and III (see BMZ 2017, p.12-13). The intervention is also in line with the partner government´s strategic documents including the PFM reform strategy, the medium- and long-term development strategies and the relevant government programs and laws. Firstly, the PFM reform strategy, which was developed in the framework of a Worldbank program, forms the basis for the government´s efforts to increase efficiency and accountability of public resources. The project design of this intervention picked up five of the ten components of the PFM reform program (integrated planning and budgeting, domestic revenues, IFMIS and cash management, enhanced internal audit and control, Monitoring & Evaluation) (see GRZ 2013, p.6) and works closely with the PFM reform coordinator under the Accountant General´s office in the MoF (interviews 14, 21). Secondly, both the long-term (Vision 2030) and the medium-term (7th NDP) key strategic documents of the government are an-chored in the activities that the intervention is pursuing. In action area II.b, the intervention is supporting MNDP in breaking down the strategic plans into key performance indicators and data catalogues in line with the national and sector performance frameworks in order to ultimately make the budget planning process more transparent and measurable (Interviews 13, 24). Thirdly, the Economic Stabilization and Growth Pro-gram is a medium-term action plan with the goal of "restoring fiscal fitness for sustained inclusive growth and development", which is reflected in crucial activities that the intervention is supporting its partner in (see MoF 2017a). For example, the project is addressing the issue of budget credibility, i.e. reducing deviations of actual budget from parliamentary appropriations (ibid., p.3). Moreover, the project is supporting domestic resource mobilization by facilitating an increase in taxpayer compliance through automation and enhanced taxpayer services (see GRZ 2017, p.6), which the intervention is working on in action area I. Also, the pro-gram seeks "raising the levels of accountability and transparency in the allocation and use of public financ-es" (ibid. p.8), which are key points reflected in the module and programme objectives of the intervention. Lastly, the project is building its activities in action areas II and III on the expected National Planning and Budgeting Bill, which is seen as a crucial reference point for the PFM reform process (interviews 3, 29, 37).	
	3. How was the country´s implementation and accountability for the Agenda 2030 set up and what support needs were identified?	1. description of Zambia´s strategy to implement Agenda 2030 ((Y/N - to be specified if yes) 2. description of relevant support		Interviews with intervention staff and partners	analysis	fair	7th National Development Plan is based on SDGs (Interview 10, 23). The project is supporting the government in breaking down the the goals into KPIs through the National Performance Framework (area 2b, Interview 13, 24).	

		needs brought up to the intervention (Y/N - to be specified if yes)						
	4. Which SDGs does the project contribute to? To what extent is the contribution of the intervention to the national/global SDGs reflected in the ToC?	1. The intervention's contribution to one or more SDGs can plausibly be explained. 2. The intervention's contribution to one or more SDGs is reflected in the results models.	Intervention's proposal and annual progress reports, results models	Interviews with partners and intervention staff	Contribution analysis	Evaluators' assessment, possible subjectivity bias mitigated by data, researcher and method triangulation	There is a close link between SDG 16.6. (develop effective, accountable and transparent institutions at all levels) and the module objective (transparency, credibility and effectiveness of public finances have increased). All outputs seek to contribute to one or several of the aspects stated in the module objective.	
	5. Cross-sectoral change strategies, etc. Where has work been carried out on a supra-sectoral basis and where have such approaches been used to reinforce results/avoid negative results?	1. Cross-cutting issues are appropriately taken into account in the planning of the intervention a.) Gender b.) LNOB c.) do-no-harm	Intervention's proposal, Gender analysis, results models	Interviews with partners and intervention staff	Contribution analysis	Evaluators' assessment, possible subjectivity bias mitigated by data, researcher and method triangulation	A guideline for gender-sensitive budgeting has been developed, but did not get to the stage of implementation - presumably due to the low priorization of the topic on behalf of the partner (interviews 12, 15, 21). Principles like LNOB and do-no-harm were not systematically addressed by the program.	overlap with Questions on Impact and Sustainability?
	6. To what extent are the interactions (synergies/trade-offs) of the intervention with other sectors reflected in conception and ToC – also regarding the sustainability dimensions (ecological, economic and social)?	1. number of identified interactions of the intervention with other sectors 2. qualitative assessment of the interactions with regards to 3 sustainability dimensions	Intervention's proposal, results models	Interviews with partners and intervention staff	Contribution analysis	Evaluators' assessment, possible subjectivity bias mitigated by data, researcher and method triangulation	Given that PFM is relevant to all public policy sectors, the project´s activities potentially have an impact on all sectors. A number of sector ministries are also involved in the piloting of more effective planning (Output-based budgeting; IFMIS) and monitoring (MIS) tools (interviews 12, 22, 24). As all project activities aim to enhance the PFM system, interactions first and foremost address the issue of financial sustainability.	overlap with Questions on Sustainability?
Suitability of the strategy/ the conception to match core problems/needs of the target groups	7. To what extent was the concept designed to reach particularly disadvantaged groups (LNOB principle)? Which prerequisites were addressed for the concept and used as a basis?	1. In the planning stage of the intervention, an appropriate analysis of economic and social impacts of the intervention on particularly disadvantaged proups was conducted. 2. qualitative assessment of application of the LNOB principle	Intervention's proposal, project progress review of predecessor intervention	Interviews with partners, intervention staff and external stakeholder	Contribution analysis	Fair	The principle of LNOB was not systematically addressed by the program. THis is partly due to the fact that the program, despite having the general population stated as the target group, works with key ministries as the direct target group without direct effect on the population (with exception of area 1 that deals with SME taxation). Moreover, the evaluation of the previous phase had recommended to shift the focus away from development orientation towards budget credibility as the key problem identified in the public financial management (PFM) system (interview 6).	
	8. How are the different perspectives, needs and concerns of women and men represented in the change process and how are the objectives represented (Safeguard & Gender)?	1. the different perspectices, needs and concerns of women and men are reflected in the ToC 2. qualitative assessment of the application of gender issues in the project implementation	Intervention's proposal, Gender and HIV/Aids assessmenet, results models	Interviews with partners and intervention staff	Contribution analysis	fair	Different needs and perspectives of women and men were not reflected in the ToC. A guideline for gender-sensitive budgeting has been developed, but did not get to the stage of implementation - presumably due to the low priorization of the topic on behalf of the partner (interviews 12, 15, 21).	
	9. To what extent is the chosen TC-measures' goal geared to the core problems/needs of the target group?	1. Partners confirm that the intervention's objective is relevant to their and the ultimate target group´s needs 2. TC-measures goals are reflected by independent	intervention proposal, desk research (publications by civil society or research institutions)	Interviews with project staff and partners	Contribution Analysis	Possibility of bias because partners may be interested in follow-up intervention. Can be mitigated by differentiated probing: e.g. asking	Issues of transparency, accountability and efficiency pose a key challenge for public service provision to the general population as target group. Moreover, the growing public debt poses a risk for the population with the government risking default or serious fiscal distress over the medium term. These assessments were shared by several external interview partners (interviews 28, 29, 32, 33) and the literature review (ZIPAR 2018 budget analysis (2017),	

		researchers or civil society organizations				for examples, asking for potential for improvement	IMF Article IV Consultations (2017), GIGA Political-Economic Short Analysis (2017))	
	10. Results logic as a basis for monitoring and evaluability (Theory of Change) o Are the hypotheses plausible? o Are the risks presented plausibly?	1. The causal hypotheses in the resutls models are plausible 2. Risk and external factors are presented plausibly 3. The results model is being used by the intervention staff to reflect achievement towards the intervention's objective	Intervention's proposal, results models	Interviews with intervention staff	Contribution analysis	Evaluators' assessment, possible subjectivity bias mitigated by data, researcher and method triangulation		Maybe this question could rather be discussed under the effectiveness criterion?
The design of the project is adequately adapted to the chosen goal.	11. Is the strategic reference framework well anchored in the concept?	1. The intervention's proposal and progress reports refer to overarching principles / frameworks of development cooperation. a.) Internation frameworks, e.g. SDGs / Agenda 2030 b) BMZ strategy on GFG	Intervention's proposal and annual progress reports	/	Document analysis	Fair	offer refers to BMZ Strategy for Good Financial Governance and is implicitly in line with BMZ country strategy from 2016, aspects of BMZ Digital Agenda (using digital innovations for governance and acccountability) and SDG 16.6. (develop effective, accountable and transparent institutions at all levels).	overlap with Q1 and Q2?
	12. To what extent does the strategic orientation of the project address changes in its framework conditions. How is/was the complexity of the framework conditions and guidelines handled? How is/was any possible overloading dealt with and strategically focused?	1. Changes in the framework conditions for the intervention are reflected in the intervention's progress reports (if applicable) 2. The intervention can describe criteria for adapting the implementation to evolving needs from partners	Intervention's proposal and annual progress reports	Interviews with intervention staff, partners and external actors	Comparison between framework conditions and risks described in project documentation and by interview partners	Fair	Organizational (e.g. seperation of planning function from MoF, restructuring of ZRA) or contextual changes (tightening public debt situation, delayed national planning and budgeting bill) were reflected in the progress reports (annual progress reports 2016 and 2017). Evolving partner needs were partly addressed: In area 1, it was decided to support the cleanup of the taxpayer database, once it became evident that this poses a fundamental challenge to effective tax auditing (interviews 11, 17); in area 2, the programme was flexibily providing technical and legal advice on the key reform bills relevant to the sector (interviews 12, 22); in area 3, a peer learning visit was organized with the Treasury of South Africa to work on issues of cash management (interviews 14, 21).	
The conceptual design of the project was adapted to changes in line with requirements and re-adapted where applicable.	13. What changes have occured?	1. The intervention is capable of providing an overview of changes in implementation that resulted from changing framework conditions	Intervention's proposal, annual reports and results models	Interview with intervention staff and partners	Analysis of evolution of intervention's conception	Fair	Key changes were the worsening of the public debt situation, organizational changes (MNDP seperation from MoF, ZRA restructuring, Accountant General`s promotion) as well as political issues within the donor landscape (Western partners losing influence compared to China and others; conflict of interest with WorldBank and DFID). While the ZRA restructuring and AG´s promotion were changes that helped leverage the project´s activities (interviews 11, 16, 14, 21), the political and administrative weakness of the newly established MNDP was something the project could not sufficently counterbalance (interview 13, 30). The shifts in the donor landscape are a topic high on the agenda of the international actors, but there does not yet seem to be a clear strategy to address the issue (interviews 10, 15, 28, 36).	
	14. How were the changes dealt with?							

Effectiveness							
Evaluation dimension	**Analysis questions**	**Evaluation indicators / descriptors**	**Available data sources**	**Other planned data collection projects**	**Evaluation strategy (evaluation design, method, procedure)**	**Expected evidence strength (narrative)**	**assessment by evaluation team**
The project achieves the goal on time in accordance with the TC-measures' goal indicators agreed upon in the contract	1. To what extent has the agreed TC-measures' goal already been achieved at the time of evaluation, measured against the goal indicators?	1. The share of middle-income taxpayers complying with value added tax legislation under the purview of the Medium Taxpayer Offices (MTOs) North and South has risen to 60%.	progress report	interviews with partner, Statistical analysis of data from Zambia Revenue Authority (ZRA)	document analyis, contribution analysis	fair. Assessment may be subject to subjectivity bias, to be mitigated by data and researcher triangulation	The indicator is expected to be met (ZRA data). The information from the progress reports was triangulated with reports provided by ZRA. These figures, however, were stated in a word document without specifying sources, which increases the risk for manipulation. Several interviews (11, 17, 27) could confirm the increase, however.
		2. The deviation of the composition of actual expenditure from that of the approved budget (budget act) with regard to the individual budget heads has diminished by a quarter.	progress report	Comparison between implemented budget and approved budget, MoF data, interviews with partner	document analyis, contribution analysis	fair. Assessment may be subject to subjectivity bias, to be mitigated by data and researcher triangulation	the share was 82% for 2016 compared to 90% in 2015. The figures for the 2017 were not yet available at the time of the field study, but the positive development could be confirmed by the partner and the project expects that the target will be met by the end of the project (interviews 12, 22, 37). It was, however, con-ceded that there is a high level of uncertainty involved, for example because mid-year budget reallocations may continue in the current political climate and emergency situations may require additional resources (interviews 12, 37).
		3. A rule-based procedure for the allocation of funds (e.g. an authorisation procedure based on expenditure in accordance with budget) is applied for one ministerial budget.	progress report	Analysis of data and information from the Ministry of Finance and and the ministry concerned, interviews with partner	document analyis, contribution analysis	fair. Assessment may be subject to subjectivity bias, to be mitigated by data and researcher triangulation	A rule-based warrant system has been piloted in two ministries (of education) in parallel to rolling out Output-based budgeting (interviews 12, 22; concept note by Budget Office & Accountant General - shifting from Cash rationing to a commitment limit). Indicator expected to be achieved by the end of the project.
	2. To what extent is it foreseeable that unachieved goals will be achieved during the current project term?	1. qualitative assessment by the interviewed stakeholders of the objective achievement by the end of intervention 2. Iqualitative assessment by the interviewed stakeholders of achievement of outcome indicators by the end of intervention a.) project partners: ZRA; MoF; MNDP b.) intervention staff 3. Prognosis in progress report to BMZ on achievement of objective and indicators	progress report	Interviews with partners and intervention staff	Contribution analysis	Fair. Data and researcher triangulation to verify information from progress reports to BMZ	The majority of indicators on both output and outcome levels are expected to be achieved by the end of the project. Notable exceptions are: output indicator B3 - it cannot be expected that MIS will be rolled out in 3 sectors + MoF and cabinet office until end 2018 (interviews 13, 23). It cannot be forseen whethe indicator C2 will be met, as the risk management unit, which was supposed to take the lead in operationalizing the government´s risk management framework has not yet been established (interviews 14, 26).
The services implemented by the project successfully contribute to the achievement of the goal agreed upon in the contract.	3. What concrete contribution does the project make to the achievement of the agreed TC-measures' goal, measured against the goal indicators?	1. Interviewed stakeholders state that the intervention / the respective units make a concrete contribution of the intervention objective, measured against the indicators a.) project partners: ZRA; MoF; MNDP b.) intervention staff c.) External stakeholders	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis, Most Signifiacant Change Analysis	Fair. Data and researcher triangulation	action area I: Based on the Tax Administration Diagnostic Assessement Tool (TADAT) prepared in early 2016 with sup-port from different international partners, the intervention assisted ZRA in operationalizing the recommenda-tions made in the assessment (Rojas et al. 2016). These recommendations included cleaning up the tax-payer database in order to have a more accurate foundation for tax audits, simplifying VAT collection through a reverse charge to reduce delays and increase compliance, and supporting voluntary compliance of taxpayers through different measures. In line with the recommendations of the TADAT report, findings of short-term consultants from the Bavarian tax authorities, who were invited by the project to conduct training on indirect tax audit methods, identified a need to enhance the taxpayer database (see Neumann 2017). Based on these recommendations, the project supported the clean-up of a parts of ZRA´s taxpayer data-base through the recruitment of temporary ZRA staff (interviews 17, 27). Technical advice to the senior management of ZRA raised awareness on the benefits and technical aspects of introducing the reverse charge for VAT collection (interviews 11, 16). In order to address the last aspect of voluntary compliance, the project supported ZRA in developing tax education materials with the education ministry, which until now have not been integrated into the school curricula though (interviews 2,18). Besides the database clean-up, the project worked towards enhancing SME tax audits by conducting train-ing on indirect tax auditing methods and revising the existing audit training modules and materials (inter-views 17, 27). This capacity development was reported to be instrumental for ZRA, because it provided tax auditors with new methods and tools (e.g. a template for indirect tax audit methods and sector notes with audit recommendations for specific sectors) which were deemed very useful to
	4. Which factors in the implementation contribute successfully to the achievement of the project objectives?	1. Success factors of the intervention cited by interviewed stakeholders 2. Success factors cited in the intervention's documentation	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis, Most Signifiacant Change Analysis	Fair. Data, method and researcher triangulation	

							improve the quality of tax audits (interviews 17, 18, 19). With regards to output indicator A2 on the quantity of SME tax audits, the field study confirmed the inception report´s suspicion that a quantitative measure does not accurately reflect project progress. While the number of tax audits does influence taxpayer compliance (module indicator M1), it is a function of ZRA`s resources and management decisions that are beyond the scope of the pro-ject (interview 16). The qualitative improvements of tax audits in terms of more accurate taxpayer data and tax audit quality were nevertheless plausibly explained (interviews 11, 16, 17, 18, 20). Another aspect of the project´s approach to increase SME taxpayer compliance is its effort to improve ser-vice quality and to make compliance easier. In this respect, the procurement and introduction of the EQM system was a visible tool to reduce disorder and waiting times at taxpayer offices and it helped ZRA to bet-ter monitor and manage its service offices (interviews 11, 16, 18, 19, 20). For example, average waiting times were reportedly reduced from 2:30 hours to 20 minutes (ZRA data generated by EQM system, see ZRA 2017). Moreover, the project had worked on two Public Private Partnerships (PPP) projects together with German IT companies aiming to make tax registration and returns more accessible via technology. One of these projects will be continued beyond this project phase with additional funding (interview 11). Finally, the strategic advice provided by the project supported the development of the SME taxation strate-gy that formulated targets and measures for further increasing SME taxpayer compliance in the medium term (ZRA 2017). In particular, the project´s technical advice helped identify key areas of improvement, for example to simplify tax regulation for small businesses and to facilitate their compliances via service points on markets or via mobile applications (interviews 17, 27). The strategy was developed in a workshop facili-tated by the project and adopted by ZRA at the end of 2017. ZRA requested support from GIZ to implement two of the 50 measures set out in the strategy, which forms the basis for project activities in the remaining year of the project (interview 11). action area II.a: Several measures to improve budget planning were introduced with support of the project, which form part of output indicator B2. With regards to the legislative framework, legal and technical advice of the senior advisor in this area helped shape and apply crucial reform bills including the excess expenditure bill (inter-views 12, 22). The project supported the MoF budget office in using provisions of the new constitution to increase legislative oversight. And with the excess expenditure bill the previous practice of the government to normalize excess expenditures through appropriation bills was removed. Hence, the advisory services helped reduce the number of supplementary budgets and excess expenditures, which marks an important contribution towards module indicator M2 on reducing the deviation between approved and actual expendi-tures (interviews 12, 22). In order to help MPSAs to formulate more credible and effective budgets, the project supports the use of OBB, which links the budget back to national and sectoral plans. It had been initially piloted in the education ministries outside the framework of this project and independently evaluated with financing from the Worldbank PFM reform project in early 2016 (see Bird et al. 2016). Based on this evaluation, the project has developed OBB manuals for developers and users and conducted training with additional ministries to roll it out. For the time being, in addition to optimizing implementation with the education ministries, OBB has been introduced in the Ministry of Community Development and Social Welfare in 2017 and more MPSAs are planned for the 2019 budget (interview 12). Based on OBB, the project further aims to increase MoF capacities through the budget challenge function. For this purpose, a budget challenge manual was devel-oped that supports the analysis of ministerial budgets. In a second step, cluster hearings were facilitated that are jointly chaired by the MoF budget office and MNDP with the goal of analysing and challenging the estimated budgets of MPSAs around NDP clusters (interviews 12, 22). Beyond the direct technical partner level, the project also reached out to the political and to the citizen´s level. In order to enhance the political ownership of the budget, the project supported budget office in re-designing the supplementary budget and excess expenditure process to deepen the Cabinet`s understand-ing of the consequences of in-year political decisions. Based on this, pre-Cabinet meetings were initiated to discuss budgetary matters separately before full Cabinet meetings, which was reported to be very effective by the partner (interview 22). With regards to the citizen's level, the project provided technical advice to MoF to produce and disseminate the citizens budget in a timely manner. This activity was complemented by the GIZ civil society´s program, which supported CSO in submitting proposals to the national budget (interview 31).

							Indicator B1 (6 of a total of 54 MPSAs prepare their budgets in the IFMIS budget module), however, was found to be unsuitable during the field mission, as its achievement lies outside the sphere of influence of the project and does no longer reflect its approach. According to the PFM reform coordinator, the IFMIS budget module has been implemented in 44 MPSAs already, but it is still not functional due to different shortcomings (interview 21). Given that the Worldbank supported PFM reform program is pursuing the gov-ernment-wide roll-out, this project has chosen to work on addressing the existing shortcomings. In this con-text, the project plans to reconfigure the IFMIS module in line with OBB requirements through external con-sultants in 2018 (interview 12). This is expected to help enhance budget effectiveness by linking the budget module with national plans. The issue of more effective expenditure authorization, which is connected to the achievement of module indicator M3, was addressed by supporting the budget office in moving from a system of cash rationing towards a warrant system. This warrant system was so far piloted in two MPSAs with the goal of helping the budget office in re-focusing their work on budget analysis rather than the authorization of expenditures (interview 12, MoF 2016). During piloting difficulties emerged, because the effective use of a warrant sys-tem requires coordination between the relevant stakeholders including budget office and the cash man-agement unit und the Accountant General`s office, which is so far not sufficiently given (interviews 12, 14). When investigating this issue further, the concern was voiced that the PFM coordinator under the Account-ant General´s office is not fully informed about reform activities in the budget office, indicating that limited coordination is a critical factor (interviews 3, 21). action area II.b: The action area`s inherent ToC had changed two times throughout the course of the project, which howev-er was not reflected in changes of the indicator system. The first one was developed in 2014 with the plan to newly establish a comprehensive IT-structure at the apex level (MNDP, MoF, Cabinet) and five key minis-tries with support of the EU, DFID and GIZ. During the inception phase in 2015, it was found that this plan was too ambitious and it was narrowed down to a four-pillar approach focusing on M&E, IT, capacity devel-opment and the governance structure. After another review, a third ToC was finally formulated in July 2017, which focused on the apex level while limiting the efforts on establishing IT systems in the sector ministries. This approach was chosen with the goal of having a more realistic workplan until the end of the project term (interview 30). While formally agreeing to this approach, however, DFID and GIZ were not aligned regarding the obligations under the ongoing delegated coordination agreement, which created uncertainty on behalf of the partner on how to proceed (interviews 15, 24, 30). Moreover, the results model and indicator system for this action area remained unchanged inspite of the aforementioned strategic changes and inspite of the fact that since project beginning the national planning function had moved from MoF to the newly estab-lished MNDP, meaning that the project in this area no longer contributed to increasing MoF capacity (output B) but MNDP capacity. The implicit underlying hypothesis hence is that by developing national and sectoral performance indicators, by supporting the underlying processes and by eventually developing suitable IT systems, the capacity of MNDP to inform and guide the budget preparation process is increased. With regards to the first aspect of developing performance indicators, considerable progress was made. A common M&E strategy was developed, subsequently MNDP was assisted in breaking down the long-term development plans of Vision 2030 into national and sector performance frameworks. Key performance indicators were defined and data catalogues were established, all of which contributed to strengthening the monitoring capacity of MNDP. This progress is also reflected in the gradual reduction of time lag for the publication of annual progress reports from almost 1.5 years for the 2014 to 8 months for the 2016 report and the 2017 report expected to come on time (interviews 13, 24). With regards to the underlying processes, however, efforts to establish a suitable governance structure for MIS implementation with MNDP in lead and the accompanying capacity development activities could not be implemented as planned. In consequence, this resulted in delays regarding the implementation of IT sys-tems (interview 13). While the various existing IT tools in the ministries were analysed in terms of the possi-bility to build an integrated IT system, the project did not get to the point of advancing on the implementation of the MIS. Instead, the project focused on supporting the enhancement of the Monthly Monitoring System (MMS), which was developed by the predecessor project as an interim solution (interviews 24, 25). action area III: In spite of the delayed start of activities due to the project´s initial lack of human

							resources in this area, the project has already contributed to more effective budget execution in several ways. Firstly, a business pro-cess mapping for payroll activities was conducted by external consultants which is currently under review by the Account General and two more business process mappings are planned on investments and on revenue. These business process mappings are assessed positively and are expected to help the Ac-countant General´s office to guide activities within the PFM reform programme on further advancing the IFMIS roadmap. Based on these business process maps, the roles and responsibilities of actors involved in budget implementation can be clarified (interviews 14, 21). Secondly, the project supported the strengthening of the audit function. After completion of a first compre-hensive IFMIS audit under the previous project phase, a second IFMIS audit was conducted by the Internal Audit department with project support to analyse transactions and identify weaknesses still inherent in the current IFMIS. While the results of the audit were kept confidential by the Accountant General due to the sensitivity of the contents, the audit was considered helpful in identifying and addressing shortcomings of the IFMIS module (interviews 21, 26). Additionally, the Internal Audit department was supported in adhering to quality assurance in line with international standards through training, conference participation and tech-nical advice. Building on the activities initiated in the previous phase of the intervention, the internal audit manuals were revised to include quality assurance and a performance audit is planned of an agricultural subsidy support programme under the Ministry of Agriculture to test and further enhance the capacity of the internal audit department (interviews 25, 26). Thirdly, to facilitate issues of cash management, a peer learning visit of the South African Treasury was organized to enhance functionality of the newly established Treasury Single Account (TSA). The visit was reported to be highly successful by the partner, as it provided a first-hand learning experience and identified key areas for improvement regarding cash management (interviews 14, 21, 37). The visit was also im-portant, as it brought together actors from the Accountant General´s office (key partner in action area III) and the budget office (key partner in action area II.a), both of which had described experiencing a lack of communication and coordination from each other (interviews 21, 22). Enhanced coordination between these two actors is considered crucial for the budget office to assume more of its core function of budget analysis and planning again and allow the cash management unit under the Accountant General deal with budget allocation (interview 14). Lastly, the government-wide risk framework that was first drafted in 2015 (see MoF 2015) had been passed by Cabinet in 2017, but still needs to be operationalized. The process was spearheaded by the internal audit department with the technical advice of the project as well as its predecessor phase (interviews 24, 25). The necessary risk management unit is not yet in place, however, therefore no risk training has been conducted so far (interview 14, 24). Once realized, the risk management system is expected to contribute to an increased capacity to effectively implement the budget according to the plans by being able to identify and systematically address risks in the budget execution process (interviews 14, 25).
	5. What other/alternative reasons contributed to the fact that the objective was achieved or not achieved?	1. Description of alternative hypotheses cited in interviews 2. Description of alternative hypotheses cited in progress reports	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis	Fair. Data triangulation	action area I: In order to better assess the contribution that the above described activities and results have made to the objective of increasing SME taxpayer compliance, they were juxtaposed with a number of alternative ex-planations and discussed with the interview partners as well as internally among the evaluation team. Look-ing at the outcome indicator M1, the following alternative explanations were carved out has being relevant: Firstly, the organizational restructuring initiated by the new Commissioner General in 2016 had initially re-sulted in a drop in the number of tax audits in 2016, but subsequently both quantity and quality of audits as well as VAT compliance for medium taxpayers had increased due to the more effective organisation of ZRA units under a direct and an indirect tax division (interviews 11, 16). Secondly, the introduction of online tax registration in 2016 and a tax amnesty introduced in 2017 had created pressure and incentives respec-tively on SME to comply with the tax legislation (interviews 16, 17). A third factor was the introduction of new reverse VAT collection practices, which the project had influence on but without being a driving force. Through the MSC approach it was revealed that some interview partners viewed the project´s activities, in particular the introduction of the EQM system as the most significant positive development of the past two years (interviews 19, 20), while others considered external factors most significant, in particular the organi-zational restructuring (interview 27) and the reverse VAT collection practice (interview 16). Nevertheless, the majority of interview partners including the latter two reported that without the project´s contribution to the quality of tax audits and taxpayer services, the

							achievements would not have been possible (interviews 16, 18, 19, 20, 27). action area II.a: It was reported that the constitutional amendment from 2016, which required all in year changes to the budget to be endorsed by parliament in supplementary budgets, also contributed to the reduced deviation of planned and actual expenditures (module indicator M2). Moreover, the results of the constitutional amendment may reflect more formal than actual improvements in budget credibility, because of the stated practice of formally approving excess expenditures through appropriation bills (interviews 12, 22). However, beyond the formal effect it had, the new constitution could be used as an argumentative tool to further ad-vance the reform agenda regarding supplementary and excess expenditures, hence making an effective contribution in itself (interviews 12, 37). action area II.b: Investigating causes of the observed delays, the working note accompanying the revised ToC in July 2017 identified the following core challenges to advancing with MIS implementation: Firstly, MNDP, in particularly the M&E unit, was reported to be heavily understaffed. Secondly, a weak management capacity in MNDP combined with the lack of a crucially needed project management unit was identified. Thirdly, it was ob-served that numerous IT systems and initiatives exist in the government without proper coordination (see DFID 2017, p.4). These observations could be confirmed in the interviews conducted (interviews 13, 30, 35). Moreover, it was elaborated that MNDP holds a weak position within the government and that the con-flict between DFID and GIZ on their respective obligations and the approach to take towards further imple-mentation (see also relevance chapter) additionally impeded the project progress (interviews 13, 15, 23). Finally, the lack of an effective steering committee further added to the partner finding it difficult to move forward with its reform agenda. In this context, it was reported that the partner was left unclear about the type and degree of support to be expected from the project (interview 24, 30), which could have been pre-vented by effective use of a steering committee. Ultimately, this combination of factors lead the result that the intervention could in this action area not deliver what was initially expected. action area III: In order to put the project´s contribution to improved budget execution into perspective, potential alternative explanations for improved budget execution were analysed and discussed with the interview partners. Two key changes outside the project´s influence were identified over the last two years, which appeared to have played a role, too. Firstly, the promotion of the Accountant General to Permanent Secretary level helped him and his office to fulfil their roles in budget execution more effectively, because it strengthened their position. Secondly, the establishment of the TSA was revealed as a crucial measure to enhance budget execution by improving appropriation control, by making cash management more efficient and by allowing more complete and timely information on government resources. The key role of the Treasury Single Ac-count was confirmed through the MSC approach in interview 21. While both of these aspects were reported to be important factors for improved budget execution by different interview partners, it was also made plausible that the project could leverage its activities through these changes (e.g. South African Treasury visit to enhance TSA, address shortcomings in IFMIS after audit) (interviews 14, 21, 34). Hence, the project did make a difference towards enhancing partner capacities on budget execution.
	6. Are core, support and management processes designed in such a way that they contribute to the achievement of the objective?	1. Strengths and weaknesses of core, support and management processes according to different stakeholders a.) Intervention staff b.) Partners c.) External stakeholders	documents on the governance and steering structure	Interviews with partners and intervention staff and external stakeholders	Contribution analysis	partly evaluators' assessment, possible subjectivity bias mitigated by data and researcher triangulation	No coherent steering structure exists for the projet, which bears a risk of critical problems not being addressed adequately and on time, as was the case in area 2b. A weekly internal team-meeting as well as weekly partner-meetings on the technical level help coordinate processes effectively (interview 6).
	7. To what extent have risks (see also Safeguards & Gender) and assumptions of the Theory of Change been addressed in the	1. reflection of risks and assumptions from the ToC in progress reports 2. reflection of risks and assumptions from the ToC by project partners	progress report	Interviews with partners and intervention staff		partly evaluators' assessment, possible subjectivity bias mitigated by data and researcher triangulation	It was found that risks were stated in the results model and brought up in the discussions with project team-members (e.g. political abuse of power, misappropriation of funds, decreasing weight of Western coopera-tion partners), however they were not systematically addressed in project implementation (interviews 11, 14, 15). Especially action area I, dealing directly with (often intransparent ways of) domestic revenue genera-tion and auditors of SME taxpayers, faces a potential risks of implicitly supporting harmful or inappropriate practices (e.g. selective targeting of tax audits, unfair burden to small taxpayers). Similarly, the other areas deal with sensitive

	implementation and steering of the project?						areas of public financial management, which carries a high risk of corruption. Neverthe-less, no strategy exists to deal with such risks effectively in terms of an internal risk management, exit strat-egies or safeguard mechanisms (interview 15). The project team reported that it plans to address such issues more systematically in the next programme phase from 2019. Another potential negative unintended result is the risk of limiting the partner´s ownership by driving the PFM reform agenda too much from the outside, rather than letting the partner government set the right prior-ities by themselves (interview 14). This risk applies to this project as much as to other projects, particularly the Worldbank supported PFM reform programme, and it is also reflected in the observation that the actual political will may be lower than the formal commitment (see relevance chapter). Given the divisions of la-bour within the German development cooperation with BMZ funding and negotiating projects and GIZ im-plementing them, this risk can be mitigated only on the political level of the intergovernmental consultation process. In that regard, however, the project has limited influence, therefore it will not be considered for the assessment of the project´s effectiveness.
The occurrence of additional/ not formally agreed positive results and unintended negative results was assessed and adequately addressed where required.	Refers to Option A, Sustainability (determination of interactions in effectiveness and impact):						
	8. To what extent were risks of unintended results assessed as observation fields by the monitoring system (e.g. compass)?	1. Description of observation of risks in the intervention's workflow 2. Description of mitigations strategies adopted towards risks by intervention	Intervention proposal, progress reports	Interviews with intervention staff and partners	Contribution analysis	Limited	
	9. To what extent have the project's benefits produced results that were unintended?	1. Description of unintended results on a.) the target group of the project b.) the partner institutions	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis	Fair. Data triangulation	With regards to unintended results, additional results – whether positive or negative – were not reported, but also not specifically considered in the project´s monitoring system. In terms of the former, no interview part-ner had observed any additional results except the risk for potential negative results discussed below (in-terviews 11, 18, 20). In terms of the latter, it was reported that the results monitoring was used as a reporting tool for progress towards milestones and the indicator goals, rather than as a management instrument (in-terview 6). While the project uses an excel-based tool to monitor project progress on a weekly basis, no specific monitoring for unintended results was reported (interview 15).
	10. Which positive or negative unintended results (economic, social, ecological) does the project produce? Is there any identifiable tension between the ecological, economic and social dimensions?	1. description of unintended positive or negative results according to interview partners on the a) economic level b) social level c) ecological level 2. assessment of tension between the 3 levels	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis, Most Signifiacant Change Analysis	Limited	
	11. How were negative unintended results and interactions counteracted and synergies exploited? What measures were taken?	1. Description of mitigations strategies towards unintended results according to intervention team 2. Descirption of exploitation of synergies according to different stakeholders a.) Intervention staff b.) Partners c.) other projects, in particular DFID, EU and MTDF	progress report	Interviews with partners and intervention staff and external stakeholders	Contribution analysis	Fair. Data and researcher triangulation to verify information from progress reports to BMZ	

Impact							
Evaluation dimension	**Analysis questions**	**Evaluation indicators / descriptors**	**Available data sources**	**Other planned data collection projects**	**Evaluation strategy (evaluation design, method, procedure)**	**Expected evidence strength (narrative)**	**assessment by evaluation team**
The announced superordinate long-term results have occurred or are foreseen (should be plausibly explained).	1. To which superordinate long-term results should the project contribute (cf. module and programme proposal, if no	Impact indicator from programme proposal: 1. The tax rate has increased from 15.8% to 18%.	Narrative assessment of progress towards indicator in progress report	document analysis of IMF data, interviews with intervention staff, partners and external stakeholders	Contribution analysis	, 2017)	preliminary figures for 2016: tax rate 12.9% (source: Annual Article IV Consultations) Despite successful on SME taxation in area 1, the overall tax rate has decreased rather than increased over the last years.

	individual measure; indicators, identifiers, narrative)?	Impact indicator from programme proposal: 2. Zambia's fiscal deficit has decreased from 8.14% of the GDP to less than 5% of the GDP.	Narrative assessment of progress towards indicator in progress report	document analysis of IMF data, interviews with intervention staff, partners and external stakeholders	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	preliminary figures for 2016: fiscal deficit (Cash basis) 5.8% (source: Annual Article IV Consultations) The fiscal deficit remains high, underlining the risky PFM framework conditions.
		Impact indicator from programme proposal: 3. Zambia's sovereign debt has stabilised at the level of approximately 40% of the GDP (preliminary estimate Article IV Consultations of the International Monetary Fund, IMF, June 2015)	Narrative assessment of progress towards indicator in progress report	document analysis of IMF data, interviews with intervention staff, partners and external stakeholders	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	preliminary figures for 2016: public debt 60.5% (source: Annual Article IV Consultations) The public debt remains high, underlining the risky PFM framework conditions.
		Impact indicator from programme proposal: 4. The Open Budget Index has improved from 39 in 2015 to 50 in 2018.	Narrative assessment of progress towards indicator in progress report	document analysis ofOpen Budget Index, interviews with intervention staff, partners and external stakeholders	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	The OBI dropped from 39 to 8 in 2017.
	2. To what extent will the project contribute to the implementation of the partner country's national strategy for implementing Agenda 2030/to the SDGs?	1. the project proposal reflects Zambia`s national strategy for implementing the SDGs 2. qualitative assessment of the project`s contribution to implementing Zambia`s national strategy for implementing the SDGs by the political partner	Progress reports	Interviews with intervention staff and political partner	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	the project´s contribution to the national and international development agenda is limited. While the project in action area II.b supports the implementation of the SDGs through operationalization of the 7th NDP by establishing KPIs and seeking to integrate national planning with budgeting, it failed to move further with implementing it in terms of connecting it to the budget. Here, considerable delays in the implementation of M&E systems and necessary support structures in the government undermine the potential of achieving further impact in this area (interviews 12, 24, 30).
	3. Which dimensions of sustainability (economic, ecological, social) does the project affect at impact level? Were there positive synergies on the three levels?	Narrative assessment of the project´s impact, and where applicable synergies, with regards to the a.) economic dimension b.) ecological dimension c.) social dimension	Progress reports	Interviews with intervention staff and project partner	Contribution analysis, Most Significant Change Analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	the field study found only partly links to the three dimensions of sustainability. For this project, the economic dimension of sustainability is most relevant. With regards to the immediate target group (partner institutions), the project contributes to increasing the economic sustainability of public finances by enhancing legal and administrative processes of budget planning and execution (interviews 21, 24, 26). Social and ecological dimensions of sustainability are assumed to be addressed indirectly through the more efficient provision of social and environmental policies (see BMZ 2014).

	4. 'Leave No One Behind': To what extent have targeted marginalised groups (such as women, children, young people, the elderly, people with disabilities, indigenous peoples, refugees, IDPs and migrants, people living with HIV/AIDS and the poorest of the poor) been reached and is there evidence of the results achieved at target group level?	narrative assessment of the project´s application of the LNOB principle	Progress reports	Interviews with intervention staff and project partner	Contribution analysis, Most Significant Change Analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	Likewise the LNOB principle, which aims to make sure that particularly marginalized groups are not neglected by any development measure, is partly considered.With regards cross-cutting issues, the project did not manage to achieve any results beyond the activity phase. While gender-sensitive budgeting guidelines were developed, they were never implemented by the partner, presumably due to low priorization from the government (12, 15, 37).
The project contributed to the intended superordinate long-term results.	5. To what extent is it plausible that the results of the project on the output and outcome levels (project goal) contribute to the superordinate results? (contribution-analysis approach)	1. Narrative assessment of contribution of intervention to program objective 2. narrative assessment of the plausability of the results model (ToC)	progress reports, interventional proposal	Interviews with intervention staff, external actors and partners	Contribution analysis	Fair	There is a plausible link between the project´s output/outcome objectives and the impact level. Activities in area 1 contribute to improvements on the tax rate (programme indicator 1) as well as the public deficit and overall debt (indicators 2 and 3), areas 2 and 3 may contribute to programme indicators 2, 3 and 4 (Open Budget Index). Compared to external influence factors (global economic developments, especially copper price; increased external loans; political climate within Zambia), the project´s activities appear to play a rather small role in contributing to achieving these indicators, though (interviews 15, 36).
	6. What are the alternative explanations/reasons for the results observed? (e.g. the activities of other stakeholders)	1. Narrative assessment of alternative explanations	progress reports	interviews with intervention staff, partners and external stakeholders	Contribution analysis	Fair	When looking at further factors that positively or negatively influenced impact, the interaction with other international players played a mixed role. The Worldbank program under the $17 mio. MDTF had been the main driving force of the PFM reform process until now. Overall, the program was reported to have made an important contribution to PFM reform, without which the Zambian government would currently not stand where it is with regards to the Treasury Single Account or the rollout of IFMIS, for example (interviews 21, 34). While the project is formally aligned with the PFM reform program and has tactically adapted to lever-aging its results (see effectiveness chapter), a closer cooperation could have contributed to better address impacts in a coordinated way. Moreover, the project came under scrutiny from co-funding partner DFID in action area II.b due to delays in implementation, which were reportedly both a result of delayed processes within the government and insufficient coordination both on a technical and a political on behalf of DFID and GIZ (interview 15, 24, 30). As a result, less than 20% of the planned co-funding budget was used and this action area could not sufficiently contribute to wider impacts on the societal level. Looking beyond the coordination with other cooperation projects, the loans from China and other non-western partners for large infrastructure projects outside the regular budget framework were reported not only to increase stress on public finances but also to indicate a change in the political dynamics of interna-tional cooperation in Zambia. On the one hand, this resulted in raising questions of debt sustainability, which indicate the lack of political will for reform (interviews 28, 33). On the other hand, the government´s confidence in taking these loans in spite of IMF`s threat to withhold loans from their behalf was assessed as sign of a decreasing weight of western-style cooperation on a larger level (interviews 10, 15, 36). Put together, these points indicate that the project can only make a limited contribution to wider macro-economic impact in light of the political forces that neglect the need of a sustainable PFM system.
	7. To what extent do changes in the framework conditions influence superordinate long-term results?	1. Narrative assessment of framework conditions that influence impacts a.) macro-economic developments b.) changes in the political landscape c.) institutional environment of the partner	progress reports	interviews with intervention staff, partners and external stakeholders	Contribution analysis	Fair	Compared to external influence factors, however, most notably global economic developments such as the volatile copper price (downward trend, but increase in 2017), increased external loans from countries out-side the established donor coordination circles, and a political climate within Zambia characterized by in-transparency and a tightening government control, the project´s activities were reported to play a rather small role in contributing to achieving these indicators though (interviews 15, 28, 36). it was reported by several external stakeholders that in spite of incremental improve-ments on the technical level, larger political challenges remained untouched. These include the limited polit-ical will for reform beyond lip service, a continuation of intransparent procurement practices as well as an observed gap between written law and its practical application. These aspects were

		d.) activities by other stakeholders					reported to reduce the impact that the technical cooperation of the project could have had (interviews 8, 28, 29, 33).
	8. To what extent is the effectiveness of the development measures positively or negatively influenced by other policy areas, strategies or interests (German ministries, bilateral and multilateral development partners)? What are the consequences of the project?	Narrative assessment of positive or negative influence that the project experiences from a.) German ministries b.) DFID / UK government c.) other bilateral or multilateral development parnters	progress report	interviews with intervention staff and project partner, interview with German embassy and other development partners (DFID, Worldbank, etc)	Contribution analysis	Fair	The activities in area 2b (M&E) came under scrutiny from co-funding partner DFID due to delays in implementation, which appear to be a result of both delayed processes within the government and insufficient coordination on behalf of DFID & GIZ. As a result, less than 30% of the planned co-funding budget was used and objectives are not expected to be met (interviews 12, 24, 30). Considerable loans from China and other non-western partners of Zambia for infrastructure projects outside the regular budget framework are putting stress on public finances and pose questions of debt sustainability.
	9. To what extent has the project made an active and systematic contribution to widespread impact? (4 dimensions: relevance, quality, quantity, sustainability; scaling-up approaches: vertical, horizontal, functional or combined)? If not, could there have been potential? Why was the potential not exploited?	1. Narrative assessment of the projects contribution to widespread impact with regards to: a.) 4 dimensions b.) scaling-up approaches	progress reports	interviews with intervention staff, partners and external stakeholders	Contribution analysis, Most Significant Change Analysis	Fair	The project´s impact in terms of scaling it up was limited by the reported silo structures of the government. While certain aspects of the project were successful in reaching out further than the immediate partners at the center (e.g. EQM and new tax audit methods in the regions in action area I; involvement of other government agencies in piloting OBB in action area II.a; gradual rollout of enhanced internal audit techniques to internal auditors in the line ministries in action area III), the project falls short of scaling up further to regions, government agencies or the wider population. This stems partly from the challenging political environment as described above, but also from the silo structure within the government, which was brought up at multiple times during the field study (5,15, 23). In this context, it was reported that communication and coordination between different ministries and even units within ministries is often insufficient. For example, the PFM reform coordinator under the Accountant General´s office was reported to lack information on activities of the budget office (interview 21). Given that both the Accountant General`s office and the budget office are supported by different action areas of the project, this silo structure was mirrored by the project, which hindered the transfer of knowledge and processes to further levels (interviews 5).
	10. Referring to the three dimensions of sustainability (economic, ecological, social): How was it ensured that synergies were exploited in the three dimensions? What measures were taken?	narrative assessment of synergies between the 3 dimensions of sustainability in the implementation of the project		interviews with interventional staff and project partner	Contribution analysis	Limited	see Q3
Unintended superordinate long-term (positive or negative) results have occurred.	11. Which unintended positive and/or negative results/changes at the level of superordinate results can be observed in the wider sectoral environment of the development measure (e.g. cross-cutting issues,	1. Description of unintended results a.) Positive b.) Negative	progress reports	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	Unintended results were not observed on the impact level beyond the two aspects of potential negative results discussed in the effectiveness chapter that also influence the impact level. One such potential unintended impact is the potential unfair burden on small taxpayers discussed above. Another such potential negative impact is the fact that government may be hindered in establishing a sustainable PFM system itself due to perceived push factors from interventions such as this one (interview 14). While government representatives spoke very positively of the progress of implementing IFMIS across all MPSAs (interviews 21, 25 ,37), several other stakeholders saw it much more critically, arguing that after over ten years of implementation it is still not uniformly used and functional (interviews 14, 31, 33). Given the tendency to demand tangible results and make support conditional on quick progress, it may hence be a negative unintended result that the project is contributing to unrealistic priorizations on behalf of the partner that ultimately result in a lack of PFM reform (interview 14).

	interactions between the three sustainability dimensions)?						
	12. To what extent is the (positive or negative) contribution of the project plausible?	1. Narrative assessment of plausibility of intervention's contribution to unexpected results		Interviews with intervention staff, partners and external actors	Contribution analysis	Evaluators' assessment, possible subjectivity bias mitigated by data, researcher and method triangulation	no unintended impacts were observerd. Potential unintended impacts were discussed in Q3.
	13. What are the alternative explanations/ reasons for the results observed? (e.g. the activities of other stakeholders)	1. Narrative assessment of alternative explanations for unintended superordinate results	progress reports	interviews with intervention staff, partners and external stakeholders	Contribution analysis	Fair	-
No project-related negative results have been observed – and the project responded adequately if any negative results were determined at any time.	14. Have negative results occurred?	1. Description of unintended negative results	progress reports	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	none observed
	15. To what extent were the risks of negative, unintended, superordinate results identified and assessed in the monitoring system? To what extent were these negative results in the sense of (negative) interactions or trade-offs in the ecological, economic and social dimensions already known during the conception of the project and reflected (e.g. in the module or programme proposal)?	1. Risks of unintented negative results are included in the intervention's monitoring systens 2. Risks of unintented negative results were reflected in the module or programme proposal	Results Models, Impact Monitor	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	see Q7 on the effectivenes criterion. No further risks were observed.
	16. Was there a corresponding risk assessment in the TC-measures' proposal? How was the ability to influence these risks originally assessed?	1. The intervention's proposal includes a risk assessment 2. The risks cited in the risk assessment correspond to the risks cited by interview partners 3. Description of mitigation measures by the intervention towards risks	Intervention's proposal	Interviews with intervention staff, external actors and partners	Contribution analysis	Fair, data triangulation	The assessment of the project´s risk analysis and handling provides a mixed picture. The proposal includes a number of key risks (e.g. politically motivated reallocation of funds, insufficient rule of law, more profound limitations of IFMIS than expected) and suggests mitigation strategies, though it assessed the influencing ability of the project as low (GIZ 2015a, p. 18). The risks stated in the proposal were also brought up by interview partners (interviews 15, 32, 33, 36). Nevertheless, the project did not appear to follow a coherent strategy to address these risks, yet. It rather sought to apply its technical advisory capacity flexibly as to mitigate the risks as much as possible on an adhoc basis (interview 15). For the upcoming phase, however, the project is planning to include a number of aspects in the project design to better address existing risks in a structured way (interview 15, 36) .

	17. To what extent have the project's services caused negative (unintended) results (economic, social, ecological)? Is there any identifiable tension between the ecological, economic and social dimensions? o Economically: Impairment of competitiveness, employability, etc. o Socially: How should the impact be assessed in terms of distributive results, non-discrimination and universal access to social services and social security systems? To what extent can particularly disadvantaged population groups benefit from the results or have negative results for particularly dis-advantaged population groups been created? o Ecologically: What are the positive or negative envi-ronmental impacts of the project?	1. Description of unintended negative results along the 3 sustainability dimensions 2. narrative assessment of the tensions between 3 sustainability dimensions	progress reports	Interviews with partners, intervention staff and external stakeholders	Contribution analysis	Fair	see Q3
	18. What measures have been taken by the project to counteract the risks/negative interactions?	3. Description of mitigation measures by the intervention towards risks	Intervention's proposal	Interviews with intervention staff, external actors and partners	Contribution analysis	Fair	no strategy exists to deal with such risks effectively in terms of an internal risk management, exit strategies or safeguard mechanisms (interview 15). The project team reported that it plans to address such issues more systematically in the next programme phase from 2019.
	19. To what extent have the framework conditions for the negative results played a role? How did the project react to this?	1. Description of framework conditions that influence impacts a.) macro-economic developments b.) changes in the political landscape c.) institutional environment of the partner d.) activities by other stakeholders 2. Description of mitigation measures	Intervention's proposal, progress reports	Interviews with intervention staff, external actors and partners	Contribution analysis	Fair	While external factors such as macroeconomic factors and the political situation play a large role in explain-ing these shortcomings, the project did not take strategic measures to address the latter by addressing the political level

		by the intervention towards risks					

Efficiency							
Evaluation dimension	**Analysis questions**	**Evaluation indicators / descriptors**	**Available data sources**	**Other planned data collection projects**	**Evaluation strategy (evaluation design, method, procedure)**	**Expected evidence strength (narrative)**	**assessment by evaluation team**
The project's use of resources is appropriate with regard to the outputs achieved. [Production efficiency: Re-sources/Services in accordance with the BMZ]	1. To what extent are there deviations between the identified costs and the projected costs? What are the reasons for the identified deviation(s)?	1. Description of differences between planned and actual costs 2. Description of reasons for differences between planned and actual costs	Current cost commitment report	Interview with head of programme, heads of unit and head of administation and finance	Comparison and analysis	Fair	until the end of 2017, i.e. after two thirds of the project term, only around EUR 210.000 of the earmarked EUR 3.5 million cofunding from DFID had been used as a result of a combination of factors described ear-lier in the effectiveness chapter. While the considerable cut in spending can be seen as an efficient step to prevent the waste of resources when objectives cannot be achieved, it fails to fulfil the maximum principle of achieving the maximum possible results with the given resources and the evaluation team believes that in spite of the challenging context, a number of measures could have prevented the stalemate in project progress. Firstly, the inadequate reflexion of action area II.b in the results model and indicator system of the measure limited the capacity of the project to effectively monitor and manage the project. Several interview partners stated the overly ambitious planning of the project in the first place (interviews 5, 13, 30, 35). While the ToC had then been revised twice in 2016 and 2017, these changes were not translated into the indica-tor system. Secondly, the lack of an effective steering committee with the partner meant that critical issues could not be addressed in a timely manner and the partner was left unclear about what support to expect (interviews 23, 24). Thirdly, the poor coordination on both a technical level (i.e. what are the common objec-tives and what is the common strategy to achieve these?) and a political level (what are the modalities of the delegated cooperation agreement?) between GIZ and DFID lead to misunderstandings and considera-ble funding delays/reductions, preventing the project to efficiently make use of the resources (potentially) available. For the remaining year, the amount of resources still allocated in the framework of the cofunding agreement was still in discussion with DFID at the time of the field study, but in light of the considerable delays and the limited absorption capacity of MNDP, the stated objective was merely to foster the interim achievements (such as enhancing the interim solution of the MMS) and underlying M&E processes (inter-views 8, 13, 15), which falls short of achieving a maximum of results with the given resources.
	2. To what extent could the outputs have been maximised with the same amount of resources and under the same framework conditions and with the same or better quality (maximum principle)?	1. Qualitative assessment of possibilities for yield maximization on output level	/	Interviews with intervention staff and partners	analysis	Evaluators' assessment, possible subjectivity bias mitigated by data and researcher triangulation	Analysing aspects of implementation efficiency, the project team provided several examples of following the principle of efficiency, but also limitations in the way the project navigates the partner structure were observed. In terms of positive examples, it was reported that allowances were not paid for training activities in order to save resources and to avoid setting false incentives (interview 14). It was also stated that the team attempts to make use of their own work time as efficiently as possible and that all decisions on ex-penses were prudently taken, though the methods or approaches used could not be specified (interview 11). Regarding limitations, it was reported that the PFM reform coordinator in the MoF has limited knowledge of what the GIZ project is doing in the other action areas, creating a risk for inefficiencies (inter-view 21). In line with this observation, several other stakeholders characterized the partner structure as operating in "silo structures" (interviews 14, 15, 17, 23), which the project did not sufficiently manage to break as was conceded by the project itself (interview 5). Besides, the project´s lack of an effective formal steering structure to coordinate activities, address problems and to ensure subsequent scale-up with the partners is seen as a shortcoming (interviews 15, 21, 30). In particular, in action area II.b it became clear that the insufficient communication on various levels resulted in a situation that made an efficient and effective resource use impossible. Finally, it was noted that parts of the BMZ funding for 2016 were withheld, which then had to be spent quickly in a potentially inefficient manner the following year (interview 11). While this circumstance lies outside the sphere of influence of the project and is therefore not considered in the assessment, it re-quires mentioning as a factor inhibiting efficient resource allocation.
	3. To what extent could outputs have been maximised by reallocating resources between the outputs?	1. Qualitative assessment of possibilities for yield maximization on output level	/	Interviews with intervention staff and partners	analysis	Evaluators' assessment, possible subjectivity bias mitigated by data and researcher triangulation	the procurement of the EQM system raises questions in terms of efficiency. While it was assessed very positively and it provided a highly visible improvement of service quality (interviews 11, 16, 18, 19, 20), the lack of an output indicator assessing the crucial aspect of improving taxpayer service quality means that an evaluation of cost-effectiveness is only possible to a limited extent. An analysis of EQM data reveals a considerable reduction of waiting time in the taxpayer offices since introduction of the system, indicating an increase in service quality (ZRA data generated by EQM system, see ZRA 2017). Yet, given the achieve-

							ment of the relevant module indicator already, the resources might have been able to achieve more results in another activity or action area.
	4. Were the output/resource ratio and alternatives carefully considered during the design and implementation process – and if so, how?	1. The intervention team is able to provide examples for considering the output / resource ratio and alternatives	/	Interviews with intervention staff	Analysis	Limited, difficult to verify	a peer-learning visit of the South African Treasury was organized, resulting in the identification of key improvements for cash management with little resources spent (interviews 14, 21, 37). Comparing this activity with alternative options such as an exchange visit of relevant MoF staff in South Africa or conducting a series of training, the peer-learning visit was assessed highly cost-effective, both in terms of the little resources spent (mainly the travel cost of the two visitors) and the results achieved (15 key MoF personnel gaining first-hand experience through peer exchange) (ibid.).
	5. For interim evaluations based on the analysis to date: To what extent are further planned expenditures meaningfully distributed among the targeted outputs?	1. Qualitative assessment of the plausibility of distibution of planned costs to units / outputs	Current cost commitment report	Interview with head of programme, heads of unit and head of administation and finance	Analysis	Evaluators' assessment, possible subjectivity bias mitigated by researcher triangulation	area 3: Given the current gap between achievements until now and expected output indicators, comparatively more resources have been allocated for achieving the remaining output indicators in this action area until the end of the intervention. In this respect, the training on risk management was purposefully postponed until the risk management unit in MoF is established, which is expected for the first quarter of 2018.
The project's use of resources is appropriate with regard to achieving the TC-measures' goal (outcome). [Allocation efficiency: Re-sources/Services in accordance with the BMZ]	6. To what extent could the outcome have been maximised with the same amount of resources and the same or better quality (maximum principle)?	1. Qualitative assessment of possibilities for yield maximization on outcome level		Interview with head of programme, heads of unit and head of administation and finance	analysis	Evaluators' assessment, possible subjectivity bias mitigated by researcher triangulation	the overarching costs of 19% are high by GIZ standards, resulting in an administrative burden that could potentially be used to maximise results. Looking into the figures in more details, the tool reveals that particu-larly staff costs are high in this regard. 50% of national and 16% of international human resources as well as 59% of local consultant personal are booked as overarching costs. A likely explanation for these com-paratively high costs is the additional administrative effort required to manage and implement co-financing from DFID, which was confirmed by the project staff (interviews 3, 15). The evaluation team deems it im-portant to shed light on such costs necessary to leverage cofinancing, which needs to be taken into ac-count for potential cofinancing plans in the future.
	7. Were the outcome/resources ratio and alternatives carefully considered during the conception and implementation process – and if so, how? Were any scaling-up options considered?	1. The intervention team is able to provide examples for considering the outcome / resource ratio and alternatives	/	Interviews with intervention staff	Analysis	Limited, difficult to verify	An alternative option for the project to increase effectiveness as well as efficiency is the rethinking of the human resource concept. Currently, each action area is covered by at least one international advisor, who is embedded into the partner structure almost like a CIM-expert. While it was successful at establishing good working relationships with specific stakeholders within the partner structure, it is a relatively resource intensive approach. For the future, the project is considering a more cluster-based approach with the project´s advisors covering larger areas within the partner structure, which can be complemented with specific consultancies if needed. Especially in light of the GIZ efforts to raise capacities of regional experts in the framework of the project Regional Support to Good Financial Governance in Southern Africa (see Raetzell 2015), a shift towards international advisors in a more coordinating function and an increased use of national or regional experts might be able to achieve even more results.
	8. To what extent was more impact achieved through synergies and/or leverage of more resources, with the help of other bilateral and multilateral donors and organisations? If so, was the relationship between costs and results appropriate?	1. description of synergies/ leverage achieved through DFID cofinancing 2. narrative assessment of synergies/ leverage achieved through DFID cofinancing 3. qualitative assessment of the cost-benefit ratio of the DFID cofinancing	progress reports to BMZ, annual review to DFID	Interviews with intervention staff	analysis	Evaluators' assessment, possible subjectivity bias mitigated by data and researcher triangulation	Regarding synergies with other funding sources, leverages could be achieved with the Worldbank PFM reform program and by taking part in two German PPP projects in action area I, but they could have partly been exploited further. In the following, each of these aspects will be examined separately. Firstly, the ongoing PFM reform program implemented by Worldbank offered considerable room for syner-gies, which were mostly but not fully exploited. It was reported that initially activities were planned together with the German Development Bank (KfW) (interviews 15, 38) and the project documents make clear that both objectives as well as action areas of the project evaluated here and the Worldbank programme are clearly linked (e.g. the GIZ project working on five out of ten components of the PFM reform programme). Also, examples for leveraging each others activities and results were provided. For example, Worldbank financed follow-up training after GIZ technical advice to internal audit department (interview 25) and consul-tancies contributing to the achievement of project results were financed with Worldbank funding (interview 12). Over the course of the project term, however the evaluation team observed that tensions between between Worldbank and GIZ objectives emerged: While the timely roll-out of IFMIS to all MPSAs is a key objective of Worldbank activities, which is also reflected in output indicator B1 of the this project proposal, the project adapted its strategy to first address underlying issues, because it noted that meeting the stated quantitative indicators alone will not bring about the expected results. Therefore, the project started to first analyse and map underlying business processes in key areas of budget execution (interview 14) and to work on reconfiguring IFMIS to better meet specific needs such as OBB functionality (interview 12).

							While this adaptation appears successful at better leveraging each other´s results, changes were not implement-ed in a coordinated way with Worldbank on the management level (interviews 15, 34). This may be ex-plained by the conflict of interest between the two players (see relevance chapter), but nevertheless indi-cates that a better alignment from the start could have achieved even more. Secondly, the involvement in two PPP projects in action area I helped leverage the resources towards in-creasing the revenue side of public finances (interviews 11, 16. 28). In particular, the PPP project ZAMeT, which aims to enable SME to register as tax payers and file returns using simple mobile phones to simplify tax compliance was launched successfully in 2015, is planned to be rolled out further from 2019 with addi-tional external PPP funding.

Sustainability							
Evaluation dimension	**Analysis questions**	**Evaluation indicators / descriptors**	**Available data sources**	**Other planned data collection projects**	**Evaluation strategy (evaluation design, method, procedure)**	**Expected evidence strength (narrative)**	**assessment by evaluation team**

Prerequisite for ensuring the long-term success of the project: results are an-chored in (partner) structures	1. What has the project done to ensure that the intended effect can be achieved in the medium to long term by the partners themselves (working aid review)?	1. The project works with the partner in a participatory approach 2. partner structures share the vision & objectives of the project	annual progress report	Interviews with partners and intervention staff, document analysis of potential reports and strategy documents by partners	Contribution analysis	fair	In action area I, the project anchored its results in the SME taxation strategy, in the audit training manual and in the physical infrastructure of the EQM system. Strategic advice was institutionalised in the SME taxa-tion strategy, which forms the basis for the future work of ZRA to increase SME taxpayer compliance (inter-views 17, 18, 20). Also the support of tax audit functions through training and accompanying manuals as well as the development of sector notes for specific business sectors provided a basis for ZRA tax inspec-tors to do their work more effectively. The results of the training were used to update ZRA´s internal training manuals, making their long-term application likely (interviews 18, 20). The EQM system, which was pro-cured for three taxpayer offices through the project, is expected to help ZRA to effectively manage their service points for many years beyond the project. Partner staff was trained on its use and the analysis of data generated by the system, hence making its long-term use likely (interview 19). Finally, an important observation was made by the project in the beginning, which was considered crucial for both effectiveness and sustainable impact of the project. It was found that the taxpayer data base as the foundation for effec-tive tax collection was inaccurate, providing false information on some business, containing various organi-sations no longer operating and omitting a large part of active taxpayers. Without an accurate taxpayer database, it was reported, it will not be possible to institutionalize higher quality tax audits (interview 11). In consequence, the project financed the recruitment of ten short-term staff for ZRA with the goal of cleaning-up and updating the database. While this is a one-time activity that needs to be follow-up and continuously improved to have an impact (interview 19), it can be seen as a plausible first step to showcase the potential of a solid taxpayer database for more effective tax revenue generation in the future. In action area II.a, the senior advisor made sure to consolidate achievements of technical advisory services through legislative or procedural measures. On the legal side, the project provided expert advice to the partner on various different reform bills including the excess expenditure bill and the long-awaited budget and planning bill (interviews 12, 22). This resulted in the institutionalization of critical reform initiatives in the adopted and expected law. While there is always the risk of a discrepancy between formalized law and its application, these legal measures are nevertheless a necessary starting and reference point for reform toward more transparent and credible budget planning in the long run (interview 12). With regards to proce-dural measures to anchor project activities in the partner structures, the project focused on institutionalizing suitable cooperation formats within the government to foster integrated budgeting processes involving all relevant stakeholders. In this respect, cluster hearings with groups of MPSAs there are jointly chaired by MoF with MNDP were introduced with the objective to institutionalize aligned planning and budgeting pro-cesses. As a supporting tool the budget challenge function was established and accompanied by a manual for budget office to help budget analysis in scrutinizing budget proposals coming from the spending agen-cies (interviews 3, 22). Moreover, the supplementary budget and excess expenditure process was re-engineered to deepen the Cabinet`s understanding of the consequences of in-year political decisions and pre-Cabinet meetings were initiated to discuss budgetary matters separately before full Cabinet meetings. These initiatives sought to institutionalize the political ownership of the budget (interview 12). Finally, the piloting of OBB in three ministries and of a warrant system to replace cash rationing are measures that aim to foster more effective, efficient and transparent approaches to budget planning in the future. Again, the advisory work was accompanied by the formulation of user manuals and training (ibid.). Particularly in light of the challenging political situation, there are risks that even adopted laws are not followed as envisioned and that the established procedures are not applied in the long run (ibid.). Nevertheless, the operational partner overall assesses the sustainability of results as high (interview 22). In action area II.b, the project was only partly successful in providing the newly established MNDP with methods and tools to operationalize strategic plans in order to foster sustainable change. While data cata-logues could be established together as a first step, the project did not manage to implement the MIS as planned (interview 13, 23). Given the slow progress in terms of developing the MIS as the targeted IT sys-tem, MNDP then decided to rely on the Monthly Monitoring System (MMS), which was developed with sup-port of the previous project phase as a predecessor to the MIS (interview 24). Until the end of 2018 the goal is to enhance MMS as much as possible, but it was reported that the system will not be able to meet the monitoring requirements set by the president and by the project (interview 13). Especially in light of the sup-port from this project running out at the end of this year, there is a large risk of the initiated processes not being able to continue. This risk is exacerbated by the fact that human and organizational resources are insufficient at present to absorb advisory results (interviews 13, 23). The lack of a project management unit, which was approved by the government but never installed is a case in point. While interview partners with-in the government partly see the lack of automation as the main challenge and therefore ask for IT solutions (interview 24), the aspects of limited capacity and resources are considered more critical by the project and external stakeholders (interviews 13, 30, 35). Finally in action area III, results are mostly anchored in the partner structure, for example through training manuals and in a peer learning process initiated with the South African Treasury. Capacity of the internal audit department is enhanced sustainably by incorporating a quality assurance system into the daily work and by fostering both internal peer review as well as an external review after five years (interviews 25, 26). Internal audit manuals had been updated and their contents are gradually disseminated to internal auditors in the MPSAs through training. The work on risk management has not started, since the relevant risk man-agement unit has not yet been established. Yet, by planning to anchor the training work in the future risk management unit, the project seeks to ensure the institutionalisation of risk management beyond the project term (interviews 14, 26). Also the peer-learning visit organized with the South African Treasury to share and gain experiences can be assessed very positively in this regard, as it created ownership of the reform pro-cess and allowed the partner define its own priorities in improving the functionality of the TSA (interviews 14, 21, 37).

	2. Which advisory contents, approaches, methods and concepts of the project are anchored/institutionalised in the (partner) system?	1. Description of contents, approaches, methods, concepts developed within the intervention a. used by the partners b. not used by the partners	annual progress report	Interviews with partners and intervention staff, document analysis of potential reports and strategy documents by partners	Contribution analysis	fair	In addition to the good prerequisites for long-term success, there are many examples of partner structures owning the tools and approaches developed with the project. In the following, these will be discussed for each action area separately. Action area I: • ZRA was reported to have a sophisticated project management unit that coordinates all interna-tional projects and follows up on common milestones in a proactive way (interviews 11, 18). • It is planned to continue the training on SME tax audit further without project support, which high-lights the interest ZRA is having to follow-up on the results of the project (interviews 19, 27). Action area II.a: • The project is attempting to increase political ownership of the budget by supporting the budget of-fice in re-designing the supplementary budget and excess expenditure process and briefing Cabi-net before their meetings on budgetary implications of in-year changes (interviews 12, 22). Action area II.b: • The cluster hearings, chaired jointly by the M&E department with the MoF budget office to chal-lenge the budget proposals of MPSAs institutionalize common budget planning procedures (inter-views 23, 30). Action area III: • The peer-learning process with the Republic of South Africa that was initiated with a first Treasury visit in 2017 and is planned to continue helped put the Accountant General´s office into the position of identifying reform challenges and learning from a neighbour's experience without external inter-ference (interviews 14, 21).
	3. To what extent are these continuously used and/or further developed by the target group and/or implementing partners?	1. Description of contents, approaches, methods, concepts developed within the intervention a. further developed by the partners b. not further developed by the partners	annual progress report	Interviews with partners and intervention staff, document analysis of potential reports and strategy documents by partners	Contribution analysis	fair	
	4. To what extent are (organisational, personnel, financial, economic) resources and capacities in the partner country (longer-term) available to ensure the continuation of the results achieved?	1. qualitative assessment of organizational resources in partner institutions 2. qualitative assessment of human resources of partner institutions 3. qualitative assessment of financial resources of partner institutions	annual progress report	Interviews with partners, external stakeholders and intervention staff	Analysis	limited	challenges exist: partly insufficient human and organizational resources, especially in planning ministry (interviews 12, 23) and for internal audit and risk management - many staff in acting positions (interview 25).
	5. To what extent are national structures and accountability mechanisms in place to support the results achieved (e.g. for the implementation and review of Agenda 2030)? o What is the project's exit strategy? o How are lessons learnt prepared and documented?						Analysing the political dimension further, the influence the project has on these challenges is limited. Nei-ther the domestic political debates nor the geopolitical trends, are developments that the project with its highly technical advisory services can address. But while the project cannot be expected to offset the aforementioned risks, it was reported that it lacks the normative counterbalance as envisioned in the BMZ Good Financial Governance Strategy (BMZ 2014; interviews 28, 33, 36). At the planning stage this tech-nical focus was considered an appropriate follow-up to the predecessor phase (see relevance chapter). Yet, in light of the increasingly challenging political situation, there mere technical focus of the project may not be enough to address PFM reform sustainably (interviews 33, 36).
Are the results of the project ecologically, socially and economically balanced?	6. Evaluation of the outcome results with regard to interactions between the environmental, social and economic dimensions of sustainability	1. Narrative assessment of three dimensions of sustainability	Progress reports	Interviews with intervention staff, external actors and partners	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	Finally, with regards to the three dimensions of sustainability (economic, social and ecological), the project successfully works towards economic sustainability, but is only partly addressing the other dimensions. While the limited role of social and particularly ecological aspects in a good financial governance project is obvious, there are some areas of interactions, which provide opportunities or risks for both dimensions. In the following, the interactions with regards to each dimension will be discussed. Most prominently, the project is fostering economic sustainability, since the project works at the governance level with central government institutions on public financial management. The economic dimension is clearly addressed by supporting the government in increasing domestic revenue generation and in bringing budget execution more in line with national development plans and the planned budget, hence making public finances more sustainable as stated in the overall programme objective of the DC-measure (see GIZ 2016). This aspect is reflected in the expected results in all action areas (see effectiveness chapter).
	7. Which positive or negative intended and unintended re-sults (economic, social, ecological) does the project produce? (Assign intended and unintended results from the effectiveness evaluation	1. Description of unintended results with regards to three sustainability dimensions a.) Positive b.) Negative	progress reports	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	

	to the three sustainability dimensions)						
	8. Is there any identifiable tension between the ecological, economic and social dimensions? o Economically: Impairment of competitiveness, employability, etc. o Socially: How should the impact be assessed in terms of distributive results, non-discrimination and universal access to social services and social security systems? To what extent can particularly dis-advantaged population groups benefit from the results or have negative results for particularly dis-advantaged population groups been created? o Ecologically: What are the positive or negative environmental impacts of the project?	qualitative assessment of tensions between sustainability dimensions on the impact level		Analysis of academic literature, Interviews with intervention staff, external actors and partners	Contribution analysis	Limited as causal chains to achieve impacts at level of final beneficaires are quite long	
	9. If negative interactions have been avoided and synergies exploited, how was this ensured? What measures were taken?	1. Description of measures taken by the intervention to avoid negative interactions 2. Description of measures taken by the intervention to create synergies	progress reports	Interviews with intervention staff	Contribution analysis	fair	none identified.

Forecast of durability: Results of the project are permanent, stable and long-term resilient	10. To what extent are the results of the project durable, stable and resilient in the longer-term under the given conditions?	1. Qualitattve assesssment of aspects that foster or hinder sustainability a.) Ownership of the partners for services or tools developed within the intervention b.) human resources available for building upon results achieved d.) Political will in the partner institutions	progress reports	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	Nevertheless, the forecast of durability is limited by inhibiting factors on the operational and on the political level. In the following, these challenges on the operational and on the political level are discussed separate-ly. On the operational level, crucial factors inhibiting the durability of results related to both the partner and the project. Firstly, the available resources of the government are often not sufficient to implement reform pro-cesses (interviews 23, 25, 35). The human resource situation is particularly difficult in MNDP, which two years after its establishment is still heavily understaffed (interviews 13, 24, 30). Also in MoF staff turnover was observed. Here, the six internal auditors interviewed were currently in acting positions, and a risk man-agement unit, which was expected in 2017, is still not in place (14, 25). Secondly, the long-standing close-ness of advisors to the partner, despite its merits, creates a personal dependency. Hence, there is a risk that the role of the advisors will be hard to fill once the project is over (interview 36). And thirdly, several interview partners stated that the Zambian government is requesting IT systems that the structures are not ready for, which results in a high risk of losing gains after the end of the project. This is particularly the case in action area II.b with the delays in MIS implementation. But to a lesser extent also action areas II.a and III are affected in view of the IFMIS module, which after years of international support is still not effectively functioning (interviews 12, 14, 33). Also on the political level, risks are perceived as high for durable and resilient results in light of the overall difficult PFM framework (interviews 28, 33, 36). This can be explained by the following factors. Firstly, aca-demic research describes the political situation as worsening with corruption concerns growing and pres-sure on civil society, media and the opposition increasing (Beardsworth 2017). Secondly, the public debt remains to be high (IMF 2017). While the Zambian government has taken up talks with IMF again, there are no signs that it plans to reverse its overall debt strategy given the announcement of further loans and the continued unclarity over the total debt (interviews 34, 38). Thirdly, the political will for a more profound re-form process is perceived to be limited. As was elaborated in the relevance chapter, there is certainly a formal commitment to reform on the political level and various actors on the technical level are pushing for reform. At the same time, a more profound political will for real change may be lacking, as is indicated by the gap between written law and actual practice (interviews 8, 28, 29, 33). Lastly, the weight of Western cooperation partners is reported to be decreasing. More concretely, it was reported that a general shift of Zambia`s bilateral cooperation is taking place towards emerging players, most importantly China, while Western cooperation partners and their conditionality approach are gradually playing less of a role (inter-views 8, 28, 36).
	11. What risks and potential are emerging for the long-term protection of the results and how likely are these factors to occur? o What has the project done to reduce these risks and exploit potential?	1. Description of risks potentially affecting sustainability 2. Assessment of extent to which intervention can influence risks 3. Description of mitigation strategies adoped	progress reports	Interviews with intervention staff, partners and external actors	Contribution analysis	Fair	The current challenges and risks are expected to remain in place and possibly aggravate in the future. While the project cannot be expected to offset these risks, it can aim to support a normative counterbalance (interviews 28, 33, 36). Mitigation strategies were developed for the next phase (interview 15).

Annex 2: List of resources

Bibliography of sources used for / cited in the final report of the Central Project Evaluation "Strengthening Good Financial Governance in Zambia"

Internal sources

GIZ (2014a). Final report on TC-measure „Strengthening Good Financial Governance in Zambia" (PN 2011.2114.4).

GIZ (2014b). Indicators – a working aid, Bonn

GIZ (2015a). Project evaluation of TC-measure „Strengthening Good Financial Governance in Zambia" (PN 2011.2114.4).

GIZ (2015b). Project offer of TC-measure "Strengthening Good Financial Governance in Zambia" (PN 2014.2075.1).

GIZ (2016). Proposal of DC-programme „Good Financial Governance".

GIZ (2016b). Progress report on TC-measure "Strengthening Good Financial Governance in Zambia" (PN 2014.2075.1).

GIZ (2017a). Terms of Reference for the Central Project Evaluation of TC-measure "Strengthening Good Financial Governance".

GIZ (2017b). Progress report on TC-measure "Strengthening Good Financial Governance in Zambia" (PN 2014.2075.1).

GIZ (2017c). Concept note for a TC-measure. Follow up project to "Strengthening Good Financial Governance in Zambia" (PN 2014.2075.1).

External sources

Auditor General (2017). Report of the Auditor General on the Accounts of the Republic – For the Year ended 31st December, 2016, Office of the Auditor General, Lusaka, Zambia

Beardsworth, N (2017) ZAMBIA: Politico-Economic Short Analysis (PSA), The GIGA German Institute of Global and Area Studies, Hamburg

Bird A. et al. (2016). Zambia – Review of Output Based Budgeting System, International Monetary Fund – Fiscal Affairs Department; AFRITAC South, Washington D.C.

BMZ (undated). Delegated Cooperation Agreement with DFID on the Final Design and Implementation of a Management Information System for the Government of Zambia.

BMZ (2014). Good Financial Governance in German Development Cooperation – Promoting Good Governance in Public Finance, BMZ Sector Strategy Paper 4 | 2014e, Bonn.

BMZ (2016). Country strategy with Cooperation Country Zambia (2016 – 2021), Bonn.

BMZ (2017). Die digitale Revolution für nachhaltige Entwicklung nutzen – Die Digitale Agenda des BMZ, Referat Bildung und digitale Welt, Bonn

Chaponda, T. et al. (2017). Zambia – Public Investment Management Assessment, Technical Assistance Report, IMF Fiscal Affairs Department, Washington, D.C.

DfID (2017). Public Sector Performance (PSP) Programme, Annual Review Fact Sheet

GRZ (2006). Vision 2030 – "A prosperous Middle-income Nation by 2030", Lusaka, Zambia

GRZ (2013). Public Financial Management Reform Strategy for the Government of the Republic of Zambia

GRZ (2016). Zambia's Vision 2030 National Performance Framework: Strategic Objectives, Long-Term Outcomes, Key Performance Indicators, Pathways to Change, and National Measurement Indicators 2017 – 2030, Lusaka, Zambia

GRZ (2017a). Zambia`s Vision 2030 National Performance Framework. GRZ (2017b). 2018 Budget address by the Minister of Finance Hon. Felix C. Mutati, MP. IMF (2014). Country Report 14/5.

IMF (2017). 2017 Article IV Consultation – Press release; staff report; and statement by the executive director for Zambia, IMF Country Report No. 17/327, Washington, D.C.

IMF (2017b). Staff report on the Article IV Consultations – debt sustainability analysis.

Mayne, J. (2001). Addressing attribution through contribution analysis: using performance measures sensibly. Canadian Journal of Programme Evaluation, 16, 1-24.

MNDP (2017). 7th National Development Plan 2017-2021, Ministry of National Development Planning, Lusaka, Zambia

MoF (2015). Risk Management Framework for Government of Zambia,Lusaka, Zambia

MoF (2016). Shifting from "Cash Rationing" to a "Commitment Limit" System to Facilitatet MPSA Expenditure – A Necessary but not Sufficient Element in the Medium Term Strategy to Restore Budget Credibility, Budget Office/Office of the Account General, Lusaka, Zambia

MoF (2017a). Economic Stabilization and Growth Programme – Zambia Plus.

MoF (2017b). Zambia's Debt Management Strategy 2017-2019.

MoF (2017c). 2018 Budget Address by Minister of Finance Felix Mutati on 29 September 2017.

MoF (2017d). Annual Economic Report 2016.

Mwenge, F. & Masumbu, G. (2016). Recounting the Miseries of the Poor – A Multidimensional Measurement of Poverty in Zambia, ZIPAR Discussion Paper No.2, Zambia

Nalishebo, S. & Halwampa, A. (2014) Uncovering the Unknown – An Analysis of Tax Evasion in Zambia, ZIPAR Working Paper No. 18, Zambia

Neumann, C. (2017). Mission Report 13.02.-24.02.2017 Zambia, Trainings on Indirect Methods, Augsburg

PEFA (2017). Report on the Evaluation on the Public Financial Management System of Zambia, Assessment Report of the Republic of Zambia, PEFA Secretariat, Zambia

Raetzell, L. (2015). Mid-Term Evaluation of HCD Measure "Executive Master's in Taxation", Syspons GmbH, Berlin

Rojas, E.et al. (2016). TADAT Performance Assessment Report – Zambia, Zambia

Stern, E. et al. (2012). Broadening the range of designs and methods for impact evaluations. Working Paper 38, Department for International Development., London

United Nations (2015). Agenda 2030 and the Sustainable Development Goals.

ZIPAR (2017). Analysis of the 2018 National Budget, Lusaka, Zambia

United Nations (2015). Agenda 2030 and the Sustainable Development Goals.

ZIPAR (2016a). Analytical Brief of the Social Sector Budget 2017 " Leaving no one behind", Zambia

ZIPAR (2016b) Walking a Tight Rope to Economic Recovery – Analysis of the 2017 Budget Speech "Restoring Fiscal Fitness for Sustained Inclusive Growth and Development", Zambia

ZIPAR (2017) Creating and Wasting Fiscal Space – Zambian Fiscal Performance 2002 -2011, Lusaka, Zambia

ZIPAR (2017) Staying the Course of Fiscal Fitness – Analysis of the 2018 National Budget " Accelerating Fiscal Fitness for Sustained Inclusive Growth, Without Leaving Anyone Behind", Zambia

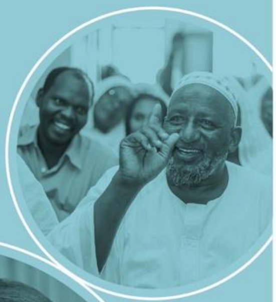
ZRA (2017) eQMS Relevance in Improving Service Delivery at ZRA-Report, Zambia

Annex 3: Terms of reference

Stabsstelle Evaluierung

Zentrale Projektevaluierung

WISSEN WAS WIRKT



Leistungsbeschreibung

Zentrale Projektevaluierung des Vorhabens

„Förderung von guter finanzieller Regierungsführung in Sambia / Good Financial Governance in Zambia“

Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Inhalt

Abkürzungsverzeichnis

AN	Auftragnehmer
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
DeGEval	Gesellschaft für Evaluation
EU	Europäische Union
FKT	Fachkrafttage
GA	Gutachter/-in
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GVR	Gemeinsamen Verfahrensreform
OECD-DAC	Organisation for Economic Cooperation and Development - Development Assistance Committee
PN	Projektnummer

1. Zentrale Projektevaluierungen der GIZ

1.1 Hintergrund und Zielsetzung

Das Evaluierungssystem der GIZ sieht sich mit einer Reihe neuer Herausforderungen konfrontiert: die zunehmende Diversifizierung von Auftrags- und Vorhabenstypen, die zunehmende Komplexität von Durchführungskontexten und Vorhaben, neue Erkenntnisinteressen der Politik (kurzfristige Ergebniserzielung, andere Evaluierungskriterien u. ä.). Hinzu kommen die neuen Evaluierungsanforderungen, die sich aus der Agenda 2030 für Nachhaltige Entwicklung und aus der Gemeinsamen Verfahrensreform (GVR) im Auftragsverfahren mit dem BMZ ergeben. Auch die Anforderungen an die Nutzung von Evaluierungen der GIZ haben sich verändert. „Lernen aus Evaluierungen" bleibt nach wie vor eine wichtige Funktion von Evaluierungen. Die Aufgabe besteht dabei vor allem darin, das von Evaluierungen generierte Wissen zielgerichtet für Entscheidungsfindungsprozesse aufzubereiten. Zum anderen haben sich die Anforderungen an Rechenschaftslegung – und somit an die Qualität und Unabhängigkeit von Evaluierungen und Evaluierungsberichten – in den letzten Jahren deutlich verschärft. Vor diesem Hintergrund hat der Vorstand der GIZ im Dezember 2016 entschieden, das Evaluierungssystem der GIZ grundlegend zu reformieren. Ziele dieser Reform sind vor allem folgende:

- **Wirkungsnachweis verbessern:** Das neue Evaluierungssystem soll die GIZ besser in die Lage versetzen, längerfristige Wirkungen sowie die Nachhaltigkeit und Verankerung der Ansätze in den Partnerstrukturen zu betrachten. Dazu sollen Evaluierungen zu einem Zeitpunkt erfolgen, zu dem Aussagen über Wirkungen und Nachhaltigkeit möglich und sinnvoll sind, und methodisch und prozedural so konzipiert sein, dass sie das leisten können.

- **Glaubwürdigkeit von Evaluierungsergebnissen erhöhen:** Wir wollen die Glaubwürdigkeit unserer Evaluierungsergebnisse weiter erhöhen, indem wir zum einen die Unabhängigkeit der Projektevaluierungen stärken. Die Projektevaluierungen werden daher zukünftig von der Stabsstelle Evaluierung gesteuert und verantwortet, die organisatorisch direkt dem Vorstand untersteht und vom operativen Geschäft getrennt ist. Die Durchführung erfolgt durch spezialisierte externe Evaluatorinnen und Evaluatoren. Die Evaluierungen werden nach national und international gültigen Standards und Qualitätskriterien durchgeführt und die Evaluierungsberichte werden veröffentlicht.

- **Projektevaluierungen auf neue Herausforderungen ausrichten:** Zentrale Projektevaluierungen sollen der zunehmenden Komplexität von Vorhaben und Durchführungskontexten, den erhöhten Anforderungen an Rechenschaftslegung sowie den evaluativen Herausforderungen aus der Agenda 2030 und der GVR Rechnung tragen.

1.2 Gestaltung der Durchführung des mehrjährigen Evaluierungsportfolios

Gegenstand der zentralen Projektevaluierungen sind regelmäßig Vorhaben, die die GIZ im Auftrag des BMZ durchführt. Zentrale Projektevaluierungen beinhalten den kritischen analytischen Rückblick auf die Wirkungen und die Implementierung eines Vorhabens. Sie können zu unterschiedlichen Zeitpunkten stattfinden. Vorhaben, die enden, werden rund 8 Monate nach Ende der meist 3-jährigen Laufzeit evaluiert (Schlussevaluierung). Vorhaben mit geplanten Folgemaßnahmen werden auch während der Laufzeit evaluiert (Zwischenevaluierung) – je nach „intended use" (Zulieferung für Planung des Folgeauftrags, Vorhabensteuerung, Berichterstattung an den Auftraggeber, Strategische Reflexion). Sowohl bei den Zwischen- als auch bei den Schlussevaluierungen werden - wenn inhaltlich sinnvoll - Vorgängervorhaben mitbetrachtet, um Aussagen über längerfristige Wirkungen und Nachhaltigkeit treffen zu können.
Im BMZ-Geschäft werden alle Vorhaben über 3,0 Mio. EUR Auftragswert in das Evaluierungsverfahren standardmäßig einbezogen. Die Auswahl der zu evaluierenden Vorhaben erfolgt in einem zweistufigen Verfahren. In einem ersten Schritt werden die zu evaluierenden Vorhaben über eine regional geschichtete Zufallsstichprobe ausgewählt. In einem zweiten Schritt wird die Stichprobe durch Evaluierungen ergänzt, die nach spezifischen Erkenntnisinteressen ausgewählt werden (kriterienbasierte Auswahl).

Insgesamt soll mittelfristig für die Grundgesamtheit aller Vorhaben über 3,0 Mio. EUR Auftragswert im BMZ-

Geschäft ein Deckungsgrad an Projektevaluierungen zwischen 30% und 50% gewährleistet werden. Im BMZ-Geschäft werden so langfristig jährlich ca. 100 zentrale Projektevaluierungen durchgeführt werden. Die Gesamtzahl der evaluierten Vorhaben soll groß genug sein, um eine repräsentative Aussage hinsichtlich der Bewertung der OECD/DAC-Kriterien aller Vorhaben der Grundgesamtheit treffen zu können.

Die Durchführung des ersten Evaluierungsportfolios wird EU-weit ausgeschrieben. Ziel sind Rahmenvereinbarungen, mit sektorfachlich und regional strukturierten Pools von Evaluatoren/-innen, die die Evaluierungen aus dieser Stichprobe bis 2020 durchführen. Da mit dem Abschluss des Vergabeverfahrens erst im zweiten Quartal 2018 gerechnet werden kann, werden die ersten Pilotevaluierungen für Ende 2017 und Anfang 2018 als einzelne Leistungen im Wettbewerb nach Short-list bzw. E-tender-Verfahren ausgeschrieben.

2. Gegenstand und Ziel der Evaluierung

2.1 Projektbeschreibung und Gegenstand der Evaluierung

In den ersten Jahren des neuen Jahrtausends war Sambia eine der am schnellsten wachsenden Volkswirtschaften weltweit. Das reale Wachstum des Bruttoinlandsprodukts (BIP) betrug durchschnittlich ungefähr 6,7% jährlich. Das starke Wachstum führte dazu, dass Sambia den Status eines Landes mit niedrigem, mittlerem Einkommen (*lower middle-income country*) erreicht hat. Trotzdem ist extreme Armut insbesondere in ländlichen Gegenden (77%) weit verbreitet und die Einkommensungleichverteilung (der Gini-Koeffizient betrug laut Weltbank 0,69 im Jahr 2015) sowie die Arbeitslosigkeit sind sehr hoch. Hohe Geburtenraten und hohe HIV/AIDS Raten verschärfen diese Probleme.

Bei Projektbeginn stellte das stark abgeschwächte Wachstum (Prognose 2016 3 - 3,5%) eine armuts-orientierte Politik vor zusätzliche Herausforderungen. Eine hohe Inflationsrate von über 20% bedrohte den bescheidenen Wohlstand großer Bevölkerungsschichten. Der Preisverfall von Kupfer, eine Reduzierung der Energielieferung an die Minengesellschaften und eine volatile Steuerpolitik führten zu einem spürbaren Rückgang des Bergbaus und zu verminderten Steuereinnahmen. Infolge der expansiven Ausgaben-Politik und zunehmenden Schwierigkeiten ihrer Finanzierung fehlte es der Regierung an Mitteln zur Abfederung des Abwärtstrends. Die sambische Regierung ist zudem bisher nur unzureichend in der Lage, ausreichend öffentliche Mittel zu generieren, den Haushalt entwicklungsorientiert zu planen und umzusetzen und die Ausgaben effizient und effektiv zu leisten.

Die finanzielle Regierungsführung hatte sich mit der Zeit verschlechtert. Das Budgetdefizit stieg 2013 auf 6,5% und blieb bei 6% im Jahr 2014. Für 2015 war ein Haushaltsdefizit von 6,9% prognostiziert. Der Schuldenstand von Sambia bewegt sich in Richtung von 40% des BIP. 2011 wurde ca. ein Drittel des Mittelvolumens des sambischen Staatshaushalts anders verausgabt als vom Gesetzgeber im Haushaltsgesetz vorgesehen war. Gemäß dem *Country Report* Nummer 14/5 des Internationalen Währungsfonds wichen nahezu alle Ausgabenkategorien von dem geplanten Ausgabenniveau substantiell ab. Politisch motivierte unterjährige Reallokationen von Finanzmitteln ohne vorherige Genehmigung durch das Parlament stellten keine Ausnahme dar. Der Haushalt in Sambia wich daher von gesetzlichen Vorgaben ab. Trotz des hohen Finanzbedarfs des Staates ist das Potenzial zur Besteuerung des informellen Sektors kaum erschlossen. Die Einnahmen aus der Besteuerung kleiner und mittlerer Unternehmen sowie der Beitrag der inländischen Mehrwertsteuer sind niedrig. Transparenz, Glaubwürdigkeit und Effektivität der öffentlichen Finanzen sind unzureichend (Kernproblem).

Das Modulziel lautet: Transparenz, Glaubwürdigkeit und Effektivität der öffentlichen Finanzen sind erhöht. Im Handlungsfeld i) Stärkung der Steuerverwaltung geht es darum die Steuerverwaltung in die Lage zu versetzen, die Steuerbefolgung der kleinen und mittleren Steuerzahler/- innen zu erhöhen. Das Handlungsfeld ii) Verbesserung der Haushaltsaufstellung zielt darauf ab, die Fähigkeit des Finanzministeriums, die Haushaltsvorbereitung gemäß der zukünftigen nationalen Planungs- und Haushaltsordnung umzusetzen, zu steigern. In Handlungsfeld iii) Verbesserung des Haushaltsvollzugs zielt das Vorhaben darauf ab, die Fähigkeit des Finanzministeriums, den Haushalt effektiv in Übereinstimmung mit dem Haushaltsgesetz auszuführen, zu erhöhen. Um dies zu erreichen, setzt das Vorhaben auf allen drei Ebenen des Capacity Development an.

Angetrieben von stark steigenden Kupferpreisen, welche sich aktuell auf dem höchsten Stand seit beinahe 15 Jahren befinden, hat sich die Realwirtschaft nach der schweren Wirtschafts- und Energiekrise 2015 / 2016 mittlerweile wieder weitestgehend erholt. Verbunden mit erhöhten Investitionen im Bergbausektor hat dies zu einer deutlichen Ausweitung der Produktion und einem steigenden Wechselkurs des Sambischen Kwacha geführt. Ausgiebige Regenfälle haben zudem zu guten Ernten und einer stabileren Stromversorgung beigetragen. Die Wachstumsprognosen haben sich daher zuletzt wieder verbessert und liegen aktuell bei 4,3%

für 2017. Auch die Inflation ist wieder auf deutlich unter 10% gesunken (6,8% im 1. Quartal 2017). In Erwartung eines in 2017 abzuschließenden IWF-Kreditvertrags scheint die wirtschaftliche Lage für Investoren aktuell stabil, die Situation der öffentlichen Finanzen ist jedoch unvermindert angespannt. Mit einer Entscheidung des IWF ist frühestens im September 2017 zu rechnen, ursprünglich anberaumt war das erste Quartal 2017. Der Haushalt für 2018 muss also spekulativ aufgestellt werden. Insgesamt sind rund 80% des Budgets für 2018 bereits für die Bedienung von Schulden (30%) sowie Lohnkosten der öffentlichen Bediensteten (50%) fest verplant. Trotz der bereits bedenklich hohen Staatsverschuldung geht die sambische Regierung weiterhin millionenschwere Verträge beispielsweise für die Beschaffung von Flugzeugen für eine nationale Fluggesellschaft ein. Somit besteht so gut wie kein finanzieller Spielraum zur Umsetzung von entwicklungsorientierten Politiken oder nachhaltigem Schuldenmanagement.

Gegenstand der Evaluierung ist das TZ-Modul (PN 2014.2075.1) mit der Laufzeit von 01/2016 bis 12/2018. Zusätzlich beinhaltet der Auftrag eine Kombifinanzierung durch das britische *Department for International Development* (DFID). In die Betrachtung einbezogen werden soll – sofern inhaltlich zweckmäßig – das Vorgängermodul PN 2011.2114.4 (12/2012-12/2015) um Aussagen über längerfristige Wirkungen und Nachhaltigkeit des Vorhabens treffen zu können.

2.2 Ziel der Evaluierung

Die Nutzungsorientierung ist eine wesentliche Säule der Evaluierungsarbeit in der GIZ. Die zentralen Projektevaluierungen betten sich darin ein und sollen Entscheidungsfindung unterstützen:

- Evaluierungsprozesse und -ergebnisse tragen dazu bei, die Entscheidungskompetenz von Entscheidungsträgern und weiteren Change Agents zu stärken;
- dies führt zu Entscheidungen, die entweder öffentliche Politiken oder die Ausgestaltung und Implementierung von GIZ-Vorhaben oder GIZ-Unternehmensstrategien verbessern;
- diese Verbesserungen führen wiederum zu einer verbesserten Leistungserbringung der Partner für ihre eigenen Bürger/-innen, der GIZ für ihre Partner und Zielgruppen sowie für ihre Auftraggeber und Mitarbeiter/-innen;
- dadurch wird schließlich die Wirksamkeit öffentlicher Politiken und GIZ-Vorhaben für die Zielgruppen sowie die Partner-, Kunden- und Mitarbeiterzufriedenheit erhöht.

Mit der Evaluierung soll der Erfolg des aktuellen Moduls (PN 2014.2075.1) bewertet werden. Dies soll auf der Grundlage der OECD-DAC-Kriterien – gestützt auf Zahlen, Daten und Fakten – und im Rahmen eines vorgegebenen Benotungssystems erfolgen. Darüber hinaus soll, wie unter 2.1. bereits angeführt, - sofern inhaltlich zweckmäßig - hinsichtlich Aussagen über längerfristige Wirkungen und Nachhaltigkeit des Vorhabens auch das Vorgängermodul mit betrachtet werden.

Da es sich hier um eine Zwischenevaluierung handelt, soll die Evaluierung darüber hinaus Hinweise für die Folgemaßnahmen erbringen, die in die Prüfung der Folgemaßnahme einfließen können.

In einem Auftaktgespräch des AN mit der Stabsstelle Evaluierung und dem/der AV des Vorhabens und ggf. Partner wird das Erkenntnisinteresse präzisiert und der Gegenstand der Evaluierung gemeinsam definiert.

3. **Prozess und Leistungen**

3.1 Verantwortlichkeiten

Verantwortlich für die Planung und Steuerung des Evaluierungsportfolios der zentralen Projektevaluierungen ist die Stabsstelle Evaluierung. Der Auftragnehmer (im Folgenden AN) ist verantwortlich für die Vorbereitung, Durchführung, Qualitätssicherung / Backstopping und Berichterstattung der Einzelevaluierung unter Berücksichtigung der unter 4. angeführten Leistungsanforderungen. Das Evaluierungsteam besteht immer aus einem Zweier-Team (internationale/r + lokale/r Gutachter/in). Die Auswahl und Einbindung des/der regionalen Gutachter/-in liegt in der Verantwortung des/der AN. Punktuell unterstützt die GIZ in den einzelnen Prozessschritten.

Die Unterstützung durch das Projekt oder Landesbüro vor Ort umfasst:

- Bereitstellung relevanter Dokumente
- Empfehlung für ein gut gelegenes Hotel
- Identifikation relevanter Interviewpersonen + Koordinierung / Erstellung des Interviewplans
- GIZ Fahrer des Projekts vor Ort wird gestellt

Der Ablauf der Evaluierung inkl. Rollenklärung ist der folgendes Prozessübersicht zu entnehmen. Das Prozesschart basiert auf den Erfahrungen der Stabsstelle Evaluierung mit dem Programm der Unabhängigen Evaluierungen sowie den dezentralen Projektevaluierungen und soll nun im Rahmen der zentralen Projektevaluierungen geprüft und bei Bedarf schrittweise angepasst werden. Hierfür ist zum Abschluss der Evaluierung eine gemeinsame Auswertung mit dem AN vorgesehen.

3.2 Prozessübersicht zentrale Projektevaluierung

Die u.a. Leistungen sind im Zeitraum vom 01.11.2017-31.03.2018 zu erbringen. Die Vor-Ort-Mission der Evaluierung wird in Sambia stattfinden.

Arbeitsschritt	Wann	Verantwortlich	Mitwirkend	Zu informieren
Vorklärungen inkl. Abstimmung Zeitpunkt der Evaluierung	Jul. 2017	Sts Eval.	AV, Partner	
Bereitstellung Dokumente	Bis KW 43 2017	AV/Projektteam (Projektdokumente), Sts Eval. (Evaluier-ungsdokumente)		
Auftragsklärung inkl. Rollenklärung GA-Team	Anfang KW 45 2017	Sts Eval.	Int. GA, lokal. GA	
Auftaktgespräch (bei Bedarf) zwecks Rollenklärung und Erkenntnisinteresse	Mitte/Ende KW 45 2017	Sts Eval.	AV, Partner, Int. GA, lokal GA	
Informationsschreiben zum Start der Evaluierung an zentrale Stakeholder (inkl. Informationen zu Prozess und Rollen)	KW 45 2017	Sts Eval.		AL, LD bzw. GL, AV, Partner, BMZ
Desk Study inkl. erster inhaltlicher Vorklärung in GIZ und bei Bedarf inkl. Vor-Ort-Check - Datenlage (u.a. WoM) - Partnersysteme - Erkenntnisinteresse der Partner	KW 45-47 2017	Int. GA/lok. GA	GIZ-MA	
Reisevorbereitung (z.T. erst nach IR möglich)	Nov./Dez. 2017	Int. GA	Lok GA, AV/ Projektteam,	

			(Landesbüro)	
Erstellung **Entwurf Inception Report (IR)** (entsprechend GIZ Vorgaben und Formatvorlage, Berichtssprache: Englisch)	Abgabe IR 27.11.2017	Int. GA	Lok. GA	
Qualitätsprüfung IR	Rückmeldung an AN: Bis 08.12.2017	Sts Eval.	AV/Partner (bzgl. sachl. Richtigkeit)	
Überarbeitung IR	KW 50 2017	Int. GA	(Lok. GA)	
Abnahme IR	KW 51 2017	Sts Eval.		BMZ
Erstellung und Abstimmung Interviewplan	KW 48-50 2017	Int. & lok. GA	AV, Partner	
Durchführung der Mission	KW 2- 4 2018 (Prüfteam ab 22.01. vor Ort)	Int. GA, Lok. GA		
Auftaktgespräche /Briefing vor Ort	Anfang KW 2 2018	Int. GA, Lok. GA	AV/Projektteam, LD, Partner, Botschaft	
Dokumentation der vorläufigen Ergebnisse für Abschlusspräsentation vor Ort (entsprechend GIZ-Vorlage)	KW 3 2018	Int. GA, Lok. GA		
Debriefing / Abschlussgespräche vor Ort	Ende KW 3 (ggf. Anfang KW 4 mit Prüfteam) 2018		AV/Projektteam, LD, Partner, Botschaft	
Auswertung, Analyse und Berichtslegung	KW 4-7 2018	Int. GA	Lok. GA	
Berichtslegung (entsprechend GIZ Vorgaben und Formatvorlage, Berichtssprache: Englisch)	Abgabe 19.02.2018	Int. GA	(Lok. GA)	
Qualitätsprüfung Evaluierungsbericht	Rückmeldung an AN: KW 11 2018	Sts Eval.	AV/Partner (bzgl. sachl. Richtigkeit)	
Überarbeitung Evaluierungsbericht (inkl. sprachl.- redaktioneller Qualitätssicherung)	KW 12 2018	Int. GA	(Lok. GA)	
Abnahme Evaluierungsbericht	KW 13 2018	Sts Eval.		
Abschlussgespräch per Skype (gemeinsame Auswertung der Evaluierung)	KW 13 2018	Sts Eval., Int. GA	(Lok. GA)	AV
Veröffentlichung des Evaluierungsberichts	April 2018	Sts Eval.		GA, AV, Partner

4. **Spezifische Leistungsanforderungen**

Die Leistungen sind, wie oben unter 3.2. bereits angeführt, im Zeitraum vom 01.11.2017 bis zum 31.03.2018 zu erbringen. Der Evaluierungsbericht ist bis zum 19.02.2018 in Englischer Sprache abzugeben, eine evtl. Überarbeitung auf Basis der Rückmeldung des AG muss bis Ende KW 12 2018 erfolgen (für den detaillierten Ablauf siehe die Prozessübersicht unter 3.2.).

4.1 Qualitätsanforderungen der zentralen Projektevaluierungen

Die GIZ orientiert sich bei ihren Evaluierungen an den Standards für Evaluation der Gesellschaft für Evaluation (DeGEval): Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit sowie den Qualitätsstandards für die Entwicklungsevaluierung des OECD-DAC. Als Basis für die Entwicklung von Instrumenten zur Qualitätssicherung formuliert die Stabsstelle Evaluierung Qualitätsstandards für die Prozessqualität, die methodische Qualität und die Produktqualität.

Die **Nützlichkeit** einer Evaluierung gewährleistet, dass das Erkenntnisinteresse der Nutzer/-innen der Evaluierung berücksichtigt wird und diesen die gewünschte Information bereitgestellt wird.

- Identifizierung der Beteiligten und Betroffenen: Die am Evaluierungsgegenstand beteiligten oder von ihm betroffenen Personen bzw. Personengruppen sollen identifiziert werden, damit deren Interessen geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können.
- Klärung der Evaluationszwecke: Es soll deutlich geklärt sein, welche Zwecke mit der Evaluierung verfolgt werden, so dass die Beteiligten und Betroffenen Position dazu beziehen können und das Evaluierungsteam einen klaren Arbeitsauftrag verfolgen kann.
- Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin: Wer Evaluierungen durchführt, soll persönlich glaubwürdig sowie methodisch und fachlich kompetent sein, damit bei den Evaluierungsergebnissen ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.
- Auswahl und Umfang der Informationen: Auswahl und Umfang der erfassten Informationen sollen die Behandlung der zu untersuchenden Fragestellungen zum Evaluierungsgegenstand ermöglichen und gleichzeitig den Informationsbedarf des Auftraggebers und anderer Adressaten und Adressatinnen berücksichtigen.
- Transparenz von Werten: Die Perspektiven und Annahmen der Beteiligten und Betroffenen, auf denen die Evaluierung und die Interpretation der Ergebnisse beruhen, sollen so beschrieben werden, dass die Grundlagen der Bewertungen klar ersichtlich sind.
- Vollständigkeit und Klarheit der Berichterstattung: Evaluierungsberichte sollen alle wesentlichen Informationen zur Verfügung stellen, leicht zu verstehen und nachvollziehbar sein.
- Rechtzeitigkeit der Evaluierung: Evaluierungsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungsprozesse bzw. Verbesserungsprozesse einfließen können.
- Nutzung und Nutzen der Evaluierung: Planung, Durchführung und Berichterstattung einer Evaluierung sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluierung aufmerksam zur Kenntnis zu nehmen und ihre Ergebnisse zu nutzen.

Die **Prozessqualität** entspricht den DeGEval Standards Durchführbarkeit und Fairness. Die Art und Weise, wie der Prozess einer Evaluierung gestaltet wird, ist für die Nutzung der Evaluierung entscheidend. Um die Evaluierung so nützlich wie möglich für Entscheidungsprozesse zu gestalten, sollten folgende Standards erfüllt werden:

- Angemessene Verfahren: Evaluierungsverfahren, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, sollen so gewählt werden, dass Belastungen des Evaluierungsgegenstandes bzw. der Beteiligten und Betroffenen in einem angemessenen Verhältnis zum erwarteten Nutzen der Evaluation stehen.
- Diplomatisches Vorgehen: Evaluierungen sollen so geplant und durchgeführt werden, dass eine möglichst hohe Akzeptanz der verschiedenen Beteiligten und Betroffenen in Bezug auf Vorgehen und Ergebnisse der Evaluation erreicht werden kann.
- Effizienz von Evaluierung: Der Aufwand für Evaluierung soll in einem angemessenen Verhältnis zum Nutzen der Evaluierung stehen.

- Formale Vereinbarungen: Die Pflichten der Vertragsparteien einer Evaluierung (was, wie, von wem, wann getan werden soll) sollen schriftlich festgehalten werden, damit die Parteien verpflichtet sind, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese neu auszuhandeln.
- Schutz individueller Rechte: Evaluierungen sollen so geplant und durchgeführt werden, dass Sicherheit, Würde und Rechte der in eine Evaluierung einbezogenen Personen geschützt werden.
- Vollständige und faire Überprüfung: Evaluierungen sollen die Stärken und die Schwächen des Evaluierungsgegenstandes möglichst vollständig und fair überprüfen und darstellen, so dass die Stärken weiter ausgebaut und die Schwachpunkte behandelt werden können.
- Unparteiische Durchführung und Berichterstattung: Die Evaluierung soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluierung deutlich machen. Berichte sollen ebenso wie der gesamte Evaluierungsprozess die unparteiische Position des Evaluierungsteams erkennen lassen. Bewertungen sollen fair und möglichst frei von persönlichen Gefühlen getroffen werden.
- Offenlegung der Ergebnisse: Die Evaluierungsergebnisse sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden.

Die **methodische Qualität** einer Evaluierung bezieht sich auf die Anwendung der Methoden der empirischen Sozialforschung zur Datenerhebung und -auswertung und entspricht dem DeGEval Kriterium der Genauigkeit.

- Beschreibung des Evaluierungsgegenstandes: Der Evaluierungsgegenstand soll klar und genau beschrieben und dokumentiert werden, so dass er eindeutig identifiziert werden kann.
- Kontextanalyse: Der Kontext des Evaluierungsgegenstandes soll ausreichend detailliert untersucht und analysiert werden.
- Beschreibung von Zwecken und Vorgehen: Gegenstand, Zwecke, Fragestellungen und Vorgehen der Evaluierung, einschließlich der angewandten Methoden, sollen genau dokumentiert und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.
- Angabe von Informationsquellen: Die im Rahmen einer Evaluierung genutzten Informationsquellen sollen hinreichend genau dokumentiert werden, damit die Verlässlichkeit und Angemessenheit der Informationen eingeschätzt werden kann.
- Valide und reliable Informationen[10]: Die Verfahren zur Gewinnung von Daten sollen so gewählt oder entwickelt und dann eingesetzt werden, dass die Zuverlässigkeit der gewonnenen Daten und ihre Gültigkeit bezogen auf die Beantwortung der Evaluierungsfragestellungen nach fachlichen Maßstäben sichergestellt sind. Die fachlichen Maßstäbe sollen sich an den Gütekriterien der empirischen Sozialforschung orientieren.
- Systematische Fehlerprüfung: Die in einer Evaluierung gesammelten, aufbereiteten, analysierten und präsentierten Informationen sollen systematisch auf Fehler geprüft werden.
- Analyse qualitativer und quantitativer Informationen: Qualitative und quantitative Informationen einer Evaluierung sollen nach fachlichen Maßstäben angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluierung effektiv beantwortet werden können.
- Begründete Schlussfolgerungen: Die in einer Evaluierung gezogenen Folgerungen sollen aus Ergebnissen hergeleitet werden, damit die Adressatinnen und Adressaten diese nachvollziehen können.

4.2 Profil der Gutachter/-innen

- Evaluierungserfahrung
- Erfahrung mit komplexen Evaluierungsdesigns
- Sozialwissenschaftliche Forschungsmethoden (quant./qualitat. und partizipative Methoden)
- Sektorkenntnisse und -erfahrung: Good Financial Governance (Gute finanzielle Regierungsführung) und Public Financial Management (Management der öffentlichen Finanzen) (idealerweise auch Praxiserfahrung), Fach-, Organisations- und Strategieberatung im Rahmen von Reformprozessen, Kenntnisse von EU- und internationalen Standards im Bereich Finanzpolitik und –verwaltung, Rechenschaftsbeziehungen sowie politischen Dialog- und Stakeholderprozessen,
- Hohe politische Sensibilität
- Erfahrungen mit der GIZ
- Landeskenntnisse in Sambia

[10] D. h. gesicherte und verlässliche Informationen

- Erfahrung mit Anwendung / Bewertung der OECD-DAC Kriterien
- Sprachkenntnisse in Englisch (internationale/r und lokale/r/regionale/r Gutachter/-innen) und Deutsch (internationale/r Gutachter/-in auf) verhandlungssicherem Niveau setzen wir voraus und fließt deswegen nicht in die Bewertung ein.

Wie oben bereits angeführt, soll die Evaluierung durch ein Gutachterteam (international/regional) durchgeführt werden. Für den/die regionale/n Gutachter/-in gelten dieselben, o.a. Profilanforderungen bis auf die Deutschkenntnisse. Die einzelnen Anforderungen werden jedoch unterschiedlich gewichtet (siehe hierzu das Bewertungsschema). Wie die Zusammenarbeit und Arbeitsteilung erfolgen soll, ist im Angebot darzulegen (siehe hierzu die Vorgaben unter 5. Umfang und Inhalt des abzugebenden Angebots). Beide Gutachter/-innen dürfen aus Gründen der Unabhängigkeit weder an der Konzeption, Planung, Implementierung, Beratung oder Evaluierung des Vorhabens beteiligt gewesen sein.

4.3 Methodisches Vorgehen

Bei den zentralen Projektevaluierungen genügt es in der Regel, als Grundlage einer glaubwürdigen Rechenschaftslegung den Beitrag des betrachteten Vorhabens zur Zielerreichung *(Kontribution)* möglichst robust zu erfassen. Es geht darum, eine plausible Beziehung zwischen Vorhaben und Wirkungen aufzuzeigen, also durch methodische und Datentriangulation hinreichende Belege dafür zu sammeln, dass die beobachteten intendierten Wirkungen höchstwahrscheinlich auf das Vorhaben zurückzuführen sind. Neben der Erfassung des Projektbeitrags sollen das Verständnis und das Wissen darüber erhöht werden, *was funktioniert und was nicht*, um auf dieser Basis fundierte Entscheidungen über die zukünftige Ausrichtung des Vorhabens treffen zu können.

Um in den zentralen Projektevaluierungen einen robusten Wirkungsnachweis zu ermöglichen, gibt die GIZ einen theoriebasierten Evaluierungsansatz vor. Theoriebasierte Ansätze wie beispielsweise Realist Evaluation, Process Tracing und die Kontributionsanalyse zeichnen sich durch folgende methodischen Elemente aus:

- Ein *Wirkungsmodell*, das bei der GIZ im Projektvorschlag enthalten ist und Erwartungen an die Ursache-Wirkungs-Beziehungen des Vorhabens visualisiert sowie Pfade von den Inputs über Aktivitäten und Outputs hin zu den angestrebten Wirkungen aufzeigt.
- Eine auf dem Wirkungsmodell basierende *Theory of Change*, die *Wirkungshypothesen* und ggf. *Mechanismen* zur Erläuterung der im Wirkungsmodell hinterlegten Ursache-Wirkungs-Beziehungen formuliert, die in der Evaluierung geprüft und bewertet werden können. Dabei sind auch mögliche Risiken bei der Implementierung des Vorhabens zu berücksichtigen.
- Eine *Kontributionsgeschichte*, die auf der Grundlage belastbarer, nachvollziehbarer und glaubhafter Evidenz die beobachteten Veränderungen und den Wirkungsbeitrag des untersuchten Vorhabens aufzeigt. Dafür müssen auch alternative Erklärungsansätze (beispielsweise Kontextfaktoren oder Maßnahmen Dritter) analysiert und die *Theory of Change* ggf. angepasst werden.

In den zentralen Projektevaluierungen sollen vornehmlich theoriebasierte Evaluierungsdesigns ausgewählt werden, die auf das Erkenntnisinteresse und den Evaluierungsgegenstand abgestimmt sind. Auf der Grundlage des Wirkungsmodells und des WoM der GIZ können die im Angebot formulierten Indikatoren und im Wirkungsmodell hinterlegten Wirkungshypothesen als Grundlage der Bewertung herangezogen und auf ihre Plausibilität hin überprüft werden. Bei der Datenerhebung kommen angemessene quantitative und qualitative Methoden zum Einsatz, beispielsweise Dokumentenanalyse, explorative Einzel- und Gruppeninterviews sowie standardisierte Onlinebefragungen. In Ergänzung zur theoriebasierten Vorgehensweise müssen zusätzliche Methoden zur Erfassung nicht intendierter Wirkungen sowie zur Bewertung der Effizienz eingesetzt werden.

4.4 Partizipatives Vorgehen

Die Partnerorientierung stellt ein wichtiges Charakteristikum der zentralen Projektevaluierungen dar. Dies spiegelt sich in den verschiedenen Phasen der Projektevaluierung und im Evaluierungsmanagement wider (bspw. durch die Verankerung des Erkenntnisinteresses der Partner in den ToR, Briefing zu Beginn der Evaluierung vor Ort, Erfassung der Partnerperspektive, Debriefing).

5. **Umfang und Inhalt des abzugebenden Angebots**

Die Stabsstelle Evaluierung möchte sicherstellen, dass die Auswahl der Gutachter/-innen konform mit der notwendigen Unabhängigkeit dieser ist. Nach Definition der Stabstelle Evaluierung trifft dies auf alle Gutachter/-innen zu, welche nicht an der Konzeption, Planung, Implementierung, Beratung oder Evaluierung des Vorhabens beteiligt waren – dies gilt für den/die internationale/n Gutachter/-in sowie den/die lokale/n Gutachter/in. Es werden nur solche Angebote in der Auswertung berücksichtigt, welche dem Kriterium der Unabhängigkeit in vollem Maße entsprechen. Ist das Kriterium nicht erfüllt, führt dies zum Ausschluss des Bieters aus dem Wettbewerb.

Das abzugebende Angebot sollte folgende Aspekte umfassen und den Umfang von drei bis fünf Seiten (exklusive CV) nicht überschreiten:

- Skizzierung eines methodisch anspruchsvollen Vorgehens inklusive eines theoriebasierten Ansatzes. Sowohl das Design als auch die Methoden der Datenerhebung sollten angemessen dargestellt werden. Die Stabsstelle Evaluierung möchte in Zukunft vermehrt kontributionsanalytische Ansätze in den Projektevaluierungen verfolgen. Eine Auseinandersetzung mit der Möglichkeit diesen Ansatz umzusetzen wird in der Auswertung positiv berücksichtigt.
- Darlegung der Arbeitsteilung im Gutachterteam
- Erfahrung in der deutschen und internationalen EZ/IZ, insbesondere mit der GIZ bzw. deren Vorgängerorganisationen
- Umfang und Qualität der Evaluierungserfahrung
- Sektorkenntnis und –erfahrung bzw. andere Kenntnisse und Erfahrungen, die für die Bewertung des Vorhabens relevant ist
- Auslandserfahrung (als GA oder Langzeit- bzw. Kurzzeitfachkraft) in der Region
- Sprachkompetenz
- Referenzen

Bitte verwenden Sie die CV-Formatvorlage, die Sie im Anhang an diese Ausschreibung finden.

6**. Mengengerüst**

Das Mengengerüst sollte insgesamt 58 FKT nicht überschreiten

- Inceptionphase bis zu 18 FKT
- Durchführung der Mission vor Ort inkl. Vorbereitung und Reisetage bis zu 26 FKT
- Auswertung und Berichtslegung bis zu FKT 14

Das Verhältnis der FKT von internationale/r Fachkraft und regionaler Fachkraft sollen sich an dem folgenden Verhältnis orientieren:
Internationale Fachkraft bis zu 37 FKT (inkl. evtl. Reisetage)
Regionale Fachkraft bis zu 21 FKT (inkl. Reisetage)

Reisekosten
Neben Flugkosten nach Sambia (pauschal) sind Reisen innerhalb Sambias sind mit 1.000,- Euro zu kalkulieren (Erstattung gegen Nachweis). Zusätzlich sind Übernachtungskosten und Tagegelder zu kalkulieren.

Anlagen

CV-Format

1. Family name, first name: ....
2. Place of Residence: ....
3. Contact/Address: ....
4. Date of birth: ....
5. Nationality: ....
6. Civil status: ....

7. Education:

Institution:	....
Date: from (month/year) *to (month/year)*	....
Degree(s) or Diploma(s) obtained:	....

8. Language skills: (Mark 1 to 5 for competence)

Language	*Reading*	*Speaking*	*Writing*
....	....	....	....

9. Membership of professional bodies: ....

10. Other skills: (eg Computer literacy, etc.) ....

11. Present position: ....

12. Years within the firm: ....

13. Key qualifications: (relevant to the programme)

    ....

14. Specific Countries experience:

Country	*Date: from (month/year) to (month/year)*
....	....

15. Professional Experience Record / List of accomplished evaluations:

Date: from (month/year) *to (month/year)*	……..
Location:	……..
Company:	……..
Position:	……..
Description:	……..
Applied evaluation design and methods of data collection	……..

Date: from (month/year) *to (month/year)*	……..
Location:	……..
Company:	……..
Position:	……..
Description:	……..
Applied evaluation design and methods of data collection	……..

Date: from (month/year) *to (month/year)*	……..
Location:	……..
Company:	……..
Position:	……..
Description:	……..
Applied evaluation design and methods of data collection	……..

Date: from (month/year) *to (month/year)*	……..

Location:	……..
Company:	……..
Position:	……..
Description:	……..
Applied evaluation design and methods of data collection	……..

16. Summary of experiences in working with development cooperation institutions (German and international; name of institution and type of cooperation)

………..

17. Other relevant information (e.g., Publications)

………
………

18. References (regarding conducted evaluations)

<u>First reference</u>:
Name:
Position:
Organization:
Email:
Phone:
In respect of which conducted evaluation:

<u>Second reference</u>:
Name:
Position:
Organization:
Email:
Phone:
In respect of which conducted evaluation:

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Sitz der Gesellschaft
Bonn und Eschborn

Friedrich-Ebert-Allee 36 + 40
53113 Bonn, Deutschland
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Deutschland
T +49 61 96 79-0
F +49 61 96 79-11 15

E info@giz.de
I www.giz.de

Photo credits and sources

Photo credits/sources:
© GIZ / Ranak Martin, Carlos Alba, Dirk Ostermeier, Ala Kheir (cover and back pages). Project related photos were kindly provided by the GIZ project team in Zambia.

Disclaimer:
This publication contains links to external websites. Responsibility for the content of the listed external sites always lies with their respective publishers. When the links to these sites were first posted, GIZ checked the third-party content to establish whether it could give rise to civil or criminal liability. However, the constant review of the links to external sites cannot reasonably be expected without concrete indication of a violation of rights. If GIZ itself becomes aware or is notified by a third party that an external site it has provided a link to gives rise to civil or criminal liability, it will remove the link to this site immediately. GIZ expressly dissociates itself from such content.

Maps:
The maps printed here are intended only for information purposes and in no way constitute recognition under international law of boundaries and territories. GIZ accepts no responsibility for these maps being entirely up to date, correct or complete. All liability for any damage, direct or indirect, resulting from their use is excluded.

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 36 + 40
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Germany
T +49 61 96 79-0
F +49 61 96 79-11 15

E info@giz.de
I www.giz.de