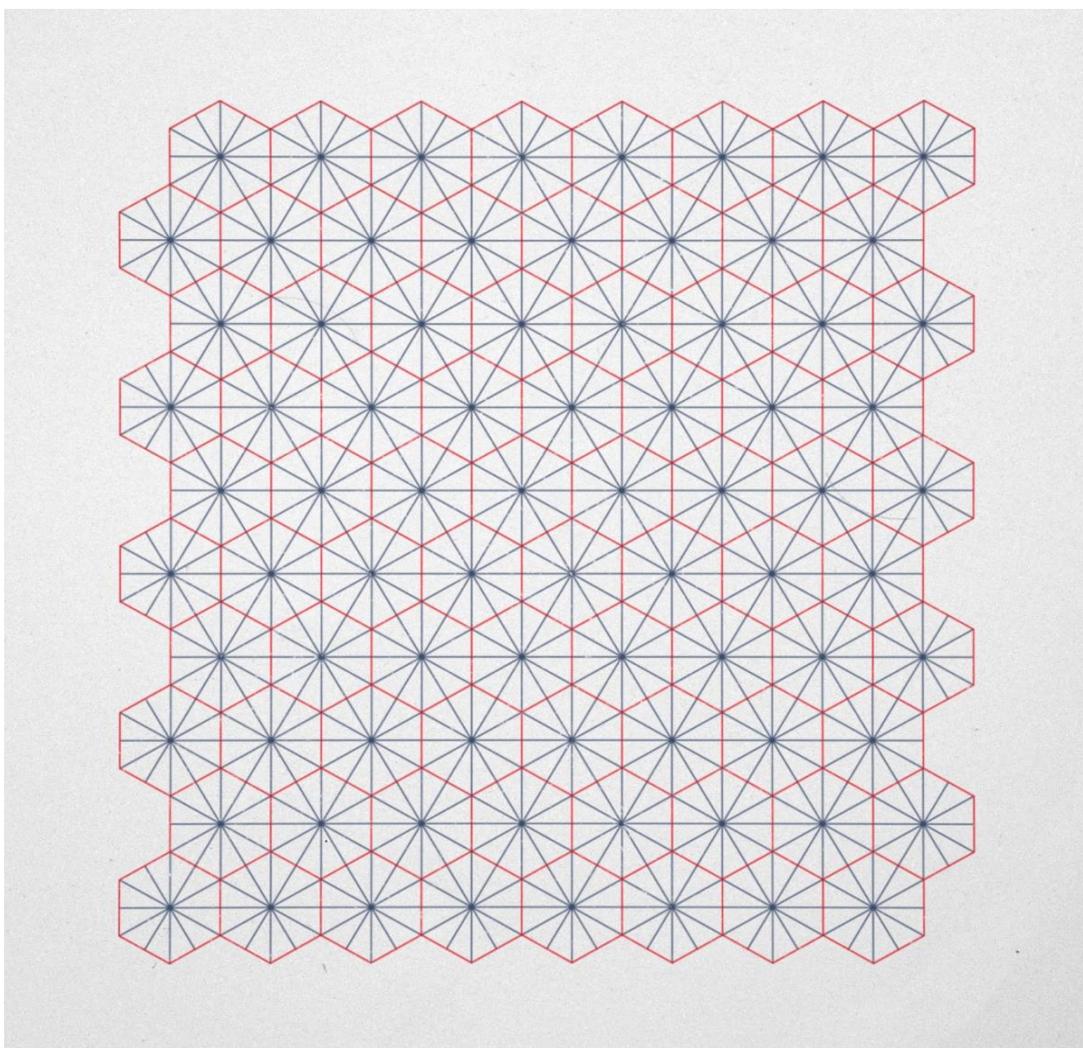


Results-Based Financing (RBF) in the Education Sector: Country-Level Analysis

Submitted to the REACH program at the World Bank



An Assessment of RBF in education in Mozambique, Nepal and Tanzania

Final Synthesis Report

July 2021

Catherine Dom, Alasdair Fraser, Joseph Holden and John Patch



Submitted by Mokoro Ltd. and Foresight Development Associates Ltd., in association with the University of East Anglia.

Mokoro Ltd.
The Old Music Hall
106-108 Cowley Road
Oxford OX4 1JE
+44 (0)1865 403179
mokoro@mokoro.co.uk

Suggested citation:

Catherine Dom, Alasdair Fraser, John Patch, Joseph Holden (2021) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Synthesis Report*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

The full set of deliverables for this assignment also includes:

Tanzania:

Catherine Dom, Alasdair Fraser, Edward Mhina (2021) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Tanzania*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

Catherine Dom, Alasdair Fraser, Edward Mhina (2020) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Tanzania*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

Nepal:

Joseph Holden and Yadab Chapagain (2021) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Nepal*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

Joseph Holden, Frances Hitchens, Yadab Chapagain (2020) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Nepal*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

Mozambique:

John Patch, Joseph Holden, Francis Ruhumbika (2021) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Mozambique*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

John Patch, Joseph Holden, Francis Ruhumbika (2020) *Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Mozambique*. Submitted to the REACH Program at the World Bank by Mokoro Ltd.

Inception:

Mokoro Ltd. Holden, J., Patch, J., Dom, C., Hitchens, F., Hodson, E., Umarji, M., Chapagain, Y., Mhina, E. *Evaluating Results-Based Financing in the Education Sector: Country-Level Analysis. Final Inception Report*. Commissioned by The World Bank., 31st July 2019

Table of Contents

Acronyms and Abbreviations	v
Acknowledgements and Disclaimer	ix
Executive Summary	x
1. Introduction	1
1.1 Theories of RBF underpinning the research	1
1.2 Change in education systems and levels of results	4
1.3 Research questions and study methodology	5
1.4 Study limitations	7
1.5 Structure of this report.....	7
2. RBF in Education	8
2.1 Overview of the three countries	8
2.2 Political economy and timeline of key events	10
2.3 Summary of country education sectors	14
2.4 RBF in education.....	19
3. Findings	23
A. Design	23
B. Results and impact.....	35
C. Financing and payments.....	48
D. Coordination, capacity and accountability	61
E. Evidence-based policy and verification	71
F. Adaptation and flexibility	79
G. Risks and unintended consequences.....	87
H. Cost-effectiveness.....	94
4. Conclusions.....	104
4.1 Steroids	105
4.2 Signposting	108
4.3 Autonomy and innovation	111
4.4 Sharpening minds	114
4.5 Labeling success	117
4.6 Aligning all actors	120
4.7 Sustaining attention	123
5. Lessons and recommendations	126
References	135

List of Tables

Table 1	Theories of how RBF works and associated risks.....	3
Table 2	Levels of metrics used in RBF in education	5
Table 3	Study methodology	6
Table 4	Mozambique, Nepal and Tanzania: summary data on characteristics.....	9
Table 5	Mozambique, Nepal and Tanzania: education high-level data	14
Table 6	RBF programming in the 2013–20 period	22
Table 7	RBF DLIs positions on the “education results chain”	29
Table 8	DLIs by technical area	31
Table 9	Summary of DLI results, their achievement, and the achievement and sustainability of associated higher-level goals	38
Table 10	RBF in sector financing across the three countries.....	49
Table 11	RBF disbursement and financial risk in the three countries.....	54
Table 12	Subnational cascading of RBF payments.....	57
Table 13	Accountability functions for the three countries.....	62
Table 14	RBF “ancillary components” in the three countries	67
Table 15	Sources of DLI information in RBF programming.....	72
Table 16	Verification process used in RBF across the three countries.....	74
Table 17	Implementation-related flexibility in RBF programs.....	80
Table 18	Technical assistance (TA), RBF dialogue and adaptation	84
Table 19	How RBF risks have materialized across the three countries.....	88
Table 20	Characteristics of different aid modalities	94
Table 21	Costs of RBF across the three countries	101
Table 22	Recommendations and lessons from the RBF assessment	127

List of Figures

Figure 1	Poverty ratios in Tanzania, Nepal and Mozambique, by geography	9
Figure 2	Timelines of key events for Nepal, Mozambique and Tanzania.....	13
Figure 3	RBF design as a program within the sector strategic plan	23
Figure 4	The different means of assessing results and impact for RBF	35
Figure 5	Mozambique lower primary PTR by province: 2013, 2017 and 2020.....	41
Figure 6	Tanzania: Measures of variation in PTRs within LGAs during EP4R period, 2015-19.....	41
Figure 7	Complexity zone and selected DLIs within the RBF programming in the three countries	44
Figure 8	Mozambique education, 2014–18, RBF potential disbursements and non-RBF actuals.....	51
Figure 9	Nepal SSDP value by donor and DLI, 2016–21 (allocations not actuals)	52
Figure 10	Tanzania RBF 2014/15-2019/20, EP4R by donor (allocations, not actuals).....	53

List of Boxes

Box 1	Levels of indicators in education systems	4
Box 2	The “big bang” of federalism in Nepal	12
Box 3	Initial steps in using RBF in Mozambique, Nepal and Tanzania	25
Box 4	Prioritization of different donors in the development of DLIs in Nepal	27
Box 5	Mozambique – development of sector RBF out of government PFM vision	28
Box 6	Tanzania – RBF funders’ mixes of incentivized results by level.....	30
Box 7	Strategies for tackling teacher deployment.....	32
Box 8	Mozambique – The unclear additionality of PFM4R financing	50
Box 9	Tanzania’s experience of disbursement compared to the recent past	56
Box 10	The use of subnational RBF payments in EP4R design in Tanzania.....	58
Box 11	Nepal – the limits of a top-down “control” approach under federalism	64
Box 12	Tanzania - An instrumental approach with insufficient support to key agents.....	64
Box 13	Mozambique PFM4R: Achieving timely disbursement of direct grants to schools (ADE) ...	65
Box 14	Management and tendering of the independent verification agent (IVA)	75
Box 15	EMIS, enrolment numbers and ghost pupils in Mozambique	77
Box 16	Nepal – Adapting DLI targets to federalism, the case of a governance and fiduciary management DLI	82
Box 17	Tanzania – Moving the goal posts on a changed pitch	83
Box 18	Flexibility in RBF programs and financial implications.....	85
Box 19	Risk to agents’ motivation – Donors failing to share the risks	90
Box 20	Nepal – Equity and counteracting DLIs.....	92
Box 21	Risks of incentivizing school performance in Tanzania – cherry picking and equity	93
Box 22	Priority alignment and its effects on RBF “success”.....	97

Acronyms and Abbreviations

ADB	Asian Development Bank
ADE	Direct Support to Schools – school grants in Mozambique (<i>Apoio Directo às Escolas</i>)
AF	Additional Financing
BRM	Budget Review Meeting (usually taking place in March/April every year, Nepal)
BRN	Big Results Now (Tanzania)
BRNEd	Big Results Now/Education (Tanzania)
BS	Budget support
CB-EGRA	Classroom-based Early Grade Reading Assessment
CD	Capacity development
CEHRD	Center for Education and Human Resource Development (Nepal)
CG	Capitation Grant (Tanzania)
CSO	Civil society organization
DEO	District Education Office (Nepal)
DFID	Department for International Development (UK) (former title)
DLI	Disbursement linked indicator
DP	Development partner
DRR	Disaster risk reduction
EC	European Commission
ECED	Early Childhood Education
EDCU	Education Development and Coordination Unit (Nepal)
EGMA	Early Grade Math Assessment
EGR	Early grade reading
EGRA	Early Grade Reading Assessment
EGRP	Early Grade Reading Program (USAID-funded program)
EMIS	Education management information system
EP4R	Education Program for Results (Tanzania)
EQUIP-T	Education Quality Improvement Program in Tanzania
ERO	Education Review Office (Nepal)
ESDP	Education Sector Development Program (Tanzania)
ESSP	Education Sector Support Project (Mozambique)
EU	European Union
EY	Ernst & Young (IVA for GPE DLIs, also role as auditor of FASE)
FASE	Education Sector Support Fund [<i>Fundo de Apoio ao Sector de Educação</i>]
FCDO	Foreign, Commonwealth & Development Office (UK)
FFBEP	Fee-Free Basic Education Policy (Tanzania)
FRELIMO	Mozambique Liberation Front [<i>Frente de Libertação de Moçambique</i>]
GBS	General budget support
GDP	Gross domestic product

GER	Gross enrolment ratio
GiZ	German Corporation for International Cooperation [<i>Deutsche Gesellschaft für Internationale Zusammenarbeit</i>]
GNI	Gross national income
GoM	Government of Mozambique
GoN	Government of Nepal
GoT	Government of Tanzania
GPE	Global Partnership for Education
HDI	Human Development Index
ICT	Information and communication technology
IDA	International Development Association
IEMIS	Integrated-Education Management Information System (refers to electronic version developed from 2014, in Nepal)
IFP	Teacher training institutes (in Mozambique) [<i>Instituto de Formação de Professores</i>]
ILI	Incentive-linked indicator (alternative term for DLI used in SSRP period in Nepal)
IMF	International Monetary Fund
IPF	Investment Project Financing
IT	Information technology
IV	Independent verification
IVA	Independent Verification Agent
JFA	Joint Financing Agreement
JFP	Joint Financing Partner
JICA	Japan International Cooperation Agency
JRM	Joint Review Meeting (usually in November every year, Nepal)
JSR	Joint sector review
KfW	<i>KfW Bankengruppe</i> (banking group), German state-owned development bank
KOICA	Korean International Cooperation Agency
KPI	Key performance indicator
KUSOED	Kathmandu University School of Education (Year 1 IVA for SSDP, Nepal)
LANES	Literacy and Numeracy Education Support (GPE program 2014-2018)
LEG	Local Education Group
LG	Local Government
LGAs	Local Government Authorities (Tanzania)
M&E	Monitoring and Evaluation
MINEDH	Ministry of Education and Human Development – <i>acronym used in Mozambique and in this report used for references</i>
MoE	Ministry of Education – <i>acronym used in body of this report for all 3 countries</i>
MOEST	Ministry of Education, Science and Technology – <i>commonly used acronym in Nepal and Tanzania, and in this report used for references</i>
MoF	Ministry of Finance – <i>acronym used in body of this report for all 3 countries</i>
MEF	Ministry of Economy and Finance – <i>acronym used in Mozambique</i>

MoFP	Ministry of Finance and Planning – <i>acronym used in Tanzania</i>
MTR	Mid-term review
NASA	National Assessment of Student Achievement (Nepal)
NCP	Nepal Communist Party
NEB	National Examination Board (Nepal)
NECTA	National Examination Council of Tanzania
NEGRP	National Early Grade Reading Program (Government Program, Nepal)
NER	Net enrolment ratio
NGO	Non-governmental organization
NPR	Nepalese Rupee
ODA	Overseas Development Aid
OOSC	Out-of-school children
OPM	Oxford Policy Management
PAD	Program Appraisal Document
PBA	Performance-Based Allocation (via the PFM4R program)
PBL	Project-Based Lending (ADB instrument)
PFM	Public financial management
PFM4R	Public Financial Management for Results Program (Mozambique)
PforR	Program for Results Financing
PIM	Program Implementation Manual (for SSDP, Nepal)
PO-PSM	President’s Office – Public Service Management (Tanzania)
PO-RALG	President’s Office – Regional Administration and Local Government (Tanzania)
POEMA	<i>Planificação, Orçamentação, Execução, Monitoria e Avaliação</i> (Planning, Budgeting, Execution, Monitoring and Evaluation) – capacity development program implemented by GiZ in Mozambique
POM	Program Operations Manual (EP4R, Tanzania)
PTA	Parent-teacher association
PTR	Pupil-teacher ratio
RBF	Results-Based Financing
RBL	Results-Based Lending (ADB lending instrument)
RC	Resource Center (Nepal)
REACH	Results in Education for All Children program
RISE	Research on Improving Systems in Education
RP	Resource Person (Nepal)
SBS	Sector Budget Support
SDEJT	District Services for Education, Youth and Technology (Mozambique) [<i>Servicos Distritais de Educaçao Juventude e Tecnologia</i>]
SDI	Service Delivery Indicator (survey, by World Bank)
Sida	Swedish International Development Agency
SIG	School Improvement Grant
SMC	School Management Committee

SMT	Senior Management Team
SQA	School Quality Assurance
SSDP	School Sector Development Plan (2016/17 – 2022/23) (Nepal)
SSRP	School Sector Reform Plan (2009 – 2015/16) (Nepal)
SWAp	Sector-Wide Approach
TA	Technical assistance
TAS	Technical Assistance Support (Tanzania)
TOC	Theory of change
ToR	Terms of Reference
TST	(Teacher) Time-spent-teaching
TU-CERID	Tribhuvan University, Center for Educational Research Innovation and Development (Year 2 IVA for SSDP, Nepal)
TVET	Technical and Vocational Education and Training
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children’s Fund
USAID	United States Agency for International Development
USD	United States Dollar
VT	Variable Tranche (part of GPE, EU and Finland financing modality)
WB	The World Bank

Acknowledgements and Disclaimer

The research team expresses its thanks to the education stakeholders whom we met in Tanzania, Nepal and Mozambique during country visits in late-2019, and interviewed at a distance in the course of 2020, including senior officials in the Ministry of Education, Sciences and Technology and the President's Office, Regional and Local Government Administration in Tanzania; the Ministry of Education and Human Development in Mozambique; and the Ministry of Education, Sciences and Technology, and Center for Education and Human Resource Development in Nepal. In addition, we would like to thank all other government officials at the national, regional and local levels, in addition to representatives from development partners, civil society organizations and non-government education providers, as well as consultants and technical assistance team members. We thank all for their time and thoughtful contributions, as well as for providing the team with all relevant documentation at their disposal. We also thank the REACH team for their support, notably in introducing the study to country stakeholders, and the World Bank country teams, for linking the team with relevant stakeholders, in addition to providing their own insights.

Quality and logistical support during the study was provided by Stephen Lister, Zoe Millington and Liam Bluer. The country study teams included Edward Mhina in Tanzania, Francis Ruhumbika in Mozambique, and Yadab Chapagain in Nepal. Quantitative analysis for the final country reports was conducted by and with support from Dr. Paul Clist at the University of East Anglia and Lucas Sempe at the Universidad Católica San Pablo.

Full responsibility for this work remains with the authors, and the views the report contains should not be attributed to the Government of Tanzania, the Government of Mozambique, the Government of Nepal, the World Bank, the REACH program, or any other stakeholder.

Executive Summary

S1. This report has been commissioned by the World Bank Results in Education for All Children (REACH) program as an assessment of the design, implementation experience, and impacts of results-based financing (RBF) in education. The study covers three countries: Tanzania, Mozambique and Nepal and seeks to explore the experience of each with RBF in education since 2013.

S2. “Results-based financing” covers a range of instruments that have been increasingly used in education and other sectors in recent years. RBF instruments are defined as those paying for a result to be achieved *only* if achievement can be credibly verified. In education, RBF includes forms of contract that disburse on results such as, for example, children enrolling and learning, or the number of classrooms constructed. This contrasts with instruments that transfer resources unconditionally or simply pay input costs, for example salaries, or materials for construction.

S3. RBF is a broad term that can cover uses such as performance-based grants to schools or local authorities, performance-based pay, development impact bonds, or conditional cash transfers, among others. RBF has, however, mainly been used as a financing modality by international aid donors. This has included the growing use of financing instruments such as Program for Results (PforR) by the World Bank, Results-Based Lending (RBL) by the Asian Development Bank (ADB), and the use of a variable tranche by both the European Union and the Global Partnership for Education (GPE). All link disbursements to results. RBF approaches have also increasingly been used by bilateral aid agencies such as the UK’s Foreign, Commonwealth & Development Office (FCDO) and the Swedish International Development Cooperation Agency (Sida).

S4. As an aid modality, the conditional nature of RBF provides an alternative to most uses of general budget support (GBS) or sector budget support (SBS). From the early 2010s, donors found these modalities problematic, with a number of evaluations raising questions on the effectiveness of

GBS and SBS in addressing systemic problems faced in key sectors such as education.¹

S5. While RBF programs are distinguished as a modality in the use of conditionality, they can vary in the degree of difference to other modalities on a number of other aspects. These include the rigor of evaluation required; the degree of alignment with government financial reporting and accountability systems; the presence of safeguards; the role of technical assistance; and whether or not they use existing basket or pooled funding mechanisms.

S6. The use of RBF in education in Mozambique, Tanzania and Nepal has been part of the shift in donor preferences with respect to financing instruments and modalities, also prevalent in other sectors such as health and infrastructure. RBF programs are structured through disbursement linked indicators (DLIs), which specify particular results for a given target date and a payment tariff associated. In each of the three countries, DLIs have generally been a subset of broader sector results covered by education sector development plans.

Methodology

S7. The research has addressed the question: “*To what extent, in what ways and under which conditions has RBF contributed to strengthen education systems to deliver results?*” This has been unpacked through a number of research questions, guiding the study and responding to the ToR. The research was undertaken in two phases, with a first phase “mapping” the RBF present in education in each country. The first phase led to a Preliminary Assessment Report for each country, finalized in February 2020, following country visits in late 2019. The second phase has included a further range of stakeholder interviews at national and local levels over the course of 2020, and led to Final Assessment Reports for each country with more in-depth analysis on the contribution of RBF to education results. This final Synthesis Report, seeks to compare and contrast the country experiences, and to identify lessons and recommendations.

¹ See, for example, DANIDA (2014) for a summary.

Table S1 Study methodology

Method	Details
Education system review	Detailed appraisal of each country education system, recent historical and political context, sector performance and trajectories, and review of sector plans and evaluations.
RBF ‘mapping’	Identification of all RBF schemes in education (covering primary and secondary level). Analysis of how they interact, and a detailed review and mapping of all DLIs in the EP4R in Tanzania; the PFM4R, GPE variable tranche and German (KfW) use of RBF in Mozambique; and the use of RBF in the SSRP (2009-16) and SSDP (2016-21) in Nepal.
Key informant interviews	Interviews with: <ul style="list-style-type: none"> • Central government stakeholders at MoE, MoF, and education agencies. • Development partner and civil society stakeholders. • Provincial and local government stakeholders. • School principals and school council chairs.
Quantitative and qualitative review of DLI performance	Review of DLI results, disbursement and financial flows; assessment of their complexity; measurement and risks that emerged. Detailed assessment of design, experience, evidence and timing of results, and links with non-incentivized measures of success and overall sector performance.
Quantitative analysis of subset of DLIs	Deeper qualitative and quantitative analysis on a selection of DLIs, notably the GPE pupil teacher ratio (PTR) DLI in Mozambique; the out-of-school children enrolment DLI in Nepal’s SSDP; and the PTR and early learning (literacy) outcome DLIs in Tanzania.
Assessment against theories of change of RBF	Analysis of RBF contribution to results, the presence of incentive effects and behavioral responses, and broader systemic benefits and costs of the use of RBF.

S8. The methodology is summarized above in Table S1. For each country, research covered the design, results, impact, and payments of RBF. In addition, the link between RBF and the coordination, capacity and accountability of institutions; the role of independent verification and evidence; and the adaptiveness of RBF programming were studied. The research has also looked at the risks and unintended consequences of RBF; and its cost-effectiveness. Findings are set out in detail in the Final Assessment Reports for each country and at a high level in this Synthesis Report.²

S9. The findings form the basis of an assessment of the plausibility of the various theories of change for RBF identified from an extensive literature review. This includes whether an incentive effect has been present (via theories of “steroids”, “signposting”, or “autonomy”), and whether other systemic effects are apparent as a result of RBF, including on evidence (“sharpening minds”), planning processes (“labeling success”), co-

ordination (“aligning all actors”), or maintaining a focus on results over time (“sustaining attention”).

S10. The research set out to look for any and all RBF in basic education in the three countries.³ The countries were selected by REACH because of their known experience with RBF through flagship programs linked to the use of the World Bank’s PforR instrument. The mapping of RBF found that despite some selective use of subnational RBF beyond these programs, the majority of RBF in terms of financial value and scope has been linked to donor aid. Where subnational RBF was found, it was mainly linked to the larger sector programs with donor support, i.e. as an incentivized component of the programs as opposed to purely ‘home grown’ uses of RBF. It is because of this finding that the research has mainly focused on the large-scale uses of RBF as an aid modality in each of the three countries, and on which this Synthesis Report also principally focusses.

² More details in: Dom et al. (2020, 2021); Patch et al. (2020, 2021); Holden et al. (2020); Holden and Chapagain (2021).

³ The study did not cover higher education or early childhood education.

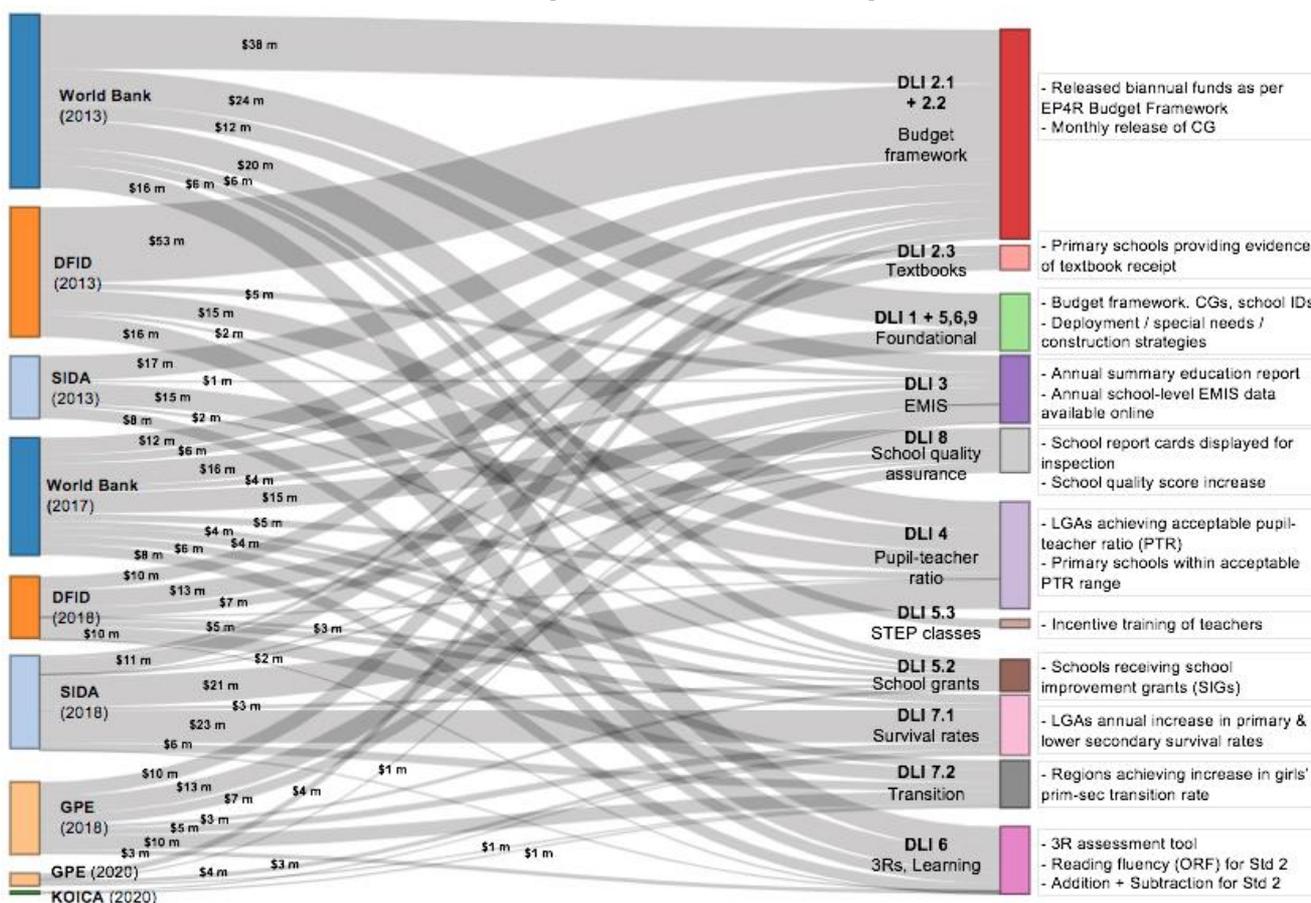
RBF in Tanzania

S11. In Tanzania RBF was introduced in support of the government’s “Big Results Now” (BRN) initiative in 2014, with sector finance linked to results for the World Bank, the United Kingdom (then via DFID) and Sida. The RBF program created a pooling mechanism, and became known as the Education Program for Results (EP4R). In total, five donors have committed over USD 500 million of potential disbursements linked to DLIs. As illustrated in Figure S1, these commitments were made in two main phases, at the outset in 2013-14, and in a significant program redesign in 2017-18. The program has been jointly implemented by the Ministry of Education, Science and Technology (MoE) and the

President’s Office – Regional Administration and Local Government (PO-RALG).

S12. As the program arose from the BRN initiative it began with high level political buy-in to its focus on learning. An important contextual change during the period was the introduction of the fee-free basic education policy (FFBEP) in 2015, implementing the then new President’s campaign promise of expanding access. This drove an increase in primary enrolment of 29 percent between 2014/15 and 2019/20 whilst teacher numbers were capped and the education budget stagnated – and created a challenging context in which the program nonetheless retained its focus on learning.

Figure S1. Tanzania education RBF 2014/15-2019/20, EP4R by donor (allocations, not actuals)



Source: Up-to-date as of June 2020. MoEST & PO-RALG (2020) provides implied exchange rates; authors' calculations

S13. Important design aspects of the Tanzanian RBF program include its significant focus on learning, accounting for nine percent of the potential disbursements, as well as transition rates for girls and survival rates (for boys and girls), with 14 percent of potential RBF disbursements. This was

combined with a focus on inputs via the budget framework DLIs on fullness of budget release for selected budget lines including capitation grants (CGs) to schools, which covered close to one-third of total potential disbursements. The program had a relatively small number of process or output

indicators in the “middle” of the results chain, although equitable pupil teacher ratio (PTR) for primary schools was a big focus, covering 15 percent of potential RBF disbursements.

S14. Technical assistance was included in the DFID program to support the EP4R as a whole. Verification for the program was required and undertaken by a consultancy hired by DFID on behalf of all funders. Verification expanded in scope, depth and degree of specification over the life of the program and became an important feature of the program’s management cycle.

S15. The program has involved a downstream payment scheme for around one-third of RBF payments, with transfers to local government authorities (LGAs). Performance has been good, for both national and LGA levels, although with substantial differences between LGAs arising at least in part from a lack of attention to equity issues.

S16. The overall disbursement rate has compared well to previous modalities used to support education in Tanzania. However, some important DLI targets were missed, including more equitable teacher deployment, and increased early learning outcomes. It became more challenging to meet these targets due to the effects of the under-resourced FFBEF-driven sector expansion.

RBF in Nepal

S17. In Nepal, RBF was introduced on a relatively small scale by the World Bank from 2013, towards the end of the School Sector Reform Plan (SSRP) period (2009 - 2015/16). This included four target results on specific process indicators covering textbooks, paying teachers and scholarships via bank accounts, and the dissemination of national student assessment (NASA) reports. Across these targets USD 20 million was disbursed against a potential total of USD 25 million.

S18. Nepal then shifted to receiving the major share of donor finance to education via RBF, in support of the government’s School Sector Development Plan (SSDP) from 2016. The five-year costed sector strategic plan (2016 – 2020/21), is principally funded by government, and to date has involved approximately 10 percent of funding from external finance. This is less than the 20 percent

external finance contribution for the previous SSRP period (based on actual spend data).

S19. The SSDP saw a major shift to using RBF as the principal financing modality for donors. Six out of nine joint financing partners (JFPs) to the SSDP linked all or part of their potential disbursements to DLIs, in total worth USD 400 million and covering over 80 percent of potential external finance to the basic education sector from 2016 to 2021.

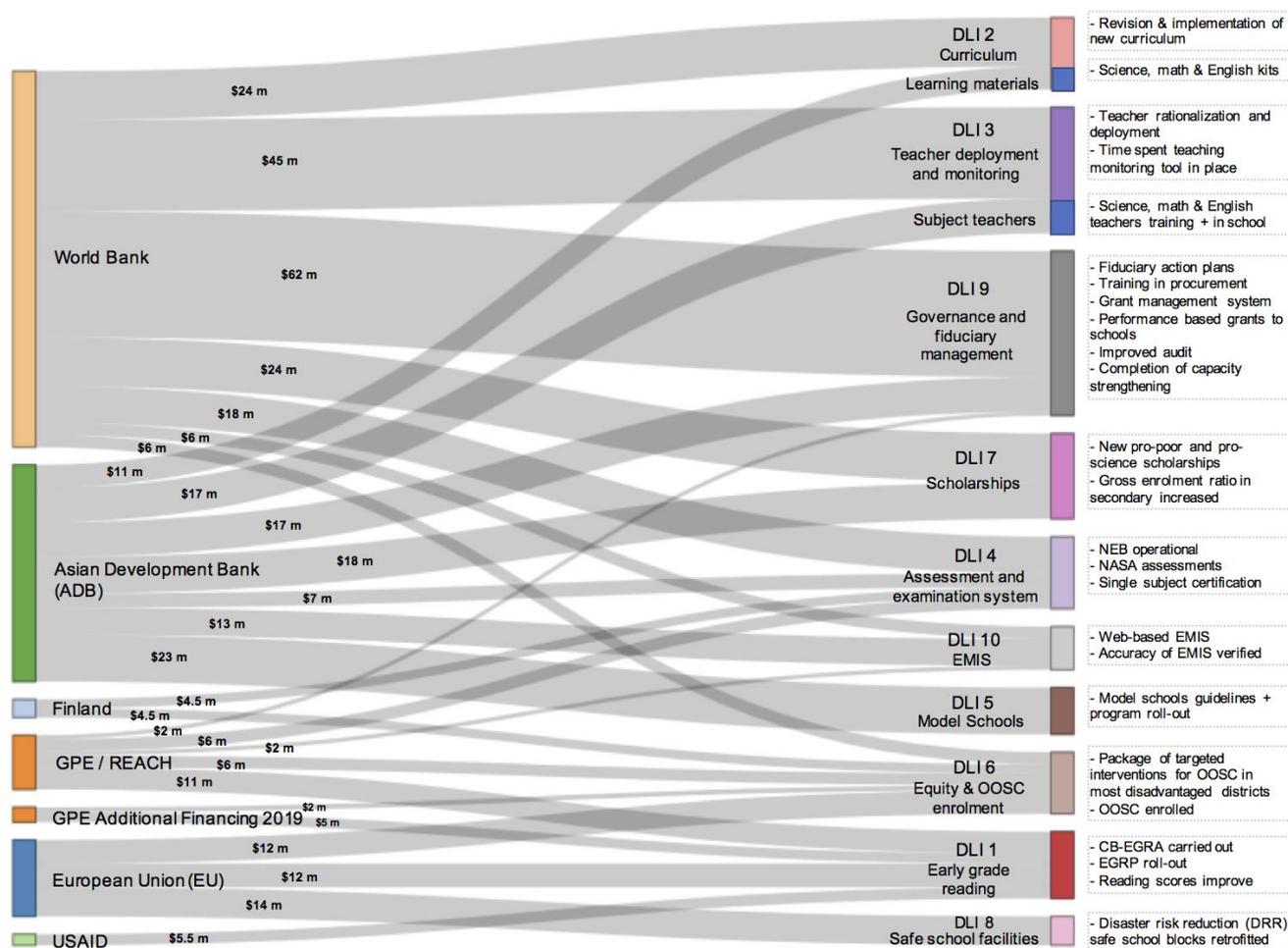
S20. The SSDP includes a large number of DLIs, as shown in Figure S2, with transfers for ten ‘DLI areas’, translating into 84 individual DLI results with a range of process, output and outcome results. The majority of potential disbursements are linked to process and output indicators, and only a small proportion of the total is linked to learning outcomes (0.1 percent of total RBF value). The largest DLI areas focus on a number of indicators relating to governance and fiduciary management, as well as a range of interventions focusing on teachers.

S21. The large-scale use of RBF took place during a period of significant change for Nepal. Firstly, the major earthquake of 2015 necessitated large-scale reconstruction, and limited the funds available for other purposes. Secondly, the Constitution of 2015 introduced federalism. A new tier of 753 local governments (LGs) replaced the deconcentrated structure of 75 districts. The new LGs took on most roles of the previous districts for education, with many changes coming into force in 2017-18.

S22. A number of DLI results have been delayed or not achieved, largely linked to the challenges of federalism – including reorganization to reporting structures, contested mandates, and upheaval to the civil service. Since the SSDP mid-term review (MTR) in 2019, DLIs have been revised to try to reflect the requirements of the transition to federalism.

S23. Verification is undertaken by one independent verification agent (IVA) for the whole plan, contracted via government on an annual basis, although not all donors require this for payment against all DLIs. Other donor support includes technical assistance to the MoE and its implementing agency, the Center for Education and Human Resource Development (CEHRD), funded by the EU and the ADB and implemented by the British Council.

Figure S2. Nepal SSDP RBF value by donor and DLI, 2016–21 (allocations not actuals)



Source: SSDP Joint Protocol updated to October 2019. Exchange rate used of €1 = USD 1.12 for currency conversions.

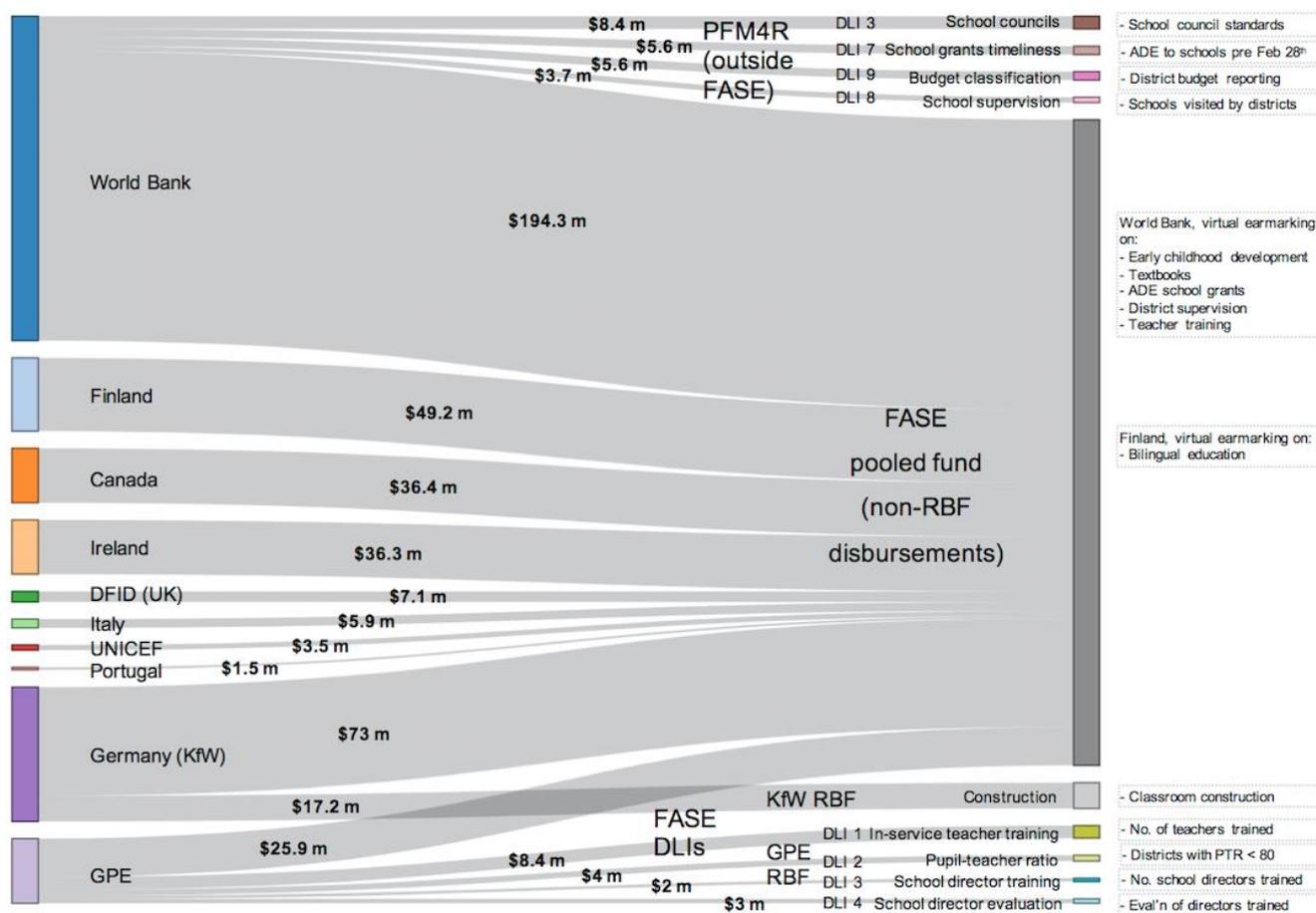
RBF in Mozambique

S24. RBF was introduced to Mozambique with the World Bank’s Public Financial Management for Results (PFM4R) program (2014-18). The program was an initiative for PFM reform, in which the education and health sectors were selected from a number of other potential candidates. The program was designed to improve subnational PFM and service delivery, with a focus on accountability in the education sector: both vertical, from the MoE to subnational government to school; and horizontal, via the involvement of communities and parents in school councils. In part, interventions were linked to data suggesting teacher attendance was particularly low in Mozambique with 45 percent not in school during unannounced visits (World Bank, 2015a). The PFM4R had four specific process DLIs which were on budgeting, on the timeliness of school grants, the amount of schools covered by district supervision, and school council functioning.

S25. The PFM4R provided funds outside the education sector pooled fund (FASE), through which the majority of external finance to the education sector has flowed in the past 20 years. This meant the PFM4R brought additional financing into the sector (approximately USD 25 million). As payments were made to the Ministry of Economy and Finance (MoF), it is not clear how the funds were all used in practice, or if they were used for education.

S26. In the same period, the GPE and Germany (KfW) used RBF in their funds disbursed to FASE. A USD 17 million GPE variable tranche (2015-17), was tied to four DLIs on some specific output results – in-service teacher training, school director training and evaluation, and lowering pupil teacher ratios (PTRs) for lower primary schools in districts where they were high (number of districts with a PTR for lower primary lower than 80). KfW focused on increasing the number of classrooms constructed in 2014 and 2015, with RBF also worth USD 17 million.

Figure S3. Mozambique education, 2014-18, RBF potential disbursements and non-RBF actuals



Source: FASE Financial Statements including EY (2018), KPMG (2019), World Bank (2014, 2015b, 2019b), KfW (2017), Orlowski (2016).

S27. In total, the three RBF funders provided USD 58 million in potential disbursements, or 12 percent of donor financing to education from 2014 to 2018, as shown in Figure S3. The diagram includes the much larger non-RBF fund flow from donors to the FASE sector pooled fund as a comparison. The vast majority of external finance to education in Mozambique in the period was therefore not conditional or based on RBF.

S28. The PFM4R program included within its design a significant amount of technical assistance in the form of program facilitators and coaches to the MoE, MoF and provinces, with training to districts also provided. This helped the program to achieve the DLI results by the final year of the program, and all funds were disbursed by its closure. Subnational financial rewards via “performance based allocations” (PBAs) were included in the program design, but did not function as intended, and may have caused some challenges at province and

district levels in terms of unmet expectations in the first years of the program.

S29. There is less evidence that the deeper PFM goals of the program, including transformation to sectoral accountability, were achieved. This is linked to the finding that the DLIs were limited to overcome “foundational” sector bottlenecks, which at the time of design were seen as providing an appropriate level of ambition in the context.

S30. The KfW use of RBF in construction saw non-disbursement for around one-third of its value and was not seen as successful in tackling underlying bottlenecks. GPE DLIs were achieved with the exception of partial achievement of the PTR DLI. Quantitative analysis for this assessment on the PTR DLI found statistically significant improvement in within-province equity of lower primary school PTRs. The DLI incentive effect was therefore found to be present. However, this was not fully recognized in payment, partially due to district reorganization in

2013 that was not reflected in the DLI baseline. DLI targets for in-service teacher and school director training undertaken through teacher training institutes were met but with very little information on the quality of provision or of any system change.

S31. A major contextual change in Mozambique was the shock to external finance and suspension of general budget support (GBS) following the “hidden loans” scandal in 2015-16, which meant that overall education sector finance stagnated, with implications including for the number of new teachers that were hired on an annual basis.

S32. The PFM4R used the country’s national audit institution, the *Tribunal Administrativo*, for the IVA role. The GPE use of RBF had a private consultancy for the IVA role. The KfW RBF included no IVA role. Complementary donor programs were also present including POEMA⁴, which provided capacity support on PFM and planning in three provinces, funded by Germany (GiZ), though not linked to the RBF program design.

Conclusions

S33. The theory of RBF is generally framed in terms of principal-agent theory. The *principal*, seeks to achieve a specific goal by incentivizing the *agent* whose action is required to achieve it, where with RBF, the incentives are financial. The theory suggests some clear ways through which RBF may work, which we categorize as the following ideas:

- the agent will put more effort or resources into achieving the specified results [*steroids*];
- effort or resources will be focused or diverted to the specified results or from other priorities [*signposting*], and
- paying for results as opposed to inputs can give the agent more autonomy to achieve results [*autonomy and innovation*].

S34. Proponents of RBF also claim other ways in which it may improve education systems. These include:

- improvements to evidence systems [*sharpening minds*];
- a focus on results improving results-based planning [*labeling success*];
- bringing actors together in cooperation and coordination to achieve results [*aligning all actors*]; and
- by “contractually” defining results, RBF provides a mechanism to focus on these over time, whatever else may be changing [*sustaining attention*].⁵

S35. While principal-agent theory provides insights as to why RBF might work, alternative theories and empirical literature, point to risks of RBF.⁶ One risk is tied to the fact that people and institutions respond to other forms of accountability than just financial incentives, for example electoral accountability, peer-to-peer accountability, and intrinsic motivation. The risk is that RBF may “crowd out” or even undermine these other forms of accountability. Other risks are often the flipside of the theoretical claims made for RBF – for example, RBF provides incentives, but perhaps the incentives are too strong, thus creating risks such as gaming, or diverting attention and resource from other important sector objectives and goals.

S36. We summarize below the conclusions of this three-country assessment against each of the main theories for RBF. This is informed by the detailed assessments of RBF programming in education in place since 2013 in each country, the experiences of sector stakeholders, and the detailed review of all DLIs including their performance and complexity. Conclusions are based on the wide range of qualitative and quantitative information and analysis conducted for the assignment, with comparisons setting out key similarities and differences across the programs studied. The findings are structured around each of the theoretical claims made for the effectiveness of RBF, and how and where risks materialized.

⁴ *Planificação, Orçamentação, Execução, Monitoria e Avaliação* (Planning, Budgeting, Execution, Monitoring and Evaluation).

⁵ These theories are based on a large literature survey, including the work of REACH, in World Bank (2017a).

⁶ See, for example Cuevas-Rodríguez et al. (2012).

“Steroids” Theory:

RBF induces more effort (and/or resources) to achieve specified results. A goal is more likely to be achieved if a payment is made on its achievement, because more effort will be exerted to achieve the goal by agents.

Potential risks:

*Gaming risks relating to the RBF metric, such as fudging, where success is in appearance only.
Demotivation, by crowding out other accountability mechanisms.*

S37. The foremost theory for RBF is that it *aligns* the incentives of the agent to achieve the targeted results. This should mean, at minimum, the agent putting more effort and/or resources into related activities (‘steroids’). For the RBF assessed, the main agent was the MoE, who could either direct change to other agents (e.g. local governments, schools) or pass on (‘cascade’) the incentive to these agents.

S38. A key condition for RBF to work is that the metric and methods by which a result is measured is truly representative of the intended goal. In the absence of this condition there are risks of gaming.

S39. The conclusions of this assessment on the “Steroids” theory of change include:

- There was evidence of increased effort and resource to results linked to RBF incentives in each of the three countries. The degree of performance on DLIs was not found to be linked to the size of the associated RBF payment.
- There were instances in which DLIs aimed to incentivize additional financing for education from the MoF. It was less clear that RBF was effective in this respect.
- Tanzania’s EP4R included subnational RBF payments in its design. Local government (LGAs) recognized this incentive, including to improve survival and transition rates. Mozambique’s PFM4R also included subnational payments in design but these did not function as intended or align to existing subnational PFM systems. In Nepal, subnational government was reorganized during the program period, and the PFM system fundamentally reshaped. This complicated the ability of the MoE to implement changes represented by DLIs.
- Tanzania’s program aimed furthest up the education results chain and could be described as more transformational in its ambitions. In Nepal, and in particular Mozambique, the RBF concentrated on lower-level results. While

Tanzania was comparatively “hands-off” in relation to how results would be reached, in Nepal and Mozambique, DLI results had more linear sequencing, similar to project results chains (e.g. process to output to outcome).

- In Tanzania, the steroids effect of RBF funding was present but the effectiveness of the learning focus was partly offset by the Presidential drive to increase access through the FFBEF (without commensurate resources), and a focus on investing in school infrastructure.
- Under the Nepal SSDP a steroids effect was clearly present for the MoE and CEHRD. The ability to achieve DLIs seemed to be constrained by their large number and complexity, and by the changes entailed by the federal transition, particularly from 2017. From that point, DLIs involving local-level delivery faced significant challenges.
- In Mozambique, foundational DLIs were well understood and aligned with MoE objectives, though it has been hard to disentangle a steroids effect from accompanying elements of the PFM4R design (technical support from coaches and facilitators). The GPE RBF did not have this support but appeared to have a steroid effect. The German (KfW) RBF was less successful in speeding up classroom construction.

S40. There were some cases of risks materializing:

- The selection of simple indicators to measure complex change at times allowed reporting of apparent success when it looks unlikely that the underlying goals had been achieved. There were limited attempts to use a range of evidence to triangulate DLI reporting and offset this risk.
- There was often an absence of strong qualitative understanding of results and the link to impact. For example, the “last mile” of school grants being spent well at schools, or changes to skills and motivation of teachers and school directors as a result of their training.

“Signposting” Theory:

RBF focuses effort (and/or resource) to achieve specified results. Within finite limits to efforts/resources, allocations of effort/resource to specified results reduce those to less important areas.

Potential risks:

Diversion risks if there is a suboptimal resource allocation, with important areas deprioritized because of the RBF. Cherry-picking risks, in which less costly or politically favored subgroups or regions see greater resources, in part because they are more likely to reach a target.

S41. The incentive of RBF to *focus* effort and resource on specified results can be viewed as a ‘*signposting*’ effect. RBF incentivizes some results but not others, implying they are more important to sector performance. This implies an opportunity cost and a risk of de-prioritizing other important goals.

S42. All the RBF programs involved design phases with DLIs agreed upon by development partners and government in a combination of multilateral and bilateral settings. During the program period (three to five years for the programs assessed) important contextual changes occurred including to sector priorities. The degree of alignment and success of any signposting could therefore change over time.

S43. The conclusions of this assessment on the “Signposting” theory of change include:

- DLIs were effective in providing signals to MoE civil servants responsible for their delivery, and in most cases the level or value of the DLI was not an important factor in this. DLIs became a central focus of annual sector reporting and dialogue for the program period.
- Tanzania had at the outset a “design lab”, with analysis and intensive discussions between government and donors, which ensured strong alignment. The high-level ownership of the BRN initiative varied over time with new political leadership. Learning goals had to compete with the legitimate but under-resourced priority given to expanding access with the FFBEP. During the 2017 redesign, there was less in-depth analysis and discussion of trade-offs than in the initial design process. Issues not addressed included important quality-focused measures such as in-service teacher training and regional equity.
- In Mozambique, the process of program development and negotiation was very focused. The MoE was effective at ensuring strong alignment with its own priorities. The PFM4R evolved out of government and World Bank

focus on PFM at sectoral and subnational level and DLIs provided a strong signposting effect on specific priorities, including for district and provincial education staff. In the case of the GPE DLIs there was a signposting effect for reducing PTRs (though not sustained). The MoE’s original goal with PTRs was to secure additional financing for new teachers from the MoF, but this was not achieved. Other GPE DLIs were likely already government priorities and may have happened without RBF.

- In Nepal, the large-scale of the SSDP RBF included DLIs in multiple thematic areas, with an array of new initiatives and priorities. Alignment or agreement with the specified goals of the DLIs by the MoE and its agencies at times was variable. Despite this, a strong signpost effect was present but with strain on the ability to achieve many DLIs in the context of the rapid and profound changes faced by government. Much of the focus was on reporting and for many process DLIs it is unclear that targeted reforms took place even when DLIs were achieved (for example targeted improvements to geographical equity in teacher allocation). Signposting was supplemented by technical assistance funded by development partners and this was often essential to achieve results.

S44. There were some cases of risks materializing:

- Results within sector frameworks that were not financially incentivized via DLIs in many cases had less focus. However, MoEs often pursued their own priorities even if they were not DLIs.
- The time dedicated to DLI reporting crowded out attention to other results.
- Equity in resource allocation and outcomes was an issue that DLIs did not systematically address. On occasions, DLIs may have had an unintended consequence of incentivizing more inequitable resource allocation.

“Autonomy and innovation” Theory:

RBF provides greater discretion regarding how to achieve specified results. Paying for results can allow the agent to work out the best way to achieve that goal and to use their local knowledge to innovate, if the principal leaves space for this.

Potential risks:

Loss aversion, with innovation curtailed, owing to perceived riskiness of actions in relation to RBF (non-) disbursement. The number of indicators, their “position” in the education result chain, and their detailed design may all be factors in reducing the space for policy discretion.

S45. A common theory for how RBF can lead to results is via greater autonomy of the agent to work out how to achieve them. The nature of this depends on the level of the results chain incentivized – for example a result at outcome level such as student learning will depend on an array of financing, strategies, policies and multiple actors to be achieved. Combined with inherent uncertainty and a range of exogenous factors, such outcomes are likely to be less under the control of the agent (MoE). Results at process or output level in contrast may incentivize a more specific set of actions, be under more control, and perhaps require less autonomy or innovation.

S46. The conclusions of this assessment on the “Autonomy / innovation” theory of change include:

- RBF programs mostly gave autonomy to central government in how finance was used. Though some looser conditionality targeted additional finance for the sector, it was only indirectly incentivized via RBF, in spite of this likely being critical to sector results and performance.
- Nepal’s SSDP had the largest number of DLI results covering the widest range of process, output and outcome targets. These were often sequenced – for example, design a scholarship scheme (process), implement the scheme (output), and improve net enrolment (outcome) – staggered over the five years of the program. This linear structure combined with detailed DLI definitions often limited autonomy in terms of which interventions government should pursue to achieve outcomes.
- The EP4R in Tanzania had a greater focus on outcomes than Nepal or Mozambique. The focus on learning provided a wide range of potential policy options to pursue. However, the relative autonomy implicit in this approach did not stimulate the policy learning/action cycle needed, and the program’s “missing middle” of outputs may have led to some inertia.

- The PFM4R in Mozambique was focused on four specific process DLIs, alongside targeted technical support. This limited its goals to some important reforms, for example improving school supervision by districts. In practice, such changes required cooperation and autonomy to achieve, particularly between the center and local levels of the education system.
- RBF incentives were combined with technical assistance, capacity development, joint sector dialogue, and other projects and technical support. This mix was likely critical to the delivery of many results. Technical assistance played other important roles, including in the complexities of RBF reporting and offsetting some of the costs to government in dealing with the reporting bureaucracy required for RBF.
- In each country, there was more autonomy and discretion at central level than for local government or schools. At local level, space for autonomy and innovation was constrained by restrictions in the use of RBF rewards and by insufficient technical support. For example, improving equity in teacher deployment would require local governments’ discretion, but often this was constrained in contexts with high centralized control, particularly with regard to decision-making in hiring and firing teachers.

S47. There were some cases of risks materializing:

- DLI frameworks could limit scope for innovation. When DLIs were achieved, successful reforms should have been built on to target subsequent reforms; but such a process was often absent.
- With complex and wide-ranging DLI frameworks, development partners often promoted their own priorities, rather than processes of learning and innovation. Complex and detailed DLI definitions could mean “micro-managing” reforms rather than allowing home-grown learning and innovation to take place.

“Sharpening minds” Theory:

RBF improves evidence and evidence-based policy. As more accurate measurement is required to document and verify results, RBF may increase the quantity and quality of evidence available and, also help to improve policy and planning.

Potential risks:

Compromised evidence systems or verification process, owing to the need for the RBF to be perceived as successful and funds to flow.

S48. An important claim for RBF is that it improves the availability and quality of evidence. This includes information gathered as DLI results, and the additional evidence and rigor from the verification mechanism. This in turn can improve policy learning and adaptation, an additional pathway to improve sector performance (‘sharpening minds’). There is a flipside risk in which there are declines in evidence quality due to the high stakes involved with disbursement and reputation linked to DLI results.

S49. The conclusions of this assessment on the “Sharpening minds” theory of change include:

- The RBF programs studied included DLIs to incentivize the strengthening of evidence systems, while also drawing on evidence from these systems. DLIs also created new evidence outside government systems. Some programs sought to introduce elements of subnational RBF using government evidence systems, but potentially before they were of sufficient quality, introducing risks in the process.
- Policy learning was not coherently incentivized by RBF; even in Tanzania, where a specific DLI was in place for this purpose. The evidence in DLI reporting directly from government was not supplemented with sufficient evidence from non-incentivized sources, or research and evaluation outside of the incentivized DLI framework, for example studies on “why” outcome results were what they were.
- A variety of approaches were taken with independent verification, including who hired the verification agent (IVA), how their scope of work and methods were agreed, the resources allocated, and the degree of donor involvement in the process. At least two donors did not require any formal verification for payments to be made for all DLIs. Each approach had limitations, and often the IVA was commissioned to provide a “tick-box” approach rather than a rigorous evaluation of results.

- Mozambique’s PFM4R used the national audit institution for the IVA role. This had the advantage of using government systems and strengthening national capacity. A challenge was the degree to which the role expected was a simple crosscheck of MoE data, as opposed to a more evaluative performance audit.
- In Tanzania, the contracting mechanism for the verification agent gave a relatively higher degree of independence from government. Combined with specified protocols this led to a detailed assessment of achievement and triangulation with qualitative evidence. All sides found this to be effective and it was a respected process.
- In Tanzania and Nepal, DLIs included a focus on learning assessments. This produced important evidence on learning performance; for example, NASA reports in Nepal are now regularly produced and include a high quality of subgroup analysis. It is less clear that this has led to specific improvements in policy.

S50. There were some cases of risks materializing:

- With the use of outcome DLIs, there were risks when ‘bad’ performance signaled failure on sensitive and complex results. Risks were to both transparency and rigor in measurement.
- There were cases where methods for assessing outcomes were not rigorous. This could create the risk of the impression that a complex outcome was achieved when the DLI result was achieved, when the reality remained unknown.
- There was tension between strengthening the DLI evidence, and developing a wider evidence base, including on whether results led to higher-level goals. For example there were gaps in analysis of drivers of out of school incidence in Nepal; constraints to accountability of school directors in Mozambique; and an absence of analysis on what caused retention to improve in Tanzania, which left stakeholders surprised when it subsequently deteriorated.

“Labeling success” Theory:

RBF induces better results-based planning. As goals are labeled as the results of interest, planning may focus more systematically on results – identifying bottlenecks to these results, and actions to address these, with associated indicators.

Potential risks:

*Lack of ambition if results are simple to achieve, or achieved within the time-scale of the plan period.
Financial flow risks because payment is typically in arrears or via non-achievement, meaning fewer resources for sector or unpredictability of fund flow.*

S51. RBF may also improve results-based management through the design process in which results are selected and the plan developed; and through the planning and reporting undertaken over time (‘labeling success’). As all parties need a coherent plan and specific, measurable indicators to contract upon, RBF design should help to identify the bottlenecks that need addressing. This can then lead to improved quality of multi-year sector plans and accompanying operational or annual implementation plans. It can also improve management of institutions and actors within the education system, for example with iterative review processes feeding back into planning and policy.

S52. The conclusions of this assessment on the “Labeling success” theory of change include:

- All three countries saw RBF reporting align to annual reporting cycles, following the structure of joint sector reviews which predated RBF and underpinned government to donor dialogue. This structured planning and dialogue approach was helpful to focus on results, though for bigger programs there was a risk that RBF crowded out dialogue on non-incentivized results.
- Collaboration in the RBF design process varied. In some cases, it was done in tandem and very closely aligned to the sector plan with DLIs selected from a wider sector results framework. In other cases, donor prioritization in planning and DLI selection was stronger. Where DLI frameworks were extensive in terms of the number and range of DLIs, the burden on planning systems was significant, reporting could take large amounts of time and resources, and this was not always sufficiently addressed. With more limited and targeted frameworks such as that of the PFM4R in Mozambique, the results-based planning approach was clearer, although potentially also less ambitious.

- In all three countries, the use of technical assistance alongside DLIs supported the planning and management of reforms as well as relationship management with the development partners. This support was often needed, but was also more often focused on DLI reporting than strategic sector planning.
- Given that basic education management is decentralized or deconcentrated, an important part of the “labeling” effect and of strengthening result-based planning depends on how this is being transmitted to subnational levels. RBF programs engaged with these levels through existing systems but did not usually give enough support or attention to subnational planning processes, and in some instances created ad hoc systems for implementation or reporting, with a very specific focus on DLIs.

S53. There were some cases of risks materializing:

- Financial risks were found and even though payments were often deferred or ‘rolled-over’, due to resource constraints even delayed payment could lead to less resources for the education sector in a given financial year.
- Detailed DLI definitions for processes and outputs meant, particularly in Nepal and Mozambique, that programming was at times “micro-managed”, leaving less room for learning and adaptation or planning towards outcomes.
- RBF “labeled success” but this could become an “indicator fixation”, in which the strategic underpinning of interventions set out in DLIs was at times perceived to be lost. This focus could take away time for problem-solving on how to remove key bottlenecks in the system and this was linked to what some government stakeholders called a “target obsession” of development partners linked to the need for DLIs to be achieved and funds to flow.

“Aligning all actors” Theory:

RBF improves coordination between key institutions. As results are clearly identified, the policy dialogue is more focused, with ensuing collective action made more likely by the incentivized nature of the framework.

Potential risks:

Dispersion of actors and coordination; for example, donors focusing on “their DLIs” as opposed to the overall sector plan.

S54. An important theory for RBF also linked to the planning process is that it may improve the coordination of stakeholders within the sector. This can be by providing a focus for improving dialogue around results between development partners and government, or in improving alignment and coordination among national government agencies and subnational levels. The use of RBF can also lead to changes in planning and management mandates as well as the PFM systems underpinning central and local level relations.

S55. The conclusions of this assessment on the “Aligning all actors” theory of change include:

- The alignment of sector actors did not automatically ensue from RBF but required attention to ownership of strategies to achieve results. At its best, consultation was broad-based, included all relevant actors, and time was taken to build consensus. Where development partners promoted reforms without sufficient buy-in on the government side, they were either not achieved, or achieved but only in a narrow sense (for example DLIs “achieved” but without associated deeper reforms).
- In Nepal, the RBF period was concurrent with the “big bang” of federalism in which the delegation, financing, motivation and reporting functions of different tiers of government underwent radical change. The new tier of local governments became responsible for most basic education management tasks. Federalism heralded greater autonomy but also led to gaps in civil service positions and technical capacity. The DLI framework was adapted, but took time to be made relevant to the new reality.
- In Tanzania, there was a sophisticated program structure, with RBF-financed subnational payments to LGAs for DLIs they were directly responsible for. This may have secured greater buy-in from local governments, as well as greater awareness of the DLIs at local level than in either of the other two countries. However,

this awareness did not appear to extend to DLIs attributed to the national level, even though some, such as the learning DLIs, also required local and school-level attention and action.

- In Tanzania, strong program governance from development partners and government, supported the EP4R’s coherence in spite of the priority shift to access with the FFBEF. Within government the program contributed to greater MoE and PO-RALG collaboration; and to improved coordination between national and local levels, although from a low base, and with a narrow focus on the DLI framework.
- Mozambique had a more stable, top-down structure of relations from the MoE to deconcentrated provinces and district services. The PFM4R was well-designed to work within this structure, and also played a role in ‘aligning actors’ by bringing together the MoF and MoE, with a focus to improve timeliness of fund flow for school grants in particular.

S56. There were some cases of risks materializing:

- There was at times a power imbalance between the donors and the government in negotiating changes to DLIs. Government was at the behest of development partners’ priorities and internal systems to ensure that DLIs could be changed, even when significant contextual changes made this essential.
- The membership of third party voices including civil society, NGOs, and advocacy organizations within joint sector fora, provides potential for enriching dialogue on sector results. There was at times a risk that simplistic narratives of success linked to DLI achievement could undermine these voices, framing their nuanced views as “dissident” and legitimate to dismiss.
- A combination of PFM, core financing, and political economy constraints were behind many of the issues addressed by DLIs, but were rarely discussed. This can create a degree of confusion and lack of transparency in sectoral discussions.

“Sustaining attention” Theory:

RBF maintains the focus on specified results during a period of change through its contractual nature, more than would be possible with a traditional (non-contractual) results framework.

Potential risks:

Inflexibility: inability to deal with or adapt to changing context over time, because of the fixed nature of the RBF “contract”.

S57. The final theory for RBF considered is that the fixed nature of an RBF contract maintains a focus on results over time. This may be beneficial in periods where the political landscape is changing, in which for example RBF may “sustain attention” on quality or on process reforms seen as essential to improving sector performance. The flipside risk to the fixed nature of the RBF contract with targets set up to five years ahead, is that DLIs may become more challenging or unrealistic, or they may become irrelevant over time. This could imply incentivising irrelevant or even counterproductive reforms.

S58. The conclusions of this assessment on the “Sustaining attention” theory of change include:

- All three countries experienced major changes to context during the programming period. Most pertinently, there were changes to the political leadership and associated priorities in Tanzania in 2015; the federalism changes in Nepal in 2017-18 and ongoing; and in Mozambique, the suspension of GBS in 2015-16, which reduced the fiscal space for the education sector.
- In practice, RBF programs have room for flexibility. The Tanzanian RBF was in two main phases with a mid-term review (MTR) leading to new areas of prioritization from 2017. The SSDP in Nepal also underwent a MTR in 2019 which was a key point in addressing the realism of program targets in light of federalism. In both countries, DLI targets were changed once formally agreed between development partner and government, but this could take time.
- In all three countries, there were cases in which non-payment occurred. While this demonstrated the “credibility” of the RBF mechanism, there were many cases where a combination of rollover provisions, program restructure, flexibility in interpretation, and redesign, were applied. This meant the value of disbursements was largely maintained, and indicates that the need and/or desire for disbursement could override considerations of credibility.

- In Mozambique, RBF was used over a relatively shorter period (mostly during 2015-17), reducing the risk that the focus might shift over time. Despite this, the suspension of budget support constrained the national budget, with likely implications for DLI targets such as PTRs in terms of hiring more teachers. While there may have been a ‘sustaining attention’ effect for the program period, there is less evidence this lasted once the RBF programs were complete.
- In Nepal, despite the SSDP moving into a ‘transition phase’ from 2018, it took two more years to change many of the DLIs that were either no longer relevant or substantially more difficult to achieve due to federalism. It is possible that the focus on DLIs may have distracted attention from deeper realignment to the new institutional reality.
- In Tanzania, the EP4R RBF focus on quality and learning had to compete with the non-RBF priority of increasing access. The latter was a change not even two years into the program, but despite this the RBF did ‘sustain attention’ on learning and quality in the 2017 revisions to DLIs. The effect was mixed, as government prioritized access by using RBF rewards to fund school infrastructure, over a number of possible more quality-focused actions.

S59. There were some cases of risks materializing:

- The contextual changes in Nepal and Tanzania did not lead to a coherent reassessment of the “stretch” or “risk level” of the program. MTRs could have provided a systematic reappraisal but were not always planned in this way.
- Flexibility in changing DLIs and in the interpretation of detailed DLI definitions may have avoided non-achievement of many DLIs. However, the process itself can be time-consuming, and may have been a distraction from the broader effort needed to adapt to major changes to the contextual environment.

Lessons and recommendations

S60. The majority of RBF found in each case was from aid-financed RBF programs, and these programs have become the main focus of the assessment including the Final Assessment Reports for each country,⁷ the findings of which this report has synthesized. Whilst some of the insights proposed may be useful to RBF programs that would be run by governments and with government financing, this is not our aim here. In a bid to be useful to aid-financed RBF practitioners on both donors and governments' sides, the suggestions we make are based on the fact that such RBF programs

always entail a donor-government relationship. We also aim to be practical, at the same time recognizing the importance of assessing the relevance of any suggestion in the particular context one is dealing with.

S61. Table S2 below sets out the recommendations and connected lessons that we hope will help those involved in either the scoping, designing, implementing or evaluating of RBF programming, to think through some of the major challenges and try to maximize the opportunities of RBF, while also minimizing the risks.

⁷ Final reports for Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021).

Table S2 Lessons and recommendations

Recommendation	Connected lessons
1. Be clear why you are using RBF while anticipating and mitigating risks	1A. There are multiple objectives of RBF, and some of them may compete, so priorities should be made clear from the outset.
	1B. Risks that are specific to RBF – information risks, motivation risks, equity risks – should be set out in the design phase, with clear strategies for mitigating each risk.
	1C. Financial and non-disbursement risks come from delays and under-performance. Risks can be offset by rollover and flexibility, but with implications on the credibility of RBF as a contractual mechanism.
2. Identify, discuss, and agree how RBF can be part of a mixed portfolio of approaches	2A. Targeted technical assistance may be needed; and this should link to clear planning on its contracting, its role and how it is structured.
	2B. RBF is only one part in a portfolio of sector support. Donors should be clear on the mix used, and plan for complementarities.
	2C. RBF alone will not address broader governance and political economy bottlenecks: a broader set of approaches are likely to be required.
3. Have clear results chains identifying the changes and goals targeted, including a balance between feasibility, desirability and flexibility	3A. While DLIs target specific results, these are also likely to link to higher-level sector goals to which the RBF design is contributing, alongside other initiatives. Ideally these goals should be made explicit at design, monitored as part of the program monitoring framework, and considered in the light of results from the other initiatives.
	3B. There can be a trade-off between using the DLIs to map each level in the chain towards outcomes (specific processes, outputs etc.) vs. fostering ownership, autonomy and adaptive management.
	3C. There may be additional costs, and increased complexity, of “larger” DLI frameworks, and when very large, there is a risk of crowding out other important results from sector dialogue.
4. Assess, discuss and explicitly agree on the degree of stretch including a focus on complexity, and the potential trade-off with disbursement	4A. Results vary in the degree of stretch for each DLI, although actors often have no consistent way of assessing it. Ideally it should be explicitly assessed and methods considered could include assessing the degree of social and technical complexity to achieve the result.
	4B. The likelihood of achievement of result targets in light of their stretch, will have implications in terms of incentives including risks of demotivation. The perception of how much stretch there is, is also likely to change over time.
	4C. Disbursement is a legitimate objective of any financing arrangement; discussions between donor and government should be clear on the trade-off in likely rates of disbursement when there is higher stretch.
5. Analyze and discuss how RBF will work with and reach out to all actors in the sector (including schools and subnational government)	5A. Significant results require a large number of agents to act, and cooperate; program documents should spell out plausible reasons why the incentive should ensure agents work together. The accountability arrangements in the sector should be continually assessed, as they are not static.
	5B. Local-level stakeholders’ engagement and action is essential to achieve results, but mechanisms may vary in how much voice and autonomy they have to act and contribute. Where they have little of both voice and autonomy, RBF may not enhance this, but should seek to strengthen the work of local stakeholders and combine with non-RBF approaches to local accountability.
	5C. Consideration should be given to how the RBF design, including DLI selection, supports and/or interacts with <i>de facto</i> and <i>de jure</i> levels of decentralization.
6. Build adaptive management in from the start	6A. The majority of DLIs have a “sell by date”, i.e. may become redundant, irrelevant, or counter-productive, and this is most true for process and output DLIs. Actors should be proactive when there is a need to remove DLIs, revise targets, or add new DLIs.
	6B. A formalized, multilateral, transparent and regular process helps to enable adaptive management; where this is not present, the processes for making changes to DLIs can be ad hoc and take time.
	6C. The method to react to <i>force majeure</i> situations could be more clearly stipulated – i.e. external or large-scale changes to context that have effects across all DLIs – and with processes to manage change across multiple donors. Any of the parties, including government, should have the means to trigger a reassessment in light of any new <i>force majeure</i> situation.

Recommendation	Connected lessons
7. Seek to use RBF in line with principles of good donor coordination	7A. Multiple donors may be involved in RBF for a particular sector plan, and should work together in design to reduce contradictions and/or duplication for government, for example with program or indicator reporting.
	7B. Where multiple donors and sectors use RBF at the same time, the risks to government in terms of finance flow can be multiplied. Donors should ideally work together to understand the collective risks that RBF poses including whether the scale of use is proportionate.
	7C. RBF links to, aligns with and potentially can strengthen the role of sector coordination and dialogue, including with non-RBF donors and projects. However, there is a risk that RBF/DLIs can come to dominate dialogue when used by multiple donors in a way that crowds out other important results.
8. Design the interaction of RBF with evidence systems so as to ensure monitoring of program performance beyond DLI reporting	8A. Plans and strategies to improve data systems are likely to be in place but with objectives beyond just RBF. The use of RBF should be clearly incorporated in these plans including data collection, the use of data in planning and policy, and the trade-offs and risks in creating high-stakes data.
	8B. The environment for communication and transparency in evidence systems and processes is a key component for how much learning and improvement to policy is likely to flow from RBF.
	8C. The use of a broad base of evidence in assessing program performance can be vital in understanding sector trends, including non-incentivized information, evaluation, and learning processes.
9. Carefully design and fund the verification system	9A. It can be challenging to ensure the independence of verification, including in the designs for the scope of work, processes of reporting, and joint management and buy-in from government and donors. A variety of approaches have been used with different trade-offs.
	9B. The amount of funds allocated, and the source of funds, to the verification process should ideally be clear and transparent at all stages. However, this is often not the case. Sufficient funds should be considered or face trade-offs with quality, sample size, rigor, and the timeline.
	9C. The verification role may or may not include a role for qualitative research and evaluation, i.e. going beyond strictly “verification”. For this to be effective clear expectations are required as well as incentives for all parties to learn from the experience rather than treating it as a bureaucratic exercise.
10. Use a combination of process, output and intermediate outcome indicators in a way that helps reformers	10A. Process reforms that are socially complex and require extensive change at local level, take time and require alignment and buy-in from a wide range of country stakeholders. They may not easily lend themselves to time-bound DLIs.
	10B. Output indicators require clear processes of design and clear project management responsibilities, as well as understanding and measurement of output <i>quality</i> in addition to more common measures of <i>quantity</i> .
	10C. For results that constitute intermediate outcomes, it may be better to have a clear and simple metric (e.g. PTRs) than complex process, output or proxy indicators.
11. Be careful, and frugal in the use of outcome indicators.	11A. For outcome indicators, their complexity and often low degree of control and knowledge on how to achieve them bring particular risks (stretch too high, risks to disbursement, and demotivation and/or to transparency). While outcomes are desirable to achieve, careful attention should be paid to these risks.
	11B. Outcome indicators can be difficult to define and measure, particularly where only one metric is used. Incentivizing the measurement of outcomes brings lower risks than incentivizing the results of outcome measures; and may therefore provide a better basis for policy learning.
	11C. Given their importance, non-incentivized measures of outcome indicators are vital, as is transparency in discussion and learning on these results.
12. Pay continual attention to equity	12A. Equity should be a consideration in RBF design and implementation including in terms of the different capacity of agents, particularly at subnational level. There can be structural inequalities in resource allocation by geographical or other criteria and there is a risk that RBF can exacerbate them.
	12B. Program-level implications for equity during design should be treated as a “crosscutting” programmatic issue, and continue to be so during implementation.
	12C. While it may be beneficial to have DLIs specifically addressing equity, it is also important to look out continually for unintended equity consequences of all DLIs.

1. Introduction

1. This report has been commissioned by the World Bank Results in Education for All Children (REACH) program⁸ as an assessment of the design, implementation experience, and impacts of RBF in education. The study covers three countries:

- **Nepal**, which following piloting of RBF by the World Bank in the 2013–15 period, shifted to receiving the major share of donor finance to education via RBF, funding the government’s School Sector Development Plan (SSDP) from 2016. RBF financing to the sector has six joint financing partners involved, and is worth USD 400 million, covering over 80 percent of potential external finance to the education sector in the period from 2016 to 2021.
- **Tanzania**, which introduced RBF alongside the “Big Results Now” initiative in 2014, with sector finance linked to results for the World Bank, the UK’s Department for International Development (DFID) and the Swedish International Development Cooperation Agency (Sida). RBF programming in the period since 2014 has been linked to over USD 500 million of potential disbursements to the sector.
- **Mozambique**, which introduced RBF with the World Bank’s Public Financial Management for Results program from 2014 to 2018, and has in the same period had RBF used by the Global Partnership for Education (GPE) and Germany (KfW). The total value of RBF to date in Mozambique has been smaller than in Nepal or Tanzania, at just under USD 60 million, or 12 percent of donor financing to education from 2014 to 2018.

2. The research has been undertaken in two phases. The outcome of the first phase was a **Preliminary Assessment Report** for each country, finalized in February 2020, following two-week country visits in late-2019. This phase provided a comprehensive “mapping” of RBF in education and important elements of the broader country context. **Final Assessment Reports** for each country are the conclusion of the second phase, with more in-depth analysis on the contribution of RBF to education results including a comprehensive review of the experience with RBF programming in education, including with further qualitative interviews undertaken over the course of 2020, and quantitative analysis on a smaller number of RBF mechanisms / disbursement linked indicator (DLI) areas selected.

3. The research led to this final **Synthesis Report** to compare and contrast the three country experiences, and identify findings and lessons that could inform education programming and the use of RBF more broadly. This Synthesis Report draws on the findings across the research and summarizes the most important and insightful findings, however any reader looking for the more comprehensive review of each country experience will find the country reports useful, including full sets of references, stakeholders consulted, and the range of findings from qualitative and quantitative research undertaken for this review.⁹

1.1 Theories of RBF underpinning the research

4. The term *results based financing* (RBF) covers a range of financing instruments that have been increasingly used in education programming, as well as in other sectors, in recent years. RBF instruments are defined as those paying for a result to be achieved *only* if achievement can be credibly verified. In education, RBF includes forms of contract that disburse on results such as, for example, children enrolling and learning, or

⁸ REACH is a multi-donor trust fund established in 2015 and funded by the World Bank Group, Norway, Germany and USAID. The main purpose of the program is to contribute to the evidence base around RBF in education.

⁹ The Preliminary Reports are referenced as, for Tanzania, Dom et al. (2020); for Mozambique, Patch et al. (2020); and for Nepal, Holden et al. (2020). The Final Assessment Reports are referenced for Tanzania as Dom et al. (2021); for Mozambique, Patch et al. (2021); and for Nepal, Holden and Chapagain (2021).

the number of classrooms constructed. This contrasts with instruments that transfer resources unconditionally or simply pay input costs, for example salaries, or materials for construction.

5. RBF is a broad term that can cover uses such as performance-based grants to schools or local authorities, performance-based pay, development impact bonds, or conditional cash transfers, among others. RBF has, however, mainly been used as a financing modality by international aid donors. This has included the growing use of financing instruments such as Program for Results (PforR) by the World Bank, Results-Based Lending (RBL) by the Asian Development Bank (ADB), and the use of a variable tranche by both the European Union and the GPE. All link disbursements to results. RBF approaches have also increasingly been used by bilateral aid agencies such as the UK's Foreign, Commonwealth & Development Office (FCDO) and Sida.

6. As an aid modality, the conditional nature of RBF provides an alternative to most uses of general budget support (GBS) or sector budget support (SBS). From the early 2010s, donors found these modalities problematic, with a number of evaluations raising questions on the effectiveness of GBS and SBS in addressing systemic problems faced in key sectors such as education.¹⁰ RBF modalities can therefore be contrasted to more "traditional" input-based financing modalities,¹¹ where activities and expenditure are accounted for more directly, independent to results achieved or with much looser conditionality. Input-based financing modalities also tend to include greater fiduciary reporting requirements.¹²

7. RBF has also been used within countries in performance-based transfers of resources from central to subnational levels of government, as well as in performance grants to schools, and to teachers in the form of performance-based pay. There is often a link between donor-led RBF and "downstream" RBF between central and subnational levels of government – and this can be in-built to the design of the programming. This is the case in each of the three countries studied, as this report will discuss.

8. The theory of RBF is often framed in terms of principal-agent theory. The *principal*, seeks to achieve a specific goal by incentivizing the recipient of funds, the *agent*, to undertake the activities required to achieve it. The theory provides some clear ways through which RBF may work – broadly these can be split into the idea the agent will put more effort or resources into achieving the specified results [*steroids*]; that the goal will see effort diverted or resources prioritized [*signposting*]; and that paying for results as opposed to inputs could give the agent more autonomy to achieve results [*autonomy and innovation*]. Proponents of RBF also claim other ways in which it may improve education systems. These include improvements to evidence systems [*sharpening minds*]; results-based planning [*labeling success*]; bringing actors together in cooperation and coordination to achieve the goal [*aligning all actors*]; and finally, by "contractually" defining results, RBF provides a mechanism to focus on the goal over time, whatever else may be changing [*sustaining attention*].¹³

9. RBF also faces risks and potential unintended consequences. Some link to the inherent challenges of measurement and the information asymmetry in which an agent knows more about what has really been achieved than the principal; and that any measure or metric is always likely to be imperfect. There are therefore risks of "gaming" in which the result is reported as being achieved when this is either not true [*cheating*] or only true in a narrow sense [*fudging*]. For certain types of indicator, there can also be equity risks where subgroups of individuals are easier to reach, and prioritized to achieve the target, possibly at the expense of harder-to-reach subgroups [*cherry-picking*]. RBF mechanisms include third party verification in order to address some of these risks, however any verification process is also subject to the same risks and challenges with measurement.

¹⁰ See, for example, DANIDA (2014) for a summary.

¹¹ This includes instruments such as Investment Project Financing (IPF) for the World Bank, Project Based Lending (PBL) for the ADB, and the use of a *fixed tranche* by the European Union and GPE.

¹² While RBF programs are distinguished as a modality in the use of conditionality, they can vary in degree of difference to other modalities. This includes the rigor of evaluation required; the degree of alignment with government financial reporting and accountability systems; the presence of safeguards; the role of technical assistance; and whether or not they use existing basket or pooled funding mechanisms.

¹³ These theories are based on a large literature survey, including the work of REACH, in World Bank (2017a).

10. The theories of how RBF may work, together with some of the countervailing risks are summarized in Table 1. These theories are used to structure the final findings and Conclusions in Chapter 4 of this report.

Table 1 Theories of how RBF works and associated risks

RBF theory	How it works ...	Risks
Steroids	RBF induces more effort (and/or resource) to achieve specified results. A goal is more likely to be achieved if a payment is made on its achievement, because more effort will be put in to achieve the goal by agents.	Gaming risks relating to the RBF metric, such as <i>fudging</i> where success is in appearance only. Demotivation by crowding out other accountability mechanisms.
Signposting	RBF focuses effort (and/or resource) to achieve specified results. Within finite limits to efforts/resources, allocations of effort/resource to specified results reduce those to less important areas.	Diversion risks if there is a suboptimal resource allocation, with important areas deprioritized because of the RBF incentive. Cherry-picking risks, in which less costly or politically favored subgroups or regions see greater resources, in part because they are more likely to reach a target.
Autonomy and innovation	RBF provides greater discretion regarding how to achieve specified results. Paying for results on a goal can allow the agent to work out the best way to achieve that goal and to use their local knowledge to innovate, if the principal leaves space for this.	Loss aversion , with innovation curtailed, owing to perceived riskiness of actions in relation to RBF (non-) disbursement. A large number of indicators reduces the space for policy discretion.
Sharpening minds	RBF improves evidence and evidence-based policy. As more accurate measurement is required to document and verify results, RBF may increase the quantity and quality of evidence available on results and, further, help improve policy and planning.	Compromised evidence systems or verification process, owing to the need for the RBF to be perceived as successful and funds to flow.
Labeling success	RBF improves results-based planning. As goals are labeled as the results of interest, planning may focus more systematically on <i>results</i> – identifying bottlenecks to these results, and actions to address these, with associated indicators.	Lack of ambition if results are simple to achieve, or achieved within the time-scale, with years still to go to the end of the plan period. Financial flow risks because payment is in arrears or via non-achievement, meaning fewer resources for sector or unpredictability of fund flow.
Aligning all actors	RBF improves coordination between key institutions. As results are clearly identified, the policy dialogue is more focused, with ensuing collective action made more likely by the incentivized nature of the framework.	Dispersion of actors and coordination; for example, donors focusing on “their DLIs” as opposed to the overall sector plan.
Sustaining attention	RBF maintains the focus on specified results during a period of change through its contractual nature, more than would be possible with a traditional (non-contractual) results framework.	Inflexibility: inability to deal with and adapt to changing context over time, because of the fixed nature of the RBF “contract”.

Source: Authors, drawing on a large literature survey, including the work of REACH, in World Bank (2017a).

11. While principal-agent theory provides insights as to why RBF might work and its risks, alternative theories¹⁴ and empirical literature, point to other risks of RBF. Notably people and institutions respond to other forms of accountability than just financial incentives – for example electoral accountability, peer-to-peer accountability, and intrinsic motivation. As a result, there can be risks of *demotivation*, where the introduction of a monetary incentive crowds out other incentives for carrying out the same actions; and *loss aversion*, in which actions are taken with perceived low risk due to the high stakes of the RBF incentive. Potential unintended consequences also include the risk of *diversion*, in which the RBF incentive leads to deprioritizing other important

¹⁴ This includes *stakeholder theory* – individuals’ attitudes are not always selfish / relationships are interdependent, that is, principal and agent can influence each other; *institutional theory* – individuals adapt to the system of norms, values and beliefs within their institutional environment; and *prospect theory* – individuals base their decisions on subjective perceptions of change, and may aim more to minimize loss than maximize gain. See Cuevas-Rodríguez et al. (2012).

goals in a way that is suboptimal. Finally, there are risks relating to the financing mechanism itself, that is, the flow of finance linked to the RBF mechanism. As payment is generally *in arrears*, but the costs of activities are experienced *up-front*, if results are not achieved non-payment may lead to shortfall in finance that may have knock-on effects.

1.2 Change in education systems and levels of results

12. Results-orientation and policy prioritization in education implicitly start and end with the student at the center. Thus, outcomes and impact of the education system are measured through the results for students, in their progression, their learning and their life chances from education including the skills and qualifications they attain. While there is some degree of consensus around certain measures of student goals – for example basic literacy and numeracy skills for children in primary education as a pre-condition for effective secondary and further education – no single measure will capture all of the goals, outcomes or impacts that we care about. Notably, student well-being in terms of safety, physically and mentally, is the foremost concern of any parent, but is often not captured by quantitative metrics of education systems.

13. Together with student learning and life chances, a number of other areas are critical and can be framed as other *outcomes* of the system – for example, access to school, including attendance, retention, and progression; and equity according to gender, ethnicity, or disability. *Intermediate outcomes* – for example, the quality of service delivery in terms of the teacher’s presence in the classroom, and their capabilities, alongside manifold *processes*, *inputs*, and *outputs*, from teacher training facilities, to the condition of classrooms, provision of learning resources such as textbooks, and set of accountability systems, standards, and supervision, underpin and ultimately determine the performance of the system in terms of student outcomes.

14. The education system includes these inter-related subsystems that drive the results for students. RBF can involve results across this spectrum, and the level of results that is targeted interplay with some of the theoretical drivers of benefits and costs of RBF. This includes the types and interrelationships of institutions involved in education delivery, including the ministry of education and central agencies, and subnational levels of government, and down to the school, the teacher and in the classroom where learning ultimately takes place. Table 2 below highlights the common definition for these different “levels” of DLI within the education system, and particular types of indicator that might be tackled. Further delineation of the types of indicators that can be chosen for RBF is set out in Box 1.

Box 1 Levels of indicators in education systems

Results-based financing (RBF) involves the selection of specific indicators – metrics or measures – that will underpin disbursement. Often known as disbursement linked indicators (DLIs), these form the primary focus for the goals or changes that the RBF seeks to incentivize. As these changes or goals form part of education sector strategies and planning processes they can focus on different levels of the system as well as different actors within the system. The classification of where a particular DLI focusses on is often separated between *outputs*, *processes*, *intermediate outcomes* and *outcomes*. The general assumption is that RBF will not specify particular *inputs* or *activities* as these are the means to achieve a result rather than the result itself.

Donor financing using RBF, such as that explored in this assessment, tends to include DLIs across a mixture of these levels. This is a reflection of the many types of changes and results targeted in the complex setting of an education sector plan. This adds some complexity to the theories of RBF in that there is an implicit (or explicit) causal logic between different levels – i.e. that *these* process and output changes will lead to *these* intermediate outcomes, and in turn this will lead to *these* outcomes for students. For example, improved teacher training institutions [process], with numbers of teachers trained [output], leading to improved classroom practices [intermediate outcome], in the hope this will improve learning of students [outcome]. In the case where all of these levels are made to be DLIs within an RBF program, this is likely to limit the theory of change on autonomy and innovation – i.e. in the case above the strategy to achieve greater learning is already set out, and the strategy itself may be incentivized by DLIs.

Table 2 Levels of metrics used in RBF in education

Level of indicator	Definition	Examples of area of focus	Important considerations
Process	A change to sub-systems of education that determine how resources are managed, people and institutions are held accountable, or information is generated.	Grant management system Accuracy of EMIS data Curriculum framework	The number of people and institutions involved in process changes may determine complexity. Some processes are likely to link to political economy considerations.
Output	The provision of a good or service that is required for delivering education but is far from a predictor of success of the system.	Number of textbooks delivered Number of teachers trained	Degree to which there is a measure of the quality of the output as well as its quantity.
Intermediate outcome	An indicator that is a strong predictor of student outcomes but that does not constitute a goal in itself of the education system.	Pupil teacher ratios Teachers employing improved classroom practice	Measurement can be very difficult and costly, particularly to be nationally representative.
Outcome	A clear achievement of a goal for the intended beneficiaries of education.	Learning outcomes Gross or net enrolment ratio	Averages need to be disaggregated to subgroups to understand equity.

Source: Authors

1.3 Research questions and study methodology

15. This assessment seeks to explore the experience of each country with RBF in education in the period since 2013, addressing the following overarching evaluation question: ***“To what extent, in what ways, and under which conditions has RBF helped education systems to deliver results?”*** The question is unpacked through a number of research questions that frame the study. These include:

- A. Design** – What goals did RBF seek to address? What indicators were chosen and at which level of the results chain? What was the design process? How was RBF envisioned to lead to improved results and were other elements of design, such as technical assistance, used?
- B. Results and impact** – What were the results achieved via the use of RBF? To what extent did the achievement of indicators represent the achievement of underlying goals? What was the contribution of RBF to the achievement of indicators and goals? Were the results of RBF sustained over time and beyond the period of RBF programming?
- C. Financing, payments, and financial flows** – What were the financial flows and payments made linked to RBF? Was RBF financing additional to education sector finance, or would financing have been still provided by other financial mechanisms without RBF? Were financial incentives passed on to different tiers of government?
- D. Co-ordination, capacity and accountability** – How did RBF relate to or change the coordination mechanisms between donors and governments? Did RBF change accountability relationships between central and subnational levels of government? How important was the role of capacity development, including technical assistance, in the results of RBF?
- E. Evidence-based policy and verification** – How was RBF used to address evidence systems? Did the process of results generation and verification built into RBF mechanisms lead to improved evidence of results? Did changes in evidence systems lead to changes in national systems of accountability and feed into improved education sector programming?
- F. Adaptiveness and ongoing change** – Was RBF, either at the level of individual indicators or the overall programming, adaptable to either the context, the perceived ambition of targets, changes in policy priorities, or perceived adequacy of metrics? Did adaptiveness and ongoing change in RBF represent an appropriate use of flexibility?

G. Risks and unintended consequences – Were there specific cases of perverse or unintended consequences linked to RBF for specific goals or indicators of change (for example, diversion of effort, gaming, etc.)? Were there other unintended consequences of the RBF programming linked to the overall use of RBF as a modality (for example, volatility of financing flows, quality of evidence systems, level of flexibility of programming, etc.)?

H. Cost-effectiveness – What were the major costs associated with (or attributable to) the use of RBF? What were the major benefits? What effect has RBF had on the overall cost-effectiveness of spending on education?

16. This final report includes the overall findings from both phases of research and is structured around these research questions. To address them, the study team has used a combination of qualitative and quantitative methods. This has included extensive documentation review, quantitative analysis of results and financial flows, and interviews with a wide range of education sector stakeholders including in government at central and subnational levels, as set out in Table 3.

Table 3 Study methodology

Method	Details
Education system review	Detailed appraisal of each country education system, recent historical and political context, sector performance and trajectories, and review of sector plans and evaluations.
RBF ‘mapping’	Identification of all RBF schemes in education (covering primary and secondary level). Analysis of how they interact, and a detailed review and mapping of all DLIs in the EP4R in Tanzania; the PFM4R, GPE variable tranche and German (KfW) use of RBF in Mozambique; and the use of RBF in the SSRP (2009-16) and SSDP (2016-21) in Nepal.
Key informant interviews	Interviews with: <ul style="list-style-type: none"> • Central government stakeholders at MoE, MoF, and education agencies. • Development partner and civil society stakeholders. • Provincial and local government stakeholders. • School principals and school council chairs.
Quantitative and qualitative review of DLI performance	Review of DLI results, disbursement and financial flows; assessment of their complexity; measurement and risks that emerged. Detailed assessment of design, experience, evidence and timing of results, and links with non-incentivized measures of success and overall sector performance.
Quantitative analysis of subset of DLIs	Deeper qualitative and quantitative analysis undertaken on a selection of DLIs, notably the GPE pupil teacher ratio (PTR) DLI in Mozambique; the out-of-school children enrolment DLI in Nepal’s SSDP; and the PTR and early learning (literacy) outcome DLIs in Tanzania.
Assessment against theories of change of RBF	Analysis of RBF contribution to results, the presence of incentive effects and behavioral responses, and broader systemic benefits and costs of the use of RBF.

17. The research set out to look for any and all RBF in basic education in the three countries.¹⁵ The mapping of RBF found that despite some selective use of subnational RBF beyond these programs, the majority of RBF in terms of financial value and scope has been linked to donor aid. Where subnational RBF was found, it was mainly linked to the larger sector programs with donor support, i.e. as an incentivized component of the programs as opposed to purely ‘home grown’ uses of RBF. It is because of this finding that the research has mainly focused on the large-scale uses of RBF as an aid modality in each of the three countries, and on which this Final Synthesis Report also principally focusses.

¹⁵ The study did not cover higher education or early childhood education.

1.4 Study limitations

18. The assignment has been affected by the COVID-19 pandemic, with the outbreak occurring between the first and second phases of the work. This means that the original plan to travel to each country for the second phase was curtailed, and interviews were carried out in large part remotely via telephone or the internet. While the second phase of research was still able to cover a wide range of stakeholders, including at subnational levels of government, the pandemic inevitably limited the ability to further probe some of the sampled RBF mechanisms in more depth – in particular to triangulate the quantitative results of RBF with the more qualitative perspectives of local stakeholders, including at the school level.

19. The research followed a process evaluation approach, and therefore aimed to find plausible contribution of RBF to results and systems change. Given the complexity of education systems, and each country context in terms of political economy, public financial management (PFM) systems, and a range of external and exogenous factors and events in the period studied, there is always a challenge around distinguishing factors of causality. The approach taken was to use the different sources – key informant interviews, document review, data analysis and regression analysis for a subset of DLIs – to closely analyze the experience of all DLIs within the RBF programming, while also looking beyond them to data and research that was not incentivized by this programming. At times stakeholders had different perspectives and opinions on how change had occurred and critically whether changes were positive or negative for sector performance. We have aimed to reflect the diversity of views, though inevitably at times we have taken a position based on the rounded view of evidence in terms of plausible contribution.

1.5 Structure of this report

20. This introduction has attempted to provide some important contextual and background theoretical factors that a reader will need to understand to follow the rest of the report. Chapter 2 – will set out brief background on Nepal, Mozambique and Tanzania, on their education systems and results, and will introduce the main uses of RBF in the period from 2013 to 2020. The central findings section – Chapter 3 – follows the evaluation framework and questions set out above. Chapter 4 then draws conclusions from the research on each of the major theories for RBF set out in Table 1 of Section 1.1 above. Finally, Chapter 5 draws together the findings from the research into some key lessons and recommendations for stakeholders involved in the design and implementation of RBF programming.

21. The country reports provide a wealth of information including the detailed country findings, as well as annexes with lists of documents consulted, stakeholders interviewed, quantitative analysis conducted, and the detailed DLI results. These reports can be viewed as annexes to this Synthesis report.¹⁶

¹⁶ The final reports are for Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021).

2. RBF in Education

22. This assessment draws together three countries which have had some of the earliest experiences with results-based financing (RBF) in education. RBF was introduced as part of external financing to the education sectors of Nepal, Tanzania¹⁷ and Mozambique in the 2013–14 period and these countries have therefore been at the forefront in the shift in sector financing. Since its launch in 2012, the World Bank has developed a portfolio of USD 6 billion for its Program for Results (PforR) lending instrument for education, including use in all three countries of this assessment. In 2013, the Asian Development Bank (ADB) launched its Results-Based Lending (RBL) instrument and since developed a portfolio of USD 2 billion for education across seven countries including Nepal. The Global Partnership for Education (GPE) includes a variable tranche for its grants to education and has done so for around 20 countries to date, again including the three countries in this study.¹⁸ Bilateral donors such as the (then) Department for International Development (DFID) of the United Kingdom and the Swedish International Development Cooperation Agency (Sida) have also shown increasing interest and use of RBF and have been integral to the programming in Tanzania; Germany (KfW) has used RBF in Mozambique; the United States Agency for International Development (USAID) and Finland in Nepal; and in Tanzania the Korean International Cooperation Agency (KOICA) also experimented with RBF towards the end of the study period, on a small scale but specifically intending to learn from this for future programming.

23. This chapter provides a summary overview of the three countries' key characteristics, recent events in their political economy, and summary of education sectors including performance. Finally, a summary of the RBF "mapping" undertaken is presented and the main RBF instruments that have been used in each case set out. It is important to note that our assessment does not include the period in which the COVID-19 pandemic unfolded across the world, including in these three countries. COVID-19 may impact the three countries differently, with Tanzania pursuing a notably divergent response. All three countries closed schools at some stage; as of October 2020, Nepal had partly opened schools; schools in Tanzania had been fully reopened for several weeks; and Mozambican schools were remaining shut.¹⁹

2.1 Overview of the three countries

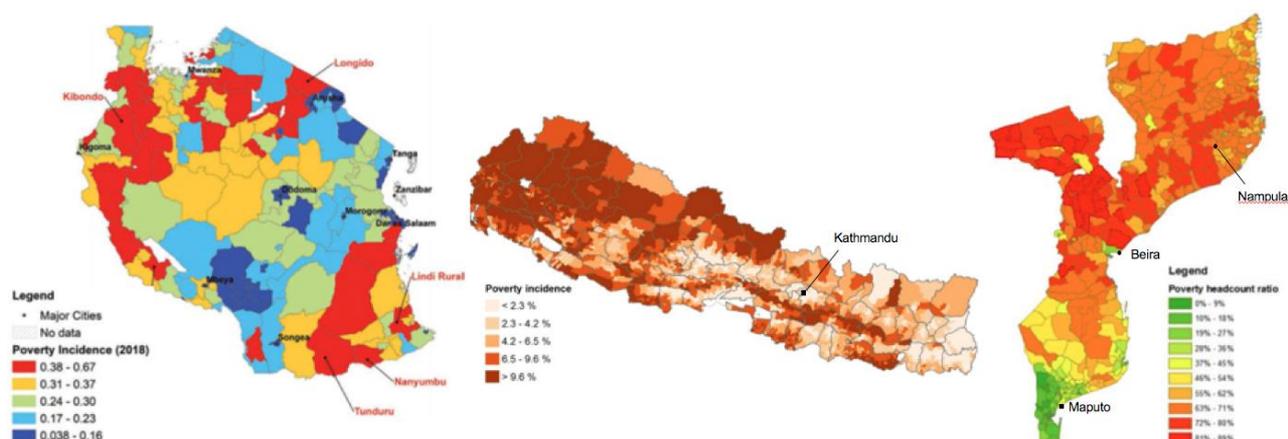
24. The three countries have some similarities, though despite two being neighbors, they also have a large number of important differences: in their recent history, their political economy and the characteristics and performance of the education systems analyzed for this assessment. The UN's Human Development Index (HDI), combines indicators on life expectancy, years of education, and gross national income (GNI), into a simple measure of development. On the HDI, Tanzania, Nepal, and Mozambique ranked 159, 147, and 180 respectively in the world in 2019. The ranks do not reflect the trends of absolute progress achieved on these indicators in recent decades. However, widespread and multidimensional poverty persists in each country. As shown in Figure 1, poverty rates are most concentrated in certain geographic areas: in the great lakes region in the north west of Tanzania, as well as in the southern borders; in the remote, mountainous northwest, northern, as well as the Terai (southern plains) regions of Nepal; and the central and northern regions of Mozambique are significantly poorer than the south. In each country, rural poverty predominates, although an increasing share of the population live in urban areas.

¹⁷ Tanzania refers to mainland Tanzania in this paper at all times, unless otherwise stated.

¹⁸ As of September 2020. World Bank PforR [chart of operations](#). ADB RBL [program portfolio](#). GPE key [data](#).

¹⁹ UNESCO (2020) [COVID-19 Impact on Education data](#), accessed October 20, 2020.

Figure 1 Poverty ratios in Tanzania, Nepal and Mozambique, by geography



Source: World Bank (2019a) for Tanzania. World Bank (2013) for Nepal. IMF (2016) for Mozambique.

25. Across a number of other dimensions, there are some important differences between the three countries. For example, the Nepalese population is better off in terms of access to sanitation and electricity, as well as having lower prevalence of HIV. Mozambique is more unequal, significantly poorer on gross domestic product (GDP) per capita, more aid dependent (in terms of ODA as a share of GNI), and has significantly lower life expectancy. Tanzania has around double the population of the other two countries, is slightly richer than Nepal on GDP per capita in current USD, although poorer when using a purchasing power parity measure, and also faces more severe human development challenges than Nepal. According to the 2020 Fragile States Index, Nepal is ranked as the 49th most fragile country in the world though has had a large improvement in its fragility rating since the civil war ceasing in 2006. Mozambique is ranked 27th, in the third-highest “Alert” category of fragility, having worsened significantly over the past decade. Tanzania is ranked as less fragile, in the fifth-highest category and ranked 61st, having secured prolonged post-independence stability, though challenges remain.

Table 4 Mozambique, Nepal and Tanzania: summary data on characteristics

Indicator	Mozambique	Nepal	Tanzania
Population	31.3 million	29.1 million	59.7 million
% of population in rural areas	63%	79%	65%
% of population 0-14 years old	44%	29%	44%
Life expectancy	61 years	71 years	66 years
Gross Domestic Product / capita (current USD) (2019)	USD 492	USD 1,071	USD 1,122
Gross Domestic Product / capita, purchasing power parity (constant USD) (2019)	USD 1,334	USD 3,558	USD 2,771
Gini coefficient (measure of inequality, higher value = less equal) (2010 for Nepal, 2014 for Mozambique, 2017 for Tanzania)	54.0	32.8	40.5
ODA (overseas development assistance) as % of gross national income (GNI) (2019)	14%	5%	5%
Rank on UN Human Development Index 2019 (out of 189 countries)	180 th	147 th	159 th
Fund for Peace, Fragile States Index Rank (out of 178 countries; 1 st being the most fragile country)	27 th	49 th	61 st
Access to electricity (% of population) (2018)	31%	94%	36%
Probability of a child’s survival to age 5 (% probability)	93%	97%	95%
HIV prevalence (% of population aged 15-49) (2019)	12.9%	0.1%	4.8%
Use of basic sanitation (% of population) (2017)	29%	62%	30%

Source: Data for 2020 unless stated, World Bank World Development Indicators (data accessed in October 2020).

2.2 Political economy and timeline of key events

26. The broader political economy sets the context for education reform, including the accountability structures in which this takes place. This section sets out some short but important points of context for each of the three countries including some recent history that will enable the reader to follow the RBF findings set out in Chapter 3 more easily.

27. **Mozambique** has seen one-party rule since the civil war ended in 1992, with parliamentary elections won by incumbents, the Mozambican Liberation Front (FRELIMO), in all six presidential general elections held since then. The President is Head of State and all branches of the Executive, and appoints the Prime Minister, provincial governors, and other key institutional figures.²⁰ The strong executive power base links to relatively weak provinces. FRELIMO historically has been a southern force,²¹ while the opposition RENAMO draws from groups in the center of the country. The political disparity further exacerbates the economic divide between the south from the center and north of the country.²² Observer reports for the most recent election in October 2019 noted severe irregularities and brutality towards opposition party agents.

28. Around half of Mozambicans live in the 52 urban municipalities, which hold local elections and have a semblance of local democracy. The other half live in the 111 rural districts, which have no local democracy and are ruled directly by FRELIMO-appointed officials. The subnational bureaucracy provides tens of thousands of jobs and this is understood to be one of the main means of retaining legitimacy for the ruling party.²³ This means that provincial and district appointments can be made on the basis of political affiliation and party loyalty, and not for merit-based reasons. Connected to this system of clientelism, Mozambique faces entrenched corruption, described as “endemic” across the government system,²⁴ including bribes and side-payments for access to many public services.²⁵ In December 2018, the government announced that an investigation over two years had uncovered that 30,000 out of 348,000 civil servant workers were “ghost” workers – with a cost to the Treasury of USD 250 million between 2015 and 2017.²⁶ Mozambique has a young and growing population of 30 million people, which is increasing at 3 percent per annum, with high levels of unemployment.

29. **Tanzania** is known for post-independence political stability and its homegrown “African socialism” conceived by *Mwalimu* (“teacher”) Julius Nyerere, Tanzania’s first President, replaced since the 1990s by a more classical economic paradigm. The formal introduction of multi-partyism in the mid-1990s has not resulted in political change thus far, with the Chama Cha Mapinduzu party of Nyerere winning all elections since then, albeit with reduced margins in 2010 and again in 2015. The current (2015-) presidency’s anti-corruption stance and “results drive” was accompanied by perceptions of reduced transparency and space for dissent according to Freedom House and other international observers.

30. The past ten years saw high government-reported growth and some, although disappointing, poverty reduction, due to both the concentration of employment in slow-growing sectors and the sluggish rate of transformation of the economy. In 2020, approximately 37 percent of the population lived in urban areas. The

²⁰ Norwegian Institute of International Affairs (2017): “...the National Assembly appoints the Citizen Ombudsman (Provedor de Justiça), but it remains a weak institution with few real powers. Both the National Assembly and the Tribunal Administrativo (auditor general) can inspect the national accounts, and the Central office for the Combat against Corruption (GCCC) can theoretically hold government officials accountable, although immunity rules for politicians are impediments to accountability.”

²¹ Though because of natural resource finds in the northern province of Cabo Delgado and the President coming from this area, the political geography has become more complicated in recent years.

²² This has been reflected in some limited success for RENAMO at local elections in 2018, in which they won 8 out of 52 municipalities in the country, which were within central and northern provinces.

²³ Norwegian Institute of International Affairs (2017), Macamo (2006), and Pérez Niño and Le Billon (2016)

²⁴ Nuvunga and Orre (2019)

²⁵ GAN Integrity (2020)

²⁶ BBC, December 11, 2018 “Mozambique busts ‘30,000 ghost workers’” - <https://www.bbc.co.uk/news/world-africa-46520946>

under 15 age group represented 44.1 percent of the population in 2017, with another 35.5 percent in the 15–35 age group. Unemployment is concentrated among the younger generations.

31. In spite of a decentralization policy enacted in the 1990s, local government authorities (LGAs) in charge of basic services provision are constrained by centralizing trends, and an intergovernmental fiscal transfer system that ignores large variations in constraints and potentials. Dependency on ODA, which used to be among the highest of any developing country, has decreased of late. This and persistently low revenue levels have resulted in a tight macroeconomic and fiscal context during the past five years. After a decade of high budget support (BS) flows, corruption concerns and disillusion with its meagre demonstrable impact led to the quasi-disappearance of budget support (BS) after 2013, to the emergence of results-based financing as an alternative was propitious.

32. In **Nepal**, in November 2006, the Government of Nepal and Maoist rebels signed a Comprehensive Peace Accord, declaring a formal end to a ten-year civil war. The Maoists joined a transitional government, and an Interim Constitution came into place in January 2007 to guarantee a “democratic, federal system”. A complex and drawn-out, eight-year process followed, leading to the 2015 Constitution of Nepal in September of that year. The new Constitution marked the end of Nepal’s monarchy and the beginning of a new era as a federal republic.²⁷ Elections since have led to coalitions between three (now two) dominant parties: the Nepali Congress Party, the Communist Party of Nepal (Maoist Center), and the Communist Party of Nepal (Unified Marxist-Leninist). The two communist parties merged in May 2018 to form the Nepal Communist Party, the ruling party in the country today. In November 2017, federalism was instituted with elections for new province and local government (LG) levels. There are said to be linkages between political and business interests, including in ownership or part-ownership interests in private schooling. The country experiences significant corruption, including in the education sector.

33. Nepal’s economy is tied to that of its giant neighbor, India. The countries share a porous border through which nearly two-thirds of Nepal’s imports come. In 2018, Nepal had the fifth highest ratio of remittances to GDP in the world, at over USD 8 billion or 28 percent of GDP, and (pre-COVID) an estimated 5.5 million Nepalese people, mainly men, are estimated to work overseas at least part of the year, mainly in India and in the Middle East. Emigrants’ remitted earnings allow their families to invest more in education and in healthcare than would be possible in the absence of this income, and this has been linked to the size and growth of private education in the country.²⁸

34. Nepal is diverse ethnically and linguistically with 123 languages spoken as a mother tongue according to the 2011 census. Intertwined with this is a complex caste system, traditionally associated with Hinduism but with variants in Nepal. Lower caste groups such as those classified as *Dalit*, are associated with low social status and with significant social and economic deprivations. Elite groups are dominated by those of “upper / higher” caste.

Timeline of key events

35. The period of review has been eventful for each of the three countries. Figure 2 provides summary timelines for the countries including in major developments in political economy as well as the education sector, and major external shocks. Some important highlights of the recent history of each include:

- In **Nepal**, 2015 saw a large earthquake, killing nearly 9,000 people, with hundreds of thousands made homeless, and over 17,000 classrooms damaged and in need of reconstruction. The Constitution of 2015 instigated a series of decentralizing reforms towards a federal structure of government. An array of legislation brought federalism into practice, in what stakeholders refer to as a “big bang”. The Civil Servant Adjustment Act (2017) provided the legislation for the large reallocation of civil servants

²⁷ Throughout this report the changes brought about from the new Constitution are referred to as changes due to **federalism**, the term most country stakeholders use to describe the constitutional changes.

²⁸ Acharya & León-González (2018) find migrant households are significantly less likely to keep their children out of school and those receiving remittances had an average of 1.3 children in private schooling; compared to 0.8 for those not receiving international remittances.

between the central and new local tiers of the system; a process still ongoing as of 2020. Other legislation included the Local Government Operations Act (2017) setting out the new delegated responsibilities. The changes included the move from deconcentrated districts (District Education Offices - DEOs for education) with limited autonomy; to, in principle, large scale decentralization with devolution of responsibilities to a new tier of 753 local governments (LGs), known as the “local level”. See Box 2 below for more information.

- In **Mozambique**, major off-shore reserves of natural gas were discovered in 2010, becoming the main driver of foreign direct investment and economic growth since. The benefits (including jobs) are yet to accrue from these long-term investments large due to an insurgency in Cabo Delgado province, which started in 2017 and has intensified since 2020. Between 2014 and 2016, large secret loans were uncovered, with loans worth around USD 2 billion found to have been taken by three government-owned companies without the knowledge of either the Mozambican Parliament or the donor community.²⁹ This led to the suspension of General Budget Support (GBS) which had been a significant feature of ODA from as many as 17 joint partners. This scandal, and the subsequent suspension of GBS, had significant economic implications, including a major depreciation of the metical, increased borrowing and increasing inflation.
- In **Tanzania**, following an outcry in 2012 at low exam results and increased high-level interest in early-years learning as a pathway to improved education outcomes, a relatively country-led process heralded a new focus on learning and quality. This dovetailed with a broader process of “planning for results” (“Big Results Now”) conceived by the then President to seek to reverse the growing disaffection from the ruling party ensuing from its poor performance in basic service delivery. The arrival of President Magafuli in 2015 resulted in a policy undertaking for fee-free basic education, which emphasized access and included a higher focus on secondary-level education.

Box 2 The “big bang” of federalism in Nepal

The Constitution of 2015 instigated a series of decentralizing reforms towards a federal structure of government. The new structure created a tier of 7 provinces, and a tier of 753 local governments (LGs) known as the *local level*; both came into effect in 2017 with elections in November of that year. LGs took on many of the previous responsibilities of districts in service delivery, in particular District Education Offices (DEOs) which were the primary conduit for education sector funds from central government and down to schools. The previous support structure for schools, of over 1,000 Resource Centers (RCs) staffed by Resource Persons (RPs), were disbanded with the changes. The Constitution granted powers to the new local level for basic and secondary education, while higher education and teacher training was assigned to province level. Responsibilities devolved to LGs included the management of permanent teachers; curriculum and textbooks; examinations and student evaluation; and reporting and monitoring. Changes also created Education Development Coordination Units (EDCUs) at district level, with less staffing than the previous DEOs, and a role that is still being defined with respect to supporting LGs.

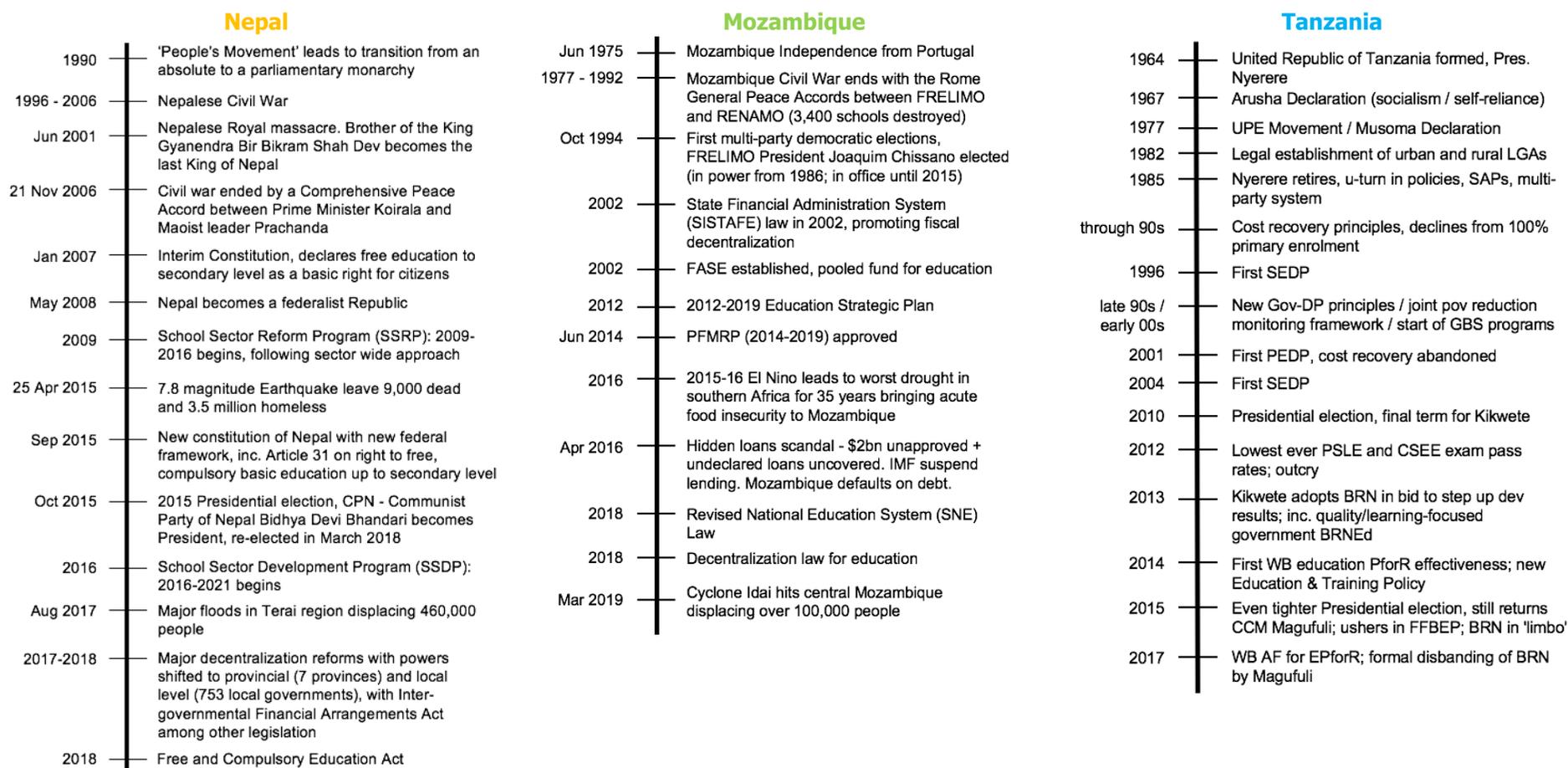
An array of legislation brought federalism into practice, in what some sector stakeholders have referred to as a “big bang”. This included the Local Government Operations Act (2017) which set out the responsibilities of LGs and put in place a radical interpretation of the decentralization set out in the Constitution. The Intergovernmental Financial Arrangements Act (2017) specified the new framework for public financial management (PFM), managing resource transfers, including the introduction of a Conditional Grant, which now channels most of the education budget from the center to the local level. The Appropriation Act (2018) brought some power back to the center, seeking to ensure that unspent money from LGs would go back to the central government, in a complex and evolving fiduciary set-up.¹

The Constitution envisaged the enactment of federal education acts to outline the new functions; however, the expected Federal Education Act has not yet been passed (as of October 2020). In the meantime, based on “model acts” passed on by Ministry of Federal Affairs and Local Development, it has been reported that many LGs have formulated their own acts, regulations and working procedures, including for the delivery of basic and secondary education. Stakeholders reported contested areas such as who should have responsibilities for the recruitment and promotion of teachers; the use of English language in tuition; and the roles with respect to curriculum.²

¹ Turner et al. (2019) + SOFRECO and FBC (2019). ² OPM (2018a).

²⁹ IMF (2019).

Figure 2 Timelines of key events for Nepal, Mozambique and Tanzania



2.3 Summary of country education sectors

36. All three countries have made progress in expanding access to education in recent years, but have equally faced the challenge of addressing quality, improving teaching and learning while the system expands. As shown in Table 5, there are some key differences in education indicators; notably, Nepal is performing better on primary completion rates and rates of enrolment at secondary. Tanzania has a greater proportion of primary school aged children out of school but this may reflect that the official primary school age goes up to 14 years, compared to just 10 years in Nepal and 13 years in Mozambique. Nepal has a much higher tertiary enrolment rate, and far higher secondary net enrolment ratio (NER) than the other two countries. A 2020 Human Capital Index (HCI) produced by the World Bank found a higher rank for Nepal than the other two countries, linked to higher school retention; Nepal also enjoys higher average learning-adjusted years of schooling, although having slightly lower quality according to the harmonized learning assessment measure, it has fewer out-of-school primary aged and secondary aged children. All three countries rank relatively low internationally despite progress in access indicators in the past two decades.

Table 5 Mozambique, Nepal and Tanzania: education high-level data

Indicator	Mozambique	Nepal	Tanzania
Literacy rate	61%	68%	78%
School net enrolment, primary (%)	94%	96%	81%
School net enrolment, secondary (%)	19%	62%	27%
Tertiary enrolment, gross (%)	7%	13%	3%
Primary completion rate (% of relevant age group)	55%	120%	68%
Out-of-school population, primary-aged %	2%	4%	13%
Rank on World Bank Human Capital Index 2020 (out of 182 countries)	175	116	160
Learning-Adjusted Years of School	4.4	6.9	4.8
Harmonized test scores, advanced attainment=625 (2007 for Mozambique, 2013 for Tanzania)	368	369	388
Out-of-school population, primary-aged %	2%	4%	13%

Source: Data for 2019 unless stated, Human Capital Index data used for learning adjusted years of school, test scores and children out of school. Literacy rate for 2015 for Tanzania, 2017 for Mozambique, 2018 for Nepal; tertiary enrolment 2018 for Mozambique. Secondary enrolment 2015 for Mozambique, 2018 for Tanzania. Enrolment, literacy and completion rates from World Bank, World Development Indicators, sourced from UNESCO Institute for Statistics.

37. Further to Table 5, there are also important challenges in key education **access/retention** indicators, with some differences between the three contexts.

- In **Mozambique**, the system nearly achieved full enrolment in the first cycle of primary education, but there is a significant inefficiency, with high levels of drop out and repetition, and low completion of the full primary cycle. Many children are over age for their grade and it is estimated that around 15 percent of children of primary school age are out of school.³⁰ Secondary access has improved from a low base, but remains low, with a gross enrolment ratio (GER) of 41 percent in lower secondary level.³¹
- In **Tanzania**, earlier declines in pupil intake at pre-primary, primary and lower secondary levels were reversed by the fee-free basic education policy (FFBEP), implementing the 2015 political campaign promise of President Magufuli. The FFBEP, combined with improvements in retention and increasing school age population, led to primary enrolment increasing by 20 percent between 2014/15 and

³⁰ World Bank (2020).

³¹ MINEDH (2019).

2017/18, significantly exceeding the population growth rate, and total primary/secondary school enrolment projected to increase by 48 percent by 2029/30.³²

- In **Nepal**, net enrolment in primary, basic and secondary education all increased in the period from 2009. The net enrolment ratio (NER) for basic education (Grades 1–8) reached nearly 93 percent, with survival rates of nearly 90 percent for primary and 78 percent for basic education. There were increases in the share of children with experience in early childhood education (ECED) up to close to 70 percent by 2019 for those entering Grade 1, from just 11 percent in 2009.

38. Each of the countries faces significant challenges in terms of the **quality of education** and more broadly of service delivery. In Mozambique, a World Bank Service Delivery Indicator (SDI) report in 2014 had a big impact on how the government and its partners viewed sector priorities, revealing significant problems in the quality of education at school and classroom levels. The study showed that 45 percent of teachers were not in school during unannounced visits, and a further 11 percent were in school but not in the classroom when they were supposed to be teaching. The level of absence for school directors was similarly high. Teacher subject knowledge was also found to be very low in Portuguese language, mathematics and pedagogy.³³ Teacher absenteeism contributed to children receiving less than 2 hours teaching per day, over 50 percent less than the average of countries in the survey. A second SDI survey³⁴, which showed improved indicators, was conducted in 2018, although more severe problems remain in the north of the country. These trends were confirmed by the National Learning Assessments for Grade 3, introduced in 2013 and repeated in 2016; in 2016 only 5 percent of Grade 3 children developed basic reading and writing skills at the expected level, a decline in performance from 2013.³⁵ The 2016 assessment introduced mathematics, and only 8 percent of Grade 3 children could demonstrate the problem-solving expected of them.

39. In Tanzania, the “learning crisis” realization in 2012/13 was prompted primarily by historically low pass rates for end-of-cycle exams in 2012, after several years of decline; combined with SDI and other evidence, there were found to be high levels of teacher absenteeism with nearly one in four teachers absent from school on any given day, and over half of the teachers present at the school, not teaching in class. Repetition and resulting late-age dropout are significant issues as well. Early learning levels were not assessed before 2013. The first assessment showed low learning levels, though outperforming Mozambique (in 2013 only 8 percent of Standard 2 pupils were able to read with comprehension). Over the study period (and according to the assessment approach and indicators selected for the RBF program) the level went up first (2015), then deteriorated below the 2013 level in 2017 – an effect of the FFBEP enrolment surge which, in addition, was not accompanied by a commensurate increase in government education funding. In contrast, and contributing to the teaching/learning quality decline, the PTR at primary level increased markedly during the study period – 2014 to 2019 – owing to a wage bill control program and certificate-related staff rationalization, coupled with persistent budget credibility issues and a budgetary emphasis on infrastructure.

40. In Nepal, both the 2007 Interim Constitution and the final 2015 Constitution declared free education to secondary level as a basic right for citizens. This was instituted in a Free and Compulsory Education Act in September 2018. According to Poyck et al. (2016) the 2007 declaration enabled consistently high levels of investment in the education sector and has given the needed support for several reforms. Until 2016, elementary education in Nepal lasted for five years – from grade 1 to grade 5 (ages five to ten). However, an amendment to the Education Act passed in 2016 extended the elementary education cycle and established a new system of compulsory basic education that is in theory accessible to every child in Nepal free of charge at public schools. In addition, children have the option to enroll in public early childhood development centers or private kindergartens before entering elementary school at the age of five. Few consistent learning assessments have

³² World Bank (2019c).

³³ World Bank (2015a).

³⁴ Bassi et al. (2019).

³⁵ MINEDH (2020).

been undertaken; however, measurement has become institutionalized in recent years. An EGRA assessment was carried out in 2014 by USAID (2014) and found a reading fluency measure of 14 words per minute (wpm) for Grade 2 students and 27 wpm for Grade 3 students; the latter number has become a baseline for the SSDP period. Learning results have also included quite regular national assessments on student achievement (NASA). Since 2011, there have been two rounds of NASA for grades 3, 5 and 8 covering (though not for all assessments) mathematics, Nepali, English and science. Results have been poor and the completion report for SSRP stated that the “*publication of the NASA results considerably raised parents’ and the education community’s awareness of the role of poor teaching of reading in the early grades on the consequences of educational underachievement at all levels*”.³⁶

41. There is significant **educational inequity** in each of the countries. In Nepal the SSRP put focus on disparities related to caste, ethnicity, religion and geographic factors, and there was progress towards universal primary education as well as in ECED enrolment. There remain large regional disparities with inequalities across different groups, for instance as measured through years of schooling or pupil teacher ratios (PTRs), and the most marginalized area is the Terai region. Private education plays a large role in Nepal, via Institutional schools (as compared to state Community schools), with around one-fifth of pupils attending one, and much higher rates in urban areas, which are generally found to have much better learning outcomes. In Mozambique, there is persistent inequity between the north and south/central regions, and between urban and rural areas, and girls face particular challenges. This shows up in most indicators, including learning outcomes, with the 2016 national assessment revealing only 2 percent of children in Grade 3 in Cabo Delgado and Manica provinces reading at the expected level, compared to 17 percent in Maputo City.³⁷ In northern provinces 40 percent of teachers absent from schools, compared to 19 percent in the south.³⁸ In Tanzania, long-term trends suggest that geographical and gender equity in education is improving.³⁹ However, by the end of 2019, geographical disparities in access, retention and learning levels were still pronounced, as well-documented in government annual reports,⁴⁰ and while there was progress with the girls’ primary-to-secondary transition rate slowly catching up with the rate for boys,⁴¹ the number of girls sitting for exams at the end-of-lower-secondary and Certificate of Secondary Education Examination pass rates remained below those of boys, especially in math and sciences. Inequality in education outcomes was also associated with poverty and deprivation in other development outcomes.⁴²

Education sector priorities, spending and functions

42. In Nepal, the sector has been managed through a succession of sector strategic plans, including the School Sector Reform Plan (SSRP) (2009–15/16) and the current School Sector Development Plan (SSDP) (2016–2022/23). The SSRP put focus on educational disparities related to caste, ethnicity, religion and geographic factors, and the period saw good progress towards universal primary education as well as in ECED enrolment. The SSDP placed quality foremost and sought to include a range of priorities to address the challenges in low levels of learning set out above. In Tanzania, basic education was a priority in the Big Results Now! (BRN) initiative, undertaken in 2012/13 by then-President Kikwete as a response to the widespread dissatisfaction with the lack of development progress for most people. The BRN process triggered a vigorous education planning process, leading to a “quick win”-focused BRNEd plan, which prioritized basic education quality and learning. In 2017/18, the BRNEd plan was replaced by a more traditional Education Sector Development Plan (ESDP), with quality as one of a long list of priorities. In Mozambique, the MoE has set out its reform priorities in a series of three Education Sector Strategic Plans since 1999, the most recent of which started in 2012 and ended in 2019.

³⁶ World Bank (2017c).

³⁷ MINEDH (2020).

³⁸ Bassi et al. (2019).

³⁹ MoEST Tanzania (2019).

⁴⁰ Ibid.

⁴¹ The primary to secondary transition rate was 67 percent for girls vs. 70 percent for boys in 2016; in 2019 it was 72.5 percent for girls vs. 73.5 percent for boys.

⁴² World Bank (2019a).

A new (fourth) plan has recently been approved for the period 2020–29. Over this period, these Education Sector Strategic Plans, and three-year Operational Plans, have prioritized primary education and have focused on the overarching goals of increasing access while reducing gender and regional disparities, improving quality, and strengthening the administrative system at all levels.

43. To varying degrees, each of the three countries have faced constraints in their national/domestic financing of education, funding of teacher salaries and key reforms. The flow of funds to subnational levels is also an important part of the context in which RBF has taken place.

- In **Mozambique**, spending in the sector is constrained, especially when over 90 percent of the government recurrent budget for the sector is allocated to salaries. As a proportion of GDP, education expenditure has remained close to 6 percent since 2012. Education also forms fairly consistently around 18 percent of the government budget. While these benchmarks are generally higher than average in terms of regional comparisons, as the country remains very poor, expenditure in absolute terms is still very low. In addition, with huge increases in the school population, expenditure per pupil in primary education in real terms was the same in 2018 as it had been in 2012.⁴³ Around 60 percent of the national budget for education is now spent at district level, including salaries and school grants.⁴⁴ In part due to the suspension of GBS alongside the hidden loans scandal, budgets have been strained, this has been linked to a slowdown in the expansion of teacher numbers, which as of 2020 has led to increasing PTRs (see Section 3B).
- Education expenditure by the government in **Nepal** has been stagnant in real terms in recent years, remaining at just less than USD 1 billion, and falling slightly as a share of GDP, from four percent in 2010 to three percent in 2018 and 2019. Education has also fallen as a share of total government spending, having averaged 22 percent in the 2004–09 period, falling to a 17 percent average between 2010 and 2017, and a low of 14 percent in 2018.⁴⁵ The fall is likely to be linked in part to the major allocation of resources to post-earthquake reconstruction. Given the size of the private education sector in Nepal, it is also likely that a much larger share of GDP is being spent on education than the above figure.
- In **Tanzania**, since 2016/17 the total education budget has declined in real terms. Resources allocated to basic education decreased due to a combination of whole-of-government and within-sector decisions, which favored secondary and TVET education to the detriment of pre-primary and primary education. By 2019/20, per student spending on pre-primary and primary pupils was 25 percent lower in real terms than in 2013/14. At the same time, a stringent cap on civil service recruitment which aimed to maintain the macro-fiscal balance was applied to teachers as well, so that in most years, recruitment of teachers barely compensated attrition. LGAs' discretion is highly constrained: basic education is funded primarily through tightly earmarked transfers from the MoF, and in 2017/18 salaries consumed 94 percent of these.⁴⁶ Schools get capitation grants (CG), but their real value has significantly eroded since they were set up in 2002.

44. Each country now has two layers of subnational administrations, each with some involvement in education. Nepal and Mozambique have more distinct provincial responsibilities than Tanzania, where regional administrations play a support role. Nepal and Mozambique have recently raised efforts regarding decentralization, with full federalism in Nepal, and a more deconcentrated approach in Mozambique (that is, districts have important administrative functions but limited decision-making power, with limited local democracy in place particularly in rural areas). Tanzania has the longest-established decentralization policy, with Local Government Authorities (LGAs) in both rural and urban areas since 1992. However, decentralized education

⁴³ Analysis of data from MINEDH (2019).

⁴⁴ UNICEF (2017) and UNICEF (2018)

⁴⁵ World Bank WDI, indicator government expenditure on education, total (% of government expenditure).

⁴⁶ World Bank (2019c)

structures have been treated as quasi-deconcentrated units. Inspection is led by the central level in Tanzania. All countries have some local government latitude to hire and locally redeploy staff, but fiscal autonomy is almost non-existent in Tanzania, so that all but the richest urban governments have little latitude to hire additional teachers and these costs are passed on to communities. All countries manage education at the central government through a variety of agencies. Notably all countries maintain a division between the ministry setting the number of teachers and the ministry setting the budget. As explored in this report, this was often an important dynamic in RBF. Unsurprisingly, language of instruction policy and curriculum design is centralized in most countries. Teacher training, which is emerging as a key channel for learning and education quality, is controlled centrally in Tanzania.

45. The key **functions of the education system** sit within the broader government structures, which vary in important ways between the three countries, as represented in Figures 8, 9 and 10.

- In **Nepal**, the policy level is led by the Ministry of Education, Science and Technology (MoEST),⁴⁷ with the Center for Education and Human Resource Development (CEHRD) as the main implementing agency below it. Other important agencies include the Education Review Office created to “institutionalize NASA” learning assessments, the National Examination Board (NEB), the Curriculum Development Center, and the Teacher Services Commission. At lower levels, roles and responsibilities for education before federalism (pre-2017) were predominantly for District Education Offices (DEOs) and a network of Resource Centers (RCs) staffed by Resource Persons (RPs). With the changes from 2017, these roles were removed. The new local government (LG) tier took on most responsibilities with respect to basic and secondary education, with an education section staffed by at least one Education Officer. In addition, the district level tier was retained with a new tier known as Education Development Coordination Units (EDCUs), with much more limited staff and responsibilities than the previous DEOs. At the provincial level, sectoral functions include higher education and teacher training.
- In **Mozambique** the Ministry of Education and Human Development (MINEDH) leads most sector functions, setting policy and strategy, as well as targets for the sector, and communicates these to subnational levels. Roles are split between its main departments. There are two lines of accountability from schools to the provincial administration, as well as to the central MoE. School directors are appointed through provincial administrations and report to the district administrator, who reports to the provincial governor, a political appointee. Provincial governments follow the sectoral organization of the center, so provincial directorates of education and human development, implement and coordinate MoE operational plans. District education services (SDEJTs) play a key role through the supervision of schools, teacher recruitment and pay, and as a conduit for school grants.
- In **Tanzania**, LGAs manage education provision under supervision by the 26 regions and the President’s Office-Regional Administration and Local Government (PO-RALG), and technical oversight by the Ministry of Education, Sciences and Technology (MoEST). Since 2018, a new participatory School Quality Assurance (SQA) system has been rolled out to monitor schools, replacing the inspection system. The autonomous Tanzania Institute of Education reporting to the MoE, leads on curriculum development with a monopoly on textbook development and publishing since 2014. The National Examinations Council of Tanzania (NECTA, similar status to TIE) designs and conducts all exams and learning assessments. The Ministry of Finance and Planning (MoF) and President’s Office—Public Service Management (PO-PSM) play key roles in resource allocation and financial accountability, managing civil service pay scales, and (jointly) fixing limits for recruitment and promotion, including for teachers.

⁴⁷ The abbreviation **MoE** is used for all three ministries of education in this report, although MoEST is commonly used in Nepal and Tanzania; while MINEDH is used in Mozambique.

46. In all three countries there is a well-established landscape of **donor coordination**, with forums for policy dialogue. In Mozambique, this is organized around regular meetings, involving sector and thematic working groups on specific issues, and the high-level Joint Coordinating Group chaired by the Permanent Secretary with donor and civil society representatives. The Local Education Group has a broader representation and a more operational/coordination and information-sharing function. The Annual Review Meeting reviews performance against a sector performance assessment framework; it was seen as too process focused, but in recent years has been becoming more results oriented.

47. In Nepal, the development partner, non-governmental organizations (NGOs) and Civil Society Organizations (CSOs) are members of the Local Education Development Partner Group. This brings together these stakeholders with government through technical working groups, which meet on a regular basis and are convened on a range of technical areas linked to the sector strategy. The Budget Review Meeting (BRM), usually taking place around March or April, reviews the budget for the following year, and development partners outline their contribution for the next year. At the Joint Review Meeting (JRM), usually taking place in November, government reports on sector performance including in the Status Report and in DLI achievement reports. Stakeholders in Nepal are clear they have an exemplary Sector Wide Approach (SWAp), which in part underpinned the very ambitious design of RBF in the SSDP period (see below).

48. In Tanzania, sector dialogue takes place under the broader framework of the Joint Assistance Strategy which reiterates the “self-reliance” principle. However, the government-development partner dialogue through sector technical working groups set up under the sector strategy had become unwieldy and characterized by long processes with little focus, and ignoring the sector key issues (‘all about process, no substance’). Moreover, annual sector performance reviews, whilst they have improved in terms of analytical content over the past five years, have continued to lack operational focus, also not including a review of donor support programs and projects.

2.4 RBF in education

49. The first stage of this assessment leading to preliminary reports for each country,⁴⁸ started with a broad mapping of the use of RBF in the education sector in each, including as used by donors in their programming of support, and by governments in their own funding of education system reforms. A key finding from the first phase mapping was that RBF has been chiefly used by donors, but has not been taken up at scale by any of the three governments in the way they finance education. There have been some uses of RBF at subnational levels of the education system, including through performance-based components to school grants. However, these have been linked to donor supported programs, and in some cases the use of subnational RBF is still at the level of piloting. As such, the research has focused primarily on the use of RBF in the provision of external donor financing to the sector and the way this has aimed to strengthen and support system-level reforms.

50. The use of RBF was therefore introduced within a broader context of donor funding to the sector. In Nepal and Mozambique, there were already joint / pooled funding arrangements in place and RBF was introduced within these. In Tanzania, RBF represented more of a break with funding instruments and modalities prevailing in the last decade. Development partners have played a role in financing each education sector over a number of years including through mixtures of sector budget support and general budget support in previous years. The shift to RBF since 2013 has been more sizeable by value in Tanzania and in Nepal (since 2016) than its use in Mozambique. The majority of sector support in the latter case is still by non-RBF financing modalities.

51. In Nepal, the majority of sector finance up to 2016 from development partners was not linked directly to results. However, during the SSRP period there was one significant use of RBF programming. The World Bank linked USD 25 million out of USD 100 million of additional financing to the achievement of incentive-linked indicators (ILIs) in 2013/14 and 2014/15. For the following SSDP period, RBF use was scaled up significantly.

⁴⁸ Tanzania – Dom et al. (2020); Mozambique – Patch et al. (2020); Nepal – Holden et al. (2020).

The five-year costed program from 2016/17 to 2020/21 has represented a major experiment in the financing modality, in which six out of nine joint financing partners (JFPs) to the education sector linked all or part of their disbursements to RBF indicators.⁴⁹ These were split between ten disbursement linked indicator (DLI) areas, with the World Bank and ADB the largest two JFPs, linking their concessional loans to seven DLI areas each; and the European Union, the largest donor via grant funding, focusing on three DLI areas. RBF financing in total for the SSDP period linked to DLIs makes up a maximum disbursement of approximately USD 400 million out of USD 490 million of external sector finance over the five-year period, with donors providing up to 8 percent of the projected education budget in total for the period.

52. In Mozambique, eight donors currently support the sector through a ring-fenced education sector support fund, FASE (*Fundo de Apoio ao Sector de Educação*), together providing approximately USD 100 million per year. Over the years this has represented a varying share of total sector funding, from less than 10 percent up to 30 percent. While as financing expenditure classified as “investment”, which is the case for construction of schools and classrooms, FASE also covers what can be considered essential (non-salary) recurrent expenditure for the sector, including: textbooks; school grants; funds for teacher training institutions (IFPs); and district supervision of schools. RBF was introduced as part of the mix of donor support modalities to the sector, representing a relatively minor proportion of this donor funding and focusing on selected reform areas. Since 2014, three significant attempts have been made at using RBF in the education sector: The World Bank’s *Public Financial Management for Results (PFM4R)* program (worth approximately USD 24 million for education); the introduction of *variable tranche payments* under GPE funding in 2015–17 (worth USD 17.4 million, alongside USD 41 million fixed tranche);⁵⁰ and Germany (KfW) introduced two *variable tranche* payments (USD 17.2 million) to its existing funding to FASE in 2014 and 2016.

53. In Tanzania, aid modalities being used in the sector were facing challenges. There were concerns with corruption, and general and sector budget support was seen as insufficiently connected with the realities of the sector, and a large performance-linked tranche was not disbursing. There was a lack of and limited use of sectoral analysis. Projects showed poor disbursement and slow implementation (for example, a World Bank Secondary Education Development Project). During the study period (2013–19), in basic education, RBF replaced Budget Support, and was viewed as a fitting way of supporting the government’s nascent focus on education quality and learning results featured in its own government wide BRN initiative. The initial RBF aid program, also called Big Results Now for Education (BRNEd), covered a four-year period and cut across government levels (national and LGAs) and all education result levels. Towards the end of the second-year government and development partners agreed to extend the program (from then on called Education Program for Results – EP4R) for a fifth year, with new DLIs to address the new fee-free basic education policy. GPE and later KOICA also joined the program. At both stages, RBF was used together with technical assistance (TA) as part of the package. The final program, with USD 515 million of potentially available aid financing, represented approximately a third of all aid financing for the education sector.

54. Table 6, below, summarizes the main features of these RBF programs. It is notable that all of the programs include a mix of RBF with technical assistance and other capacity development. Independent verification is a key element in the design of RBF programs, and the approach to this has varied across the different programs. Nepal used national partners for the independent verification agent (IVA) role, although this was not required by all donors in the SSDP; Tanzania used an international firm (OPM) contracted through DFID; and Mozambique used the government audit institution (the *Tribunal Administrativo*) for the PFM4R, a local subsidiary of an international firm (EY) for the GPE, and an internal verification process for the German (KfW) use of RBF.

⁴⁹ The nine are Finland, the GPE, the Japan International Cooperation Agency (JICA), Norway, UNICEF, USAID, the ADB, the EU, and the World Bank. Australia was originally going to fund the SSDP but pulled out before any disbursements.

⁵⁰ GPE RBF was programmed as additional financing to the World Bank-managed (non-RBF) *Education Sector Support Project* (ESSP).

55. Performance-based payments beyond central government level were limited in implementation, except in Tanzania, where a third of RBF payments were allocated through pre-agreed formulae to local governments against achievement of results for a subset of DLIs, also a small school pilot, and central government's discretionary top-ups which supported a construction program via school, village and LGA financing implementation. The PFM4R in Mozambique included in design "performance-based allocations" (PBAs), which would be subnational payments associated with RBF results, although as set out in Chapter 3, they did not work as intended. The SSDP design did not have subnational payments built in, although one DLI area included a target result on performance-based grants to schools.

56. The following Chapter 3 will provide our detailed findings on the RBF programming above, covering in order Design (A); Results and Impact (B); Financing and payments (C); Coordination, capacity and accountability (D); Evidence-based policy and verification (E); Adaptiveness and flexibility (F); Risks and unintended consequences (G); and Cost-effectiveness (H). Chapter 4 will provide some conclusions on the synthesis against the theories of change for RBF outlined in Table 1 in Chapter 1. Finally, Chapter 5 will provide some recommendations for those involved in the appraisal, design and implementation of RBF programming.

Table 6 RBF programming in the 2013–20 period

	Mozambique	Nepal	Tanzania
RBF programs, funders and governance	Three programs with separate DLI frameworks: World Bank PFM4R; GPE Variable Tranche programmed alongside WB ESSP; Germany (KfW) RBF in construction. PFM4R program structure / coordination team in MoF. GPE contribution managed by World Bank as part of contribution to FASE, KfW also within FASE. FASE controller within MoE finance department.	Four World Bank incentive linked indicators (ILIs) in last two years of SSRP. Ten DLIs for SSDP, used by six funders (WB, ADB, EU, GPE, Finland, USAID), from total nine Joint Financing Partners of SSDP.	BRNEd (first phase): one DLI framework, three funders (WB, DFID, Sida); EP4R (second phase): to five funders (GPE & KOICA joined), each with own mix of DLIs, funding pattern & instrument (WB use of PforR). First phase part of government-wide BRN initiative, with MoE BRN steering committee. Second phase: program coordination units in MoE & PO-RALG; joint MoE/PO-RALG Senior Management Team (SMT).
DLI frameworks / focus areas	1) WB PFM4R 2014–18: focus on education & health. DLIs on school & local accountability (school grant timeliness, school supervision by SDEJTs, school council functioning) 2) Variable tranche (VT) in GPE FASE funding 2015–17, focus on teacher training, school director training, and PTRs for lower primary level. 3) KfW focus on construction (2013–16), via two separate tranches (one in 2014, one in 2016) based on targets for 1,000 classrooms per year.	1)WB SSRP (2013–15): teacher salaries and secondary students' scholarships paid via bank accounts; textbook timeliness; dissemination of learning assessment reports (NASAs) 2)SSDP (2016–21): ten DLI areas with 84 results: quality (curriculum, early grade reading, certification, learning assessments, improved exams); teachers (deployment equity & subject teachers' allocation, training, time spent teaching (TST)); access (OOSC in deprived areas, secondary scholarships & GER); secondary Model Schools, DRR classroom construction, fiduciary & governance, and EMIS.	In both phases, foundational DLIs, mostly prerequisites for subsequent recurrent DLIs. Recurrent DLIs: finance flows, data management, teacher management, school quality assessment system & level, retention rates, learning (reading and math). 1 st phase (2014–16): 6 recurrent DLIs. 2 nd phase (2017–20): 9 recurrent DLIs; several with sub-DLIs: total 16 indicators.
Total RBF value available	USD 58 million total potential RBF disbursements over the four years.	USD 25 million for SSRP + just under USD 400 million for the five years of the ongoing SSDP (up-to-date as of late 2019, restructuring since may bring new finance)	USD 515 million over 6 years. Approx. one-third of sector aid financing.
Financial incentive cascading	Planned at all levels for PFM4R as PBAs, not for other programs. Missed DLI payments rolled to next year.	None, except performance-based school grant DLI. Rollover provisions for most JFPs.	Clearly designated DLIs at central level (MoE) and LGA level (approx. 1/3 rd available RBF funding). No rule for missed RBF payments.
Evidence generation & independent verification (IV)	Government systems for primary evidence. Different verification systems: no independent verification for KfW; independent consultancy firm for GPE (EY); government national audit institution for PFM4R (<i>Tribunal Administrativo</i>).	No specified methodology for government reporting, mixture of high/low specification extent. SSRP (WB) and SSDP: IVA role by Nepalese research institutions KUSOED and TU-CERID. SSDP: IVA required for WB, EU, and partially by ADB (other DLIs via JFPs endorsement of government report). Nepalese research institutions contracted by government for IVA role.	Government systems for primary evidence. High level of specification. EP4R focus on evidence generating system strengthening (DLIs, TA). Independent verification by consultants contracted by donors since 2 nd year. High level of specification, jointly (government/donor) agreed TORs.
Ancillary components (pure vs hybrid RBF)	PFM4R was hybrid instrument, including capacity development and coaching windows. KfW parallel TA. GPE linked to TA under FASE funding but no direct support.	TA provided specifically for RBF (EMIS support, studies) and more broadly for SSDP (e.g. EU & ADB funding British Council TA/CD facility) & non-RBF funders (UNICEF hands-on TA for range of processes).	EP4R has its own TA (known as Technical Assistance Support – TAS) contracted via DFID as a within-program component. Also there is other TA/CD including from non-EP4R funders.

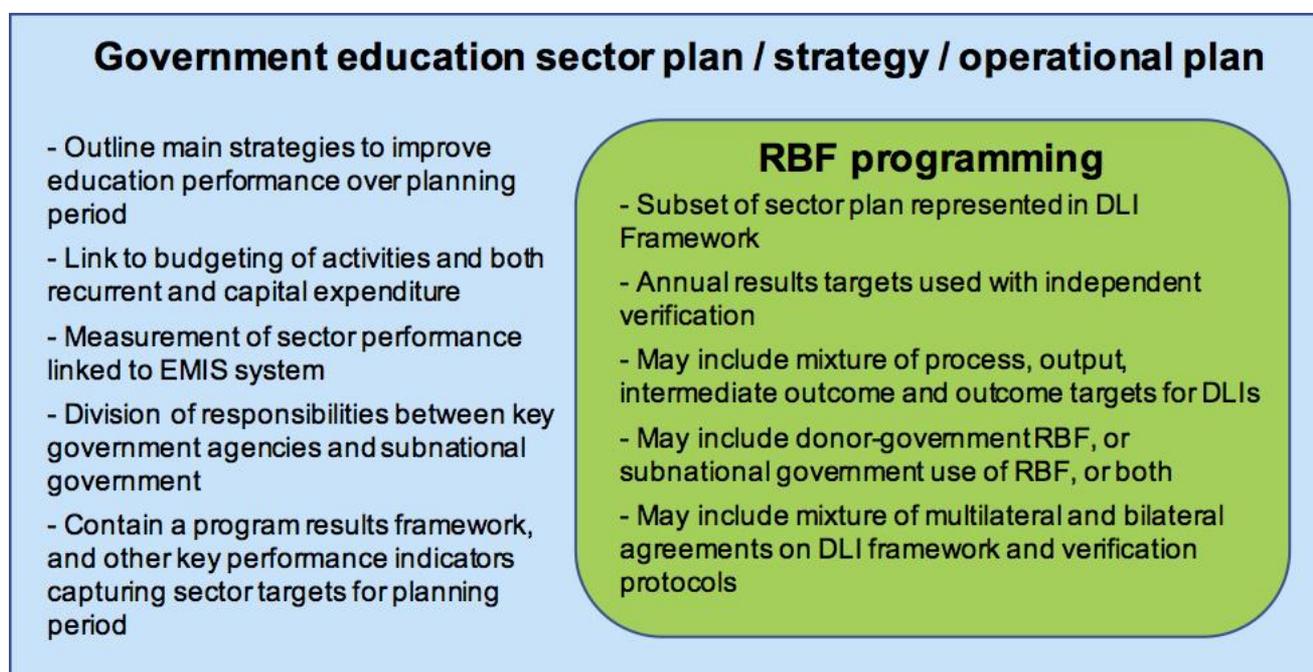
3. Findings

A. Design

57. The use of RBF in Mozambique, Tanzania and Nepal has been part of a shift in donor preferences with respect to financing instruments and modalities. This has included the growing use of the Program for Results (PforR) financing instrument for the World Bank, Results-Based Lending (RBL) for the Asian Development Bank (ADB), and the use of a variable tranche by both the European Union and the Global Partnership for Education (GPE); all linking disbursements to indicators representing sector results.⁵¹ During the same period, interest has grown in the use of RBF approaches in the headquarters of bilateral aid agencies such as the UK’s DFID (now FCDO) and Sida.⁵² The conditional nature of RBF financing has presented an alternative to general budget support (GBS) or sector budget support (SBS).⁵³

58. A consistent factor across RBF programs is that they involve a subset of sector finance provided in a way that is fungible.⁵⁴ As shown in Figure 3, RBF programming tends to include a DLI framework with annual targets covering a subset of the broader results required to achieve the goals of the country’s education sector strategic plan, and ongoing operational plans.

Figure 3 RBF design as a program within the sector strategic plan



Source: Authors

The development of education RBF programming in the three countries

59. The development of RBF across the three countries broadly followed the above model, with interlinkages between RBF programs and sector plans. However, the strength of this association varies across the RBF programs studied here:

⁵¹ The EU use of variable tranche goes back further than the other instruments, in use since at least 2000. Schmidt (2006).

⁵² These modalities can be contrasted to “traditional” input-based financing modalities, where activities and expenditure are accounted for more directly, independent to results achieved or with much looser conditionality (see Table 19 in Section 3H).

⁵³ A number of evaluations raised questions on the effectiveness of budget support in addressing systemic problems faced in key sectors such as education. See, for example, DANIDA (2014) for a summary.

⁵⁴ That is, funds are not restricted in terms of how they should be spent, or at least are broadly non-restrictive.

- In **Mozambique**, the 2012 Sector Strategic Plan set out some broad strategies that would be pursued by RBF, including a focus on vertical and horizontal accountability reforms. The World Bank's development of the PFM4R was an initiative linked to broader PFM reform, in which the education and health sectors were selected from a number of other potential sectors, with a focus on subnational PFM and service delivery. The GPE use of RBF in Mozambique was linked to the 2015–18 Operational Plan, and the development of the plan was a requirement for the GPE and World Bank additional financing to which it closely aligned. The German (KfW) use of RBF in construction (2013–16) was its own initiative, linked loosely to the government's Accelerated Plan for Construction.
- In **Tanzania**, the programming of RBF was closely associated with the preparation by the Government of Tanzania of a "quick-win-focused" "Big Results Now" plan for basic education (BRNEd 2013), squarely focusing on improving education quality and learning to address the politically highly sensitive "learning crisis". The BRNEd RBF program adopted the government's BRNEd goal and most of its key priorities. In its second phase, the RBF program, significantly expanded in 2017 and renamed Education Program for Results (EP4R), supported a number of priorities in BRNEd, but was less closely related to, the more traditional Education Sector Development Plan (ESDP) 2016/17–2020/21 which superseded the BRNEd, following the disbanding of the Big Results Now initiative by the 2015-elected President.
- In **Nepal**, the interlinkage with the sector plan was explicit, with the development of the SSDP, in which six joint financing partners (JFPs) tied finance to DLIs from the 2016/17 financial year. This compared to just one development partner using DLIs, the World Bank from 2013, in the previous SSRP period (2009 – 15/16), known at the time as incentive linked indicators (ILIs). The introduction of large-scale RBF financing in education was directly linked to the SSDP plan development, with a five-year costed plan from 2016/17 to 2020/21, and ten "DLI areas".

60. The ways in which the use of RBF was "introduced" in the sector, how initial decisions were taken to adopt this new way of working, and whether and how risks were assessed, varies across countries. As Box 3 sets out, in spite of a possible "pilot phase" in using RBF in Nepal, there seems to have been a more genuine opportunity in Tanzania to discuss how to address the challenges raised by the new way of working that RBF represented. In Mozambique, the use of RBF was focused in the PFM4R case with additional finance for the sector; but has not been as extensively used in the main education sector financing fund, the FASE, with use only by the GPE and Germany (KfW). As a result, a lower share of external education sector finance (that is, development aid) is linked to RBF in Mozambique than in the other two countries.

Box 3 Initial steps in using RBF in Mozambique, Nepal and Tanzania

Nepal was closest to having what might be called a “pilot phase” for RBF. This was with the small-scale use of RBF introduced by the World Bank in the SSRP period for the 2013/14 and 2014/15 years. This first use of RBF was not explicitly formulated as a pilot, and was not evaluated as such. Some stakeholders recall that the decision to use RBF on a much larger scale to finance the SSDP (from 2016/17), occurred mainly as the World Bank and Asian Development Bank (ADB) were keen to do so, using their new financing instruments. This brought other joint finance partners (JFPs) into the process, leading to the streamlining of the various bilateral mechanisms into a single Joint Protocol and Joint Financing Agreement, and to 80 percent of donor financing for the SSDP to be channeled as RBF. There does not seem to have been much discussion of the risk entailed by the MoE in shifting such a significant portion of external finance to RBF, in spite of the SSRP experience in which USD 5 million out of USD 25 million “at risk” was lost.

In **Tanzania**, the RBF modality fit well with the government’s focus on results through the BRN. However, there was some apprehension among both government officials and development partners that this new way of working, where disbursements follow results (for which therefore, government must find funding “up front”), might be problematic. This was initially assuaged by the use of “foundational DLIs”, focusing in the first phase on pre-requisites for the recurrent DLIs and seen as relatively easy to achieve (as “desk work”, for instance finalization and approval of the BRNEd budget framework for which releases would be incentivized through the recurrent financial DLIs), and in the second phase on developing strategies seen as priorities by the RBF partners (e.g. school infrastructure development strategy, and finalizing the design of the SQA system whose rollout would be incentivized through a recurrent DLI); and the agreement that technical assistance was indispensable and would be provided.

In **Mozambique**, the use of RBF resulted from relatively independent decisions taken by the World Bank for the PFM4R in collaboration with the MoF. Many sectors were considered for a program focusing on PFM in subnational service delivery, with health and education finally selected. The PFM4R therefore brought additional financing to the education sector. The GPE introduced its variable tranche in its 2015 education sector program implementation grant (ESPIG) with 30 percent of funding available using RBF (alongside 70 percent of non-conditional funding as a *fixed tranche*). Germany (KfW) sought to use RBF as an independent initiative, linked to longstanding technical support to construction. Some concerns were raised with regard to the use of conditional funding as part of the joint FASE funding, which was the case for the GPE and KfW RBF, though not the PFM4R. A 2016 review of FASE¹ indicated that whilst this had the potential to strengthen accountability for results, it was recommended to limit the use of such mechanisms, given the potential risk to predictability and flexibility in donor funding and the implications for FASE-financed essential education spending, including textbooks, construction, school grants and supervision.

¹ Orłowski (2016)

61. The mix of RBF funders varies across the three countries. In each, the World Bank and the GPE are among the mix, with the World Bank leading on the adoption of RBF in Tanzania, Mozambique, and Nepal (in the latter case, together with the ADB). Among the smaller uses of RBF, Germany (via KfW) in Mozambique wanted to “experiment” with RBF – an experiment which was not taken forward linked to its inconclusive outcome (see Section 3B). “Experimenting” was also an explicit objective of KOICA in Tanzania, which joined the RBF program in 2020 with a small amount of funding. In Tanzania, a reason to join the RBF program for GPE (in 2018) and KOICA (in 2020) was also the attractiveness of pooling resources behind a common set of objectives, and for KOICA, as a small donor, there was also the attractiveness of joining an influential set of larger donors. In Nepal, for smaller donors, there was also a view that RBF allows a partner to be “at the table” if funding the SSDP as a JFP, in terms of the voice and influence in joint sector working groups and in the structured annual review process. This would therefore give a donor greater policy leverage than would otherwise be the case, a significant reason to use RBF for donors spoken to for this research.

62. In Tanzania, DFID have continued to use “traditional financing” to support basic education, through a separate non-RBF project, alongside and of a similar size to their RBF funding.⁵⁵ DFID present this combination

⁵⁵ For DFID the EP4R (total £105 million) is “financial aid” whereas the EQUIP-T project (£90 million) is non-financial aid.

as a way of strengthening aid effectiveness through complementarity between RBF aiming to tackle systemic challenges, with more operational activities that support the implementation of RBF-incentivized reforms, allow innovation, and feed lessons from the ground into RBF-level discussions.⁵⁶ This “portfolio approach” is also present in GPE programming, with the fixed tranche-financed project supporting activities broadly aimed at the DLI results that GPE is incentivizing through the variable tranche, albeit with less emphasis on innovation and less systematic coordination than what is planned to be achieved with the DFID design. The GPE variable tranche in Mozambique was also programmed alongside a fixed tranche, and with the World Bank’s Education Sector Support Project (ESSP), aligned to the MoE’s 2015–18 Operational Plan with contributions to the FASE. Most development partners funding education in Mozambique have continued to do so through funding FASE, mostly not using RBF, though the World Bank and Finland use “virtual earmarking”, loosely tying their contribution to certain activities though not to DLIs.⁵⁷ Norway, JICA and UNICEF also fund the SSDP in Nepal but as non-RBF joint financing partners, and the GPE, the European Union and Finland provide a fixed tranche alongside their variable contribution.

The process of selection of DLIs

63. RBF as a donor modality faces trade-offs between “stretch” and the need for disbursement. An early assessment of the use of RBF by the World Bank found an “inherent tendency to shift the balance toward the disbursement objective” in the way DLIs were selected.⁵⁸ As seen above (Table 6 in Chapter 2), the study focuses on six RBF programs in three countries, hence five DLI frameworks. The DLI frameworks of the Mozambique World Bank PFM4R, GPE variable tranche (VT) and German (KfW) RBF programs, and the World Bank’s use of ILIs in the 2013–15 period in Nepal, were separate and separately negotiated between the government and the development partner concerned. The DLI frameworks of Nepal’s SSDP and Tanzania’s RBF program, while significantly larger and more complex than in Mozambique, were negotiated as a single framework by the funders as a group (initial funders in the Tanzanian case).

64. The process of DLI selection has tended to involve discussions between partners and government, with some variability in the degree of government “ownership” of strategies pursued and of the DLI framework(s). The DLIs used in Mozambique and in Tanzania had strong government ownership. In Mozambique, the MoF was involved in the design process for the PFM-focused PFM4R, which was not the case in Tanzania or in Nepal. In Mozambique, particularly for the GPE VT DLIs, the MoE was keen to ensure DLIs were manageable and negotiation involved some compromise, in a context of a sector that is highly aid dependent with FASE funding most non-salary expenditure.

65. The Tanzanian experience suggests that judiciously-used “foundational DLIs” may help in multiple ways – to assuage fears and stimulate the initial fund flow, but also, to achieve discrete actions as a necessary basis for implementing further measures needed to achieve some of the DLI results. In Nepal, many of the Year 1 DLIs are of a similar nature to the Tanzanian foundational DLIs (development of designs, guidelines etc.) and could be considered to have played a similar role, though less explicitly. In Mozambique, most of the DLIs could be viewed as, in a sense, foundational; this was seen as needed given the context in which RBF was new to the sector and country, and the sector being more dependent on external funding for essential expenditures.

66. In Tanzania the government-led BRN meant there was close collaboration between senior education officials and key development partners, over the course of several months, to identify priority pathways to improve the quality of teaching and learning, resulting in very strong ownership of RBF by government. By the time of the program redesign this ownership had deepened as RBF had been found to be not only “workable” but effective – including in terms of much higher disbursement rates than contemporary or recently used

⁵⁶ The design of a DFID successor program in education intends to strengthen the link between innovation and reform by bringing the two components under one single program.

⁵⁷ Ten development partners have provided funds to education via FASE since 2014; only Germany (KfW) and GPE used RBF.

⁵⁸ World Bank (2016)

instruments. As in Mozambique for the GPE, over the program's life, education officials in Tanzania negotiated with development partners with a view to keep a "reasonable" balance between stretch and feasibility, for instance directing new GPE and KOICA funding towards indicators for which they assessed good results would be more likely to be maintained.

67. In Nepal, in spite of the simultaneity of the processes of developing the SSDP and the RBF DLI framework, the extent of "donor drive" behind the selection of DLIs is unclear, but appears to be higher than in Tanzania with different development partners having different priorities, as further unpacked in Box 4 below. As a result, the DLI framework is composed of two subsets of priorities for two subgroups of donors, with only limited overlap, the World Bank and ADB supporting one subset and the four other smaller funders the other.

Box 4 *Prioritization of different donors in the development of DLIs in Nepal*

In Nepal, there has been a longstanding and strong sector-wide approach (SWAp) through which government and development partners jointly work through sector working groups including in the development of sector strategies. The SWAp enabled what was described as an iterative and collaborative process of development of the DLIs for the SSDP program period alongside the development of the plan (the SSDP) itself. However, the process also involved substantial bilateral effort, and in many cases, may have reflected the priorities of development partners as much as government though stakeholders gave contrasting assessments on the extent of alignment. Some emphasised good alignment, whereas for others, development partners were keen to push their own agendas. The decision itself to use RBF was principally led by the World Bank and ADB through their respective PforR and RBL financing instruments. Other JFPs joined, resulting in single Joint Protocol and Joint Financing Agreement.

The negotiation process was then centered on getting the *right* set of DLIs together, splitting allocations and priorities between donors, and combining the DLIs into one Joint Protocol. The World Bank and ADB had the largest number of DLIs and the highest-value DLIs, followed by the European Union (EU) and two rounds of support from GPE. Finland and USAID are much smaller RBF funders, the latter joined later in the program period. In the DLI framework there are two fairly distinct subsets of priorities, some for the World Bank and ADB, and others for the other four funders. That said, there was heterogeneity within each subgroup too. There were key areas of overlap between the two banks' priorities (notably, teacher management and financial governance), but also areas that each had championed (e.g. monitoring of time spent teaching (TST) for the World Bank, Model Schools for the ADB). The other subgroup of RBF funders shared a focus on early grade learning, two of them also incentivized addressing the issue of out of school children (OOSC), and each also had their own priorities, such as on Grade 8 examinations for Finland only, and a focus on disaster risk reduction (DRR) by the EU only.

68. As set out in Box 5, the Mozambique PFM4R was an unusual case among those studied, with both a sectoral and PFM focus. It involved an interesting multi-sectoral design approach born out of earlier PFM reforms and responding to the need for improved PFM for subnational service delivery. The main program structure was based in the MoF, although there appears to have been limited cross sectoral work/learning in practice. In education this approach led to the selection of DLIs incentivizing the strengthening of key vertical and horizontal accountability links, thereby contributing to strengthen school governance, which was a central concern of education stakeholders. The DLIs reportedly were strongly owned in the MoE.

69. In contrast to the PFM4R in Mozambique, all the other RBF designs grew firmly out of the education sector strategies and priorities, with more limited engagement on education PFM/financing issues. In Tanzania the PFM focus was particularly limited, considering the size and ambition of the program, although two high-value "financial DLIs" focused on release of government funds for school grants and a few other "priority lines" in the education budget. The objective was to address the well-known weak budget credibility and unreliable budget releases affecting all sectors. However, arguably these high value DLIs were "low hanging fruit" with just 5 percent of the sector budget protected, and the government rewarded for any level of release, as RBF disbursements were scaled proportionally. In Nepal there was more significant engagement with financial governance issues through multiple indicators within DLI 9.

Box 5 Mozambique – development of sector RBF out of government PFM vision

The PFM4R program grew out of the government's PFM Vision (2009) which led to a USD 157 million program of PFM support (2010-14), which virtually every donor in the country supported. The recognition that this program did not sufficiently support subnational or sectoral PFM led to the development of the PFM4R (2014–18). The development of the program involved shortlisting sectors to work with, leading to the selection of health and education and the respective line ministries. It therefore brought “additional” funds to education, outside of the FASE pooled fund, that were new to the sector. Reforms were an attempt to improve PFM in a context of increasing fiscal decentralization, with district services then managing 60 percent of the sector budget.

Reforms sought to improve vertical accountability and PFM (accountability of provinces and districts to central government), as well as horizontal accountability (accountability of school directors and teachers to school councils, representing parents and communities). To ensure greater transparency in the use of funds at the district level, district services were upgraded to the status of budget management units in an attempt to improve transparency in classification of their budgets (previously the central MoE had little oversight of how district funds were allocated across subsectors / priorities). The program sought to unblock the long-standing bottleneck of slow disbursement of the school grants (*Apoio Directo às Escolas*, ADE), to ensure they arrived on time at the start of the school year, including the link to the MoF required for disbursement. The program's focus on school governance, supervision of schools by the district services, and the functioning of school councils, in part was then meant to ensure ADE funds would be well spent.

70. In all three countries, the logic underpinning the valuation of individual DLIs and their respective weights in relation to the total RBF funding was not explicit and not very clear – resulting in part from the relative fragmentation in the negotiations on DLIs' selection and values, even in the case of the more joined-up set-up of Nepal, and Tanzania at the initial design stage. In Nepal and Tanzania the range of values of individual DLIs was very large. This was less the case in Mozambique. In addition:

- In **Tanzania**, at program outset, among the six recurrent DLIs the financial DLI (school grant and priority budget lines releases) represented half of the RBF funding potentially available, around seven times the value of the “smallest” DLI. Towards the end of the program (early 2020) the range in DLI values had reduced, with nine recurrent DLIs and additional RBF funding available. The largest-value were still the financial DLIs – slightly less than a third of the total program funding, roughly double the DLIs next in value (equitable PTR), and still seven times the lowest-value DLI (school grant improvement). Whilst there was initially some logic in this allocation of development partners' funding, it was diluted by the process of repeated redesign of the allocation framework (mid-term review, additional donors, restructuring).
- In **Nepal**, the largest-value DLI area out of the ten DLI areas in the framework, was the World Bank and ADB-financed governance and fiduciary management DLI (including 19 specific DLI results across the five years). This was six times larger in value than the lowest-value DLI (EU-financed disaster risk reduction DLI). The very high financial value allocated to the governance and fiduciary management DLI (USD 81 million at the outset), may have been a risky strategy considering the imminent rollout of federalism, which was quite likely to affect the sector governance and PFM set-up. The level of other DLI values had a wide range but the main driver seemed to be which donor chose to fund them, as the World Bank and ADB brought higher values than the other joint financing partners (for example the World Bank had potential disbursements at an average of USD 7 million per result, compared to USD 1 million for Finland).
- In **Mozambique**, the valuation processes of the three programs were separate. For the PFM4R, the largest-value DLI was the school councils' strengthening, which may count as the most ambitious. In contrast for the GPE variable tranche DLIs the highest-value DLI was, arguably, the least ambitious (number of teachers trained). This split of DLI values may have partly resulted from government officials' efforts in trying to avoid taking on too much risk.

DLI range and positions on the “education results chain”

71. In both Nepal and Tanzania a broader range of DLIs were selected, covering a wider range of areas compared to Mozambique. The DLIs of the RBF programs can be categorized in different ways. Table 7 categorizes DLIs in relation to where each indicator is positioned in an “education results chain” ranging from availability of inputs, to learning as an outcome. Improved learning was indeed a DLI focus in Nepal and Tanzania (with a considerably higher funding value in Tanzania), and this was not the case in Mozambique.

Table 7 RBF DLIs positions on the “education results chain”

Levels in results chain	Mozambique	Nepal (SSDP period)	Tanzania
Outcome		Increased early learning levels (reading) Share of RBF value: 0.1%	Increased early learning levels (reading, math) Share of RBF value: 9%
Intermediate outcome		GERs; reduction in OOSC incidence Share of RBF value: 11%	SIG; retention, transition rates; school SQA score improvement Share of RBF value: 18%
Output	In-service teacher training, school director training (numbers); equitable PTR (GPE VT) Share of RBF value: 24%	Rollout of EGRP; activity kits for science, math and English; subject teachers training; Model Schools; scholarships; classroom retrofitting Share of RBF value: 20%	Equitable PTR (local/ school levels); STEP training (number) Share of RBF value: 16%
Process	School grant timeliness; district accounting system; district-level school supervision system rollout; school level accountability system rollout (PFM4R); trained school directors evaluated (GPE VT) Share of RBF value: 37%	NEB operational; new curriculum framework; new curriculum for G9-12; TST monitoring process; single subject certification; G8 exams; DRR safety masterplan; fiduciary system; grant system; EMIS accuracy Share of RBF value: 69%	Foundational, data management, SQA rollout, policy learning Share of RBF value: 24%
Input	School grant timeliness (PFM4R); construction (KfW) Share of RBF value: 39%		Financial DLIs inc. school grant (GoT budget releases), textbook availability Share of RBF value: 32%

Note: Authors assessment. In Tanzania, percentages are proportions of total RBF available incentivizing results at each level, for the DLI framework and DLI values towards the end of the program (early 2020). In Nepal, SSDP potential value. In Mozambique, percentages for RBF programs combined.

72. In Nepal, DLIs tended to be designed with linear sequencing, similar to project results chains, i.e. with targets successively moving from processes to outputs then to outcomes over the five years – that is, starting with system development, to system rollout, to quantitative outcome targets in the program’s later years. Achieving results thus required achieving successive steps on the DLI results chain, in a manner somewhat similar to the implementation of stand-alone projects. Given this approach, later years’ quantitative targets were fixed at the outset, even though specifications of how they would be measured and delivered were developed during the first year after targets were set. In practice the DLIs increased in complexity with the rollout of federalism, as the main responsibility to achieve results was moving from central-level agencies in earlier years to local-level agents in later years (see Section 3B). In Nepal and Tanzania as RBF funders selected different mixes of DLIs this also meant that they each incentivized a different mix of levels in the education results chain. However, considering the level of complexity of the DLI frameworks, this did not derive from a systematic prioritization process. This is illustrated in Box 6 below for the case of Tanzania. In Mozambique, the MoE managed to keep DLI selection predominantly at the lower end of the result chain – mainly outputs and processes – which were seen as manageable, and for this first phase in the use of RBF at least the donors accepted this.

Box 6 Tanzania – RBF funders’ mixes of incentivized results by level

The EP4R aimed to incentivize a blend of input, process, output and outcome changes. However, the balance of RBF funding across result level repeatedly changed, at program redesign with additional DLIs and funding from existing RBF funders, then when GPE and later on KOICA joined and selected some of the existing DLIs, and also through several reallocations of unspent RBF rewards by some of the funders, at different points in time. As a result of this constant evolution, the balance of funding across levels of incentivized changes is unlikely to have been monitored.

Although the DLI framework was largely common, initially at least, development partners had different priorities. There was a significant emphasis from DFID on RBF budgetary support aspect with a high allocation of RBF funding to comparatively “easy to achieve” results on financial DLIs, at input level. The World Bank funding is in two distinct phases, which focus on quite different things; in both phases the financial DLIs (input level) are important, but at the redesign stage a strong emphasis emerged on school quality assessment (SQA) (process and intermediary outcome levels). The World Bank and DFID had the largest shares of RBF funding on the early learning targets (outcome) and on the financial DLIs (input/process), at either ends of the results chain, and on the SQA system (process, intermediary outcome). Other donors paid more attention to books on desks and students at schools. Sida had the strongest emphasis on input-based targets but also on retention rate at intermediary outcome level.

73. In Tanzania, the DLI framework also moved upward in the education results chain over the program life. However, unlike in Nepal this was done mainly at the point of redesign by adding new DLIs targeting results higher up in the education results chain (improved retention rate, improved school quality score, and one additional learning sub-indicator). The complexity of the DLI framework also increased over time as all new indicators relied on sophisticated result measurement approaches, even those targeting lower level results such as a DLI which required school-based grade- and subject-specific data on textbook availability to measure results. In the case of the data DLIs, the formulation of the DLIs did not change but the specifications of what was deemed “acceptable” in terms of coverage of the school-level statistics to be collected and of the analyses to be presented in the annual sector performance review report became more ambitious year-on-year.

Thematic coverage of DLIs

74. The Mozambique and Nepal RBF designs engaged more strongly on institutional and governance changes, with a strong focus on school/subnational PFM and accountability in both designs. In Mozambique this was the specific focus of the PFM4R, linked as seen earlier to broader PFM reforms for improving service delivery through stronger sectoral PFM. The need for this was reinforced by the findings of a World Bank Service Delivery Indicators (SDI) survey in 2014 which highlighted the importance of strengthening school-level accountability, not least to improve teacher presence in school.⁵⁹ In Nepal, as discussed in Section 3B below, this strong governance focus, with an associated large volume of RBF rewards, became problematic with the rollout of federalism and the associated shift from a deconcentrated education management model to a decentralized model in which LGs are responsible for school management. The Tanzanian RBF design engaged less strongly in this area, and altogether more lightly on school management than in Mozambique and Nepal.

75. The Tanzania RBF design engaged much more strongly on strengthening evidence generating systems, both in terms of focus of specific DLIs and in the extent to which the measurements of results on other DLIs relied on one of the targeted systems. Three foundational and two out of nine recurrent DLIs incentivized the development, strengthening and rollout of the school statistics management system (EMIS), the school quality assessment system (SQA), and the learning assessment system focusing on early grade learning (EGRA/EGMA), respectively. Moreover, for seven of the recurrent DLI targets the result measurement relied exclusively on the EMIS, the SQA system and the EGRA/EGMA. In value terms, “system” DLIs represented just less than 15 percent of the total RBF funding (EMIS, annual report, SQA rollout) and a much larger RBF amount depended on EMIS data availability. In Nepal one DLI area – the second smallest-value, representing just 5 percent of the total –

⁵⁹ World Bank (2015a).

focused on strengthening the EMIS, but other DLIs largely relied on DLI-specific reporting. In Mozambique there was no focus on the EMIS, and while a DLI aimed to strengthen the system of district supervision of schools, the DLIs did not require the analysis of the information generated (see Section 3E).

76. All three country DLI frameworks included some focus on education quality. In Mozambique the focus was relatively limited, through the GPE DLI on the number of teachers given in-service training, although accountability reforms had longer results chains with a view to quality as a goal. In Nepal, several DLIs focused broadly on quality including curriculum reform, learning materials, EGR, examinations and assessments, teacher training among others. In Tanzania, textbook availability along with school quality assurance in Tanzania also aimed to contribute to improved quality. Importantly, there was a focus on learning outcomes in both Tanzania and Nepal; with significantly higher value in Tanzania, with 9 percent of the total RBF potentially available, compared to just 0.1 percent of the RBF value linked to learning in Nepal.

77. Table 8 presents a second method of classifying DLIs, looking at which technical area each DLI aimed to impact. The table focuses on the most direct/strongest link between a DLI and an area. Some of the DLI areas cross-over different categories and are linked to both in the table below.

Table 8 DLIs by technical area

Technical area	Mozambique	Nepal (SSDP period)	Tanzania
Teachers	Equitable PTR across districts (lower primary). Trained school directors evaluated. Teacher qualification (training). (All GPE VT)	Improved teacher management through TST monitoring. Teacher (re-)deployment. Subject teacher training. Schools with full complement of math, science and English teachers.	Equitable PTR. 2 nd phase (foundational) – Updated teacher deployment strategy.
Governance and PFM	School grant timeliness. District accounting system. District-level school supervision system rollout. School-level accountability system rollout. (All PFM4R)	Improved governance & strengthened fiduciary management (procurement, accounting, school grant system). Enhanced reliability & transparency of EMIS system.	Improved school quality score (2/6 SQA scores). Reliability of school grant funding; GoT budget releases for priority lines (5% total sector budget). School stats, annual sector performance review report.
Access	Construction (KfW).	Model Schools. OOSC enrolment in most disadvantaged districts. Poor student and girls' secondary scholarships. Disaster risk reduction.	2 nd phase (foundational) – Strategy for school infrastructure development; inclusive education strategy.
Equity	Equitable PTR across districts (GPE VT).	Poor student scholarships. OOSC enrolment in most disadvantaged districts.	Equitable PTR. Girls' primary to secondary transition rate. 2 nd phase – inclusive education strategy.
Quality	Teacher training (GPE VT).	Model Schools. Curriculum revised.	Textbook availability (2 nd phase).
Learning	None.	Reform of exams & assessment system. Early grade reading program and learning outcome.	Retention & transition rate. Improved early grade reading & math skills. School Improvement Grant.

Note: For most DLIs, achieving results might also have less direct/large effects on a number of other technical dimensions. This is not shown here. For Tanzania this excludes foundational DLIs; and reflects the final DLI framework and DLI range of values (as in early 2020).

78. Equity, whilst an issue in all three countries, was only marginally covered in the RBF design in Mozambique and Tanzania. In Mozambique, the only equity indicator was for district-level PTRs, and the DLI as

designed focused on equity *within* provinces rather than *between* provinces. The Nepal RBF design presented some internal inconsistency with regard to equity, with a number of equity-focused indicators (OOSC enrolment, scholarships for poor students) but also a counter-equity focus on 200 “Model Schools” representing roughly 1 percent of public schools in Nepal and likely to be selected from among the already better resourced schools located in more urban areas. In Tanzania where financial incentives cascaded down to local government level, there was also no attention to inter-LG equity in the design and implementation of the cascade and reward allocation systems. Moreover, only Nepal addressed demand-side issues on a targeted basis (scholarships for poor secondary students; “targeted interventions” to bring OOSC to school in more deprived districts).

79. School financing was a DLI-targeted area in the three countries, in different ways. In Mozambique the focus was specifically on increasing school grant timeliness in relation to the start of the school year, which had initially been poor. In Tanzania a similar DLI focused on the completeness of school grant release. The shift from channeling the grants through LGAs to a system of direct MoF-school channeling addressed both completeness and timeliness issues; however, towards the end of the program timeliness was again deteriorating, with late releases from MoF, and this was not captured by the DLI. In Nepal the ambition was to reform the whole school grant system including in terms of allocation, and school accountability. Issues such as the adequacy of the school funding level and/or of the grant allocation system were not considered in Mozambique and Tanzania, and there was no attempt at incentivizing the “next step” of improved use of school funding in any of the three countries, except indirectly through the variable indicators focusing on school governance.

RBF focus on teachers

80. In all of the three countries there was also a focus on teachers, including a specific emphasis on improving teacher deployment, though with different approaches taken (see Box 7). This took place in contexts of complex political economy surrounding teacher management in all three countries; and in which teachers’ motivation and effectiveness was affected by a range of issues going well beyond (re-)deployment. The results of these initiatives are also looked at more closely in Section 3B.

Box 7 Strategies for tackling teacher deployment

Diagnostic – The inequitable deployment of the teachers across different geographical areas was known to be a long-standing issue in all three countries at the time of design of the RBF programs.

- In Tanzania it was shown to affect both deployment across LGAs, and across schools within the same LGA, as teachers found ways to be posted in comparatively more attractive/less remote areas at both levels. The issue was perceived to be particularly severe in primary schools, with secondary schools generally better staffed at least in terms of numbers of teachers, although many faced gaps in specific disciplines.
- In Mozambique, the issue of teacher deployment was particularly acute at lower primary level. At that level, the country’s average PTR had fluctuated over time, with a peak following the abolishment of school fees in 2006, followed by a steady improvement until 2011, and plateauing since then. However, teacher deployment was historically strongly imbalanced, with central and northern provinces facing much higher PTRs than southern provinces. Within provinces some districts faced an even starker lack of teachers, and very high average PTRs.
- In Nepal, there are four main types of teachers, with different pay and conditions often linked to different qualification levels. Three of these groups are paid via central government grants. The fourth group is paid from the schools’ own resources; these teachers are often the least qualified yet they are indispensable, representing a third of the teaching force. The highest-paid, permanent tenure teachers have historically been the most difficult to deploy equitably in part due to teacher preferences to locate in more urban, relatively more affluent areas. In addition, population density in the country is highly variable, with more sparsely populated mountainous regions comparing to the densely populated Terai regions; and this compounds challenges in the equity of deployment. The long-standing issue of rationalizing teacher deployment became even more complex with federalism, so that in 2017 the respective roles and

responsibilities in relation to the management, deployment, hiring and firing of teachers were somewhat unclear and contested.

RBF strategies for tackling teacher deployment and indicators – The challenges across the three countries therefore had their own distinct characteristics, though leading to the same result, inequitable PTRs. Tanzania and Mozambique used the PTR itself as the measure, while Nepal approached this through a more complicated measure:

- In Tanzania the DLI was designed in reference to an acceptable PTR range. One target result focused on bringing an increasing number of LGA-level PTR within this range, the other focused on the same goal for schools within each LGA. LGAs had to use some of the RBF rewards they were earning (for this and a number of other DLIs) to finance teacher transfers – although, unlike with the foundational DLIs for the national level, there was no funding to “prime the pump” for LGAs and they had to initially find funds to transfer teachers from other sources. At redesign, a foundational DLI incentivized the finalization of a comprehensive strategy with a view to further improving teacher deployment. The strategy involved the recentralization of posting of new teachers by PO-RALG allocating them down to school level, through an IT system whose development the RBF-associated technical assistance supported.
- In Mozambique the focus was on lower primary PTR, and the GPE variable tranche DLI was to reduce the number of districts with average PTR above 80, with a target of reducing from 12 to 2 districts in the category for the 2014–17 period. By focusing on a within-province measure, the DLI largely left aside the imbalance in PTRs between provinces. This was to be done during a period when budget constraints put limits on new teacher recruitment at province level.
- In Nepal the DLI focused on rationalizing the deployment of teachers – incentivizing in the first year, the preparation of a policy, guidelines and a plan to deploy and reallocate teachers. Thereafter, the DLI incentivized action based on the approved policy and plan. It was expressed as a percentage reduction of the number of teachers needing “to be redeployed”, a fairly complicated indicator that only indirectly related to the geographical equity of teacher deployment. The DLI was to be measured in Years 3 and 5 – and Year 3 turned out to be the first year where the rollout of federalism seriously affected all DLIs concerned by the shift from deconcentrated to decentralized basic education management (see Section 3B).

81. The strategies to address teacher deployment were only part of the engagement of RBF with teachers. Nepal’s SSDP included four different strategies within its teacher DLI (DLI 3) including deployment. The others were targets incentivizing a teacher “time spent teaching” (TST) monitoring system, as well as the training of, and schools with a “full complement” of science, English and math teachers. In Tanzania there was less attention on teachers outside of the PTR DLI, although there was a DLI aiming to raise teacher motivation by eliminating teachers’ non-salary arrears – the budget line to this effect was one of those prioritized in terms of ensuring releases of the government budget, through the financial DLI. Tanzania’s RBF program design also initially incentivized the training of teachers on the new early grade curriculum, but the DLI was discontinued at redesign. Teacher pay had also been tackled in Nepal using an ILI during the SSRP period, which was the proportion of teachers paid via bank accounts, a focus that was discontinued for the SSDP period.

82. Some other DLIs focused on the accountability of teachers more generally. The DLI-associated school supervision systems in Mozambique and Tanzania included a focus on teachers’ effectiveness as part of what would be assessed, although this did not mean that teacher absenteeism would be sanctioned if detected. In Mozambique supervision was part of a broader push from the top to address teacher absenteeism, which had been highlighted as severe in the 2014 SDI report with one of the highest rates in the world and which the accountability reforms ultimately sought to address. In Mozambique, political appointment of school directors was a well-known challenge. The GPE variable tranche DLI on training school directors was developed in part to “do something” with the rationale that the political challenges were too intractable and perhaps training of school directors could at least give them some skills for the role with an outside chance of improving their motivation.

83. In Tanzania, teacher promotion was reportedly strongly affected by nepotism and corruption going well beyond the local level and reaching right up to central-level agencies involved in the process. Later in the RBF program period the implementation of the FFBEF raised a major challenge of very large classes, as too few new teachers could be recruited. In Nepal teacher selection and deployment was said to be a “major political issue”,

with the better paid permanent positions centralized for recruitment and selection and the most-qualified permanent tenure teachers concentrated in relatively better-off areas. Teacher management in Nepal is also complex considering the four different types of teachers, the historical imbalances between geographical areas, and the difficulty of addressing this through a central plan in the federal set-up. In Mozambique the GPE variable tranche included an in-service teacher training DLI to address the general under-qualification of teachers, in recognition of previous pre-service training systems requiring little training for entry requirements, a system that is being addressed but still with major capacity constraints. The teacher training model is dependent on a “cascade” by which teachers are trained at IFPs and then train more teachers at the level of clusters of schools, and an additional cascade to the school level itself. This model has existed in various forms in Mozambique for some years but with insufficient budget or evaluation of effectiveness.

Overall design considerations

84. Overall, the RBF programs in Tanzania and in Nepal (from 2016), had much broader designs than in Mozambique. Both targeted improved learning, an outcome at the “top” of the education results chain with a much greater share of funding on learning outcomes in Tanzania. Tanzania included a few specific links on priority pathways towards this outcome, mainly processes (SQA, data, financial), as well as access (equitable PTRs and retention). In Nepal the target results of RBF were conceived as successive steps on a range of self-contained thematic areas including many process and output indicators, with around half the DLI areas focusing on quality (EGR, curricula, assessment, teachers, Model Schools) and the other half on either access (OOSC, scholarships, retrofitting classrooms) or system-level improvements (PFM, governance, EMIS). In contrast, in Mozambique the three programs exclusively focused on process or outputs, with different pathways towards improving quality: school governance, PFM and accountability (PFM4R), teacher training and school director training and PTRs (GPE VT), and improving and speeding up improvements to classroom infrastructure (Germany KfW).

85. In Mozambique, the GPE focus on teachers and school directors was seen to complement the PFM4R focus on school governance whilst not duplicating it. There was less explicit coordination with the KfW program which was a separate bilateral approach within the broader multilateral support from FASE. In Tanzania, the question of why the specific links and pathways focused on by the program were selected for the first phase of the program finds its answer in the BRNEd design process which, as seen earlier, was unanimously said to have been strongly government-led with strong ownership of the resulting priorities. For the second phase the rationale is less clear: the DLIs were selected to incentivize responses to plausible concerns emanating from the decision taken by government to implement the FFBEP, but other responses would have been equally (and perhaps more) relevant, for example, if there had been more emphasis on in-service teacher training. In Nepal there was less of a single design priority beyond a series of reforms tackling quality, i.e. the DLI framework covers a range of technical areas, capturing many activities and strategies within the SSDP.

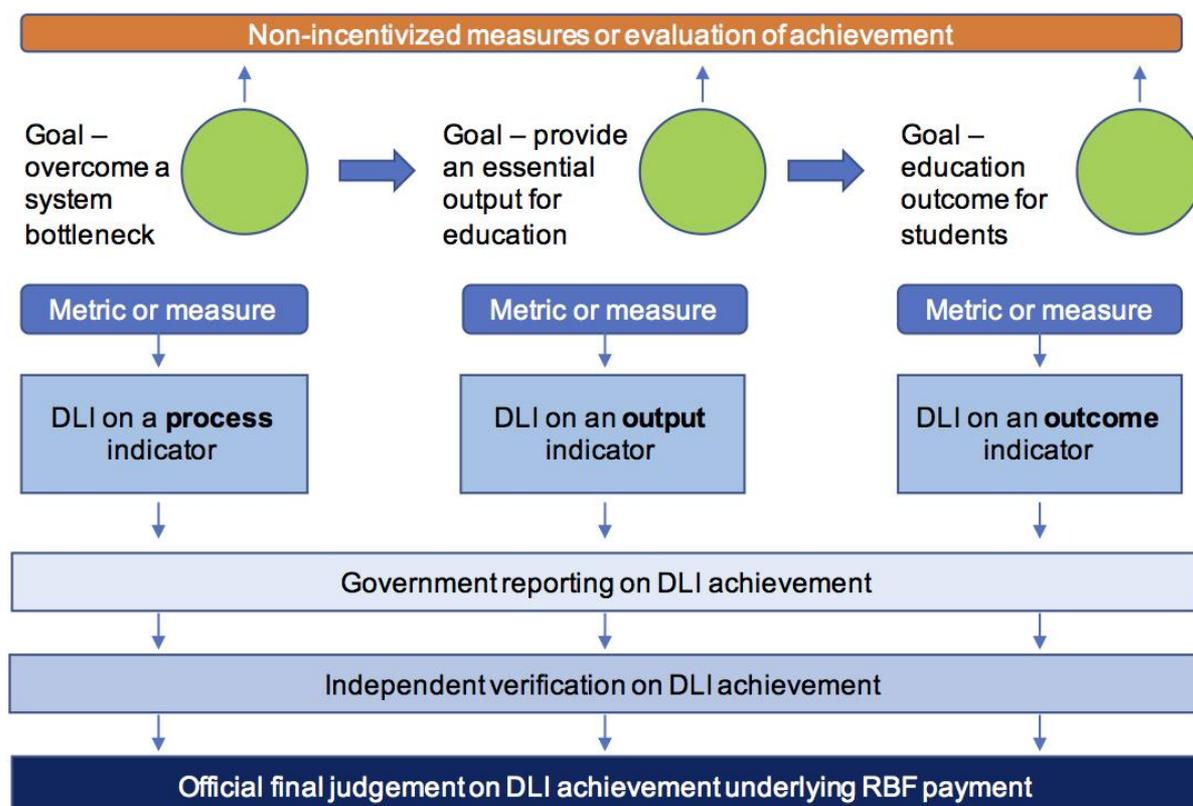
86. In Tanzania, a few years into program implementation the RBF-associated technical assistance provided system development support to implement government decision to recentralize new teacher posting to the school level and the streamlining of the teacher promotion system, which could help preventing nepotism and corruption in the posting and promotion process.

87. It is important to note that in all three countries, the RBF programs were founded on theories of change that were “incomplete” in several ways. This includes features such as, causal links between DLI-targeted results, higher-level goals, and improved students’ outcomes, not being fully outlined for all DLIs. This is the case at lower levels in the education results chain, where goals are often left implicit, for example the financial DLIs of Tanzania; or the use and expansion of Model Schools in Nepal, for which there were competing rationale; and finally, multiple theories of change for the same RBF DLI frameworks in the case of Tanzania. This must be borne in mind in the discussion of results in Section 3B below.

B. Results and impact

88. The designs of RBF programming across the three countries under review provide a wide range of different “results” as DLIs, with various mixtures of process, output, and higher-level indicators including those representing student outcomes. Results at different levels provide alternative means for programs to define success and impact on education systems. Where “results” are simply defined by the performance of DLIs, this may align with the “goals” of the system when the results are at outcome level, for example measures of students’ educational attainment. Where DLIs are at the process or output level, the goals represented may be more about overcoming a particular bottleneck, for example a specific reform to accountability mechanisms or PFM; or represent an important input or output that is a necessary (but not sufficient) condition for student outcomes, for example the deployment, training and motivation of teachers, or the timely provision of school textbooks.

Figure 4 The different means of assessing results and impact for RBF



Source: Authors

89. Any specific DLI may therefore represent a goal in itself or be a stepping stone to another goal, or both. Impact can be assessed in terms of whether these different goals are achieved. This section provides an overview of the achievement of DLIs, and also the inter-related goals, as well as the sustainability of changes experienced to date. We also assess the degree of contribution of RBF to any changes experienced, in terms of the financial incentive as well as accompanying components of program designs, such as technical assistance. Figure 4 shows some of these different means that can be used to assess results and impact, and it is important to note that the various sources may not always align – that is, a DLI could be “achieved, verified and paid upon” in cases where non-incentivized measures find that the connected goal was not achieved, or indeed vice versa; government reporting on DLI achievement could conflict with the assessment from independent verification; and a DLI and goal could be achieved unambiguously, without this leading to the next goal in a results chain.

The achievement of DLIs

90. The achievement of DLIs across the three countries saw some notable successes, although in each case some results have not been reached:

- **Mozambique** had the fewest DLIs across the three countries and also overall had the highest success rate in achieving them. This included the PFM4R program, with its four DLIs all achieved by the end of the program period and the full potential disbursement made. The four GPE DLIs were also all considered to be achieved with the exception of the PTR indicator, which only partially reached the target set (losing USD 3 million out of a possible USD 4 million). For the German (KfW) use of RBF in construction, the targets were either not met or partially so.
- In **Nepal**, performance was mixed for the SSRP World Bank (four) ILIs in the 2013–15 period. The ILIs on teachers receiving salaries through bank accounts, and on the presentation of NASA learning assessment findings, were fully achieved. The ILIs on the proportion of students receiving textbooks within two weeks of the start of the school year, and on the delivery of scholarships via bank accounts, were partially met. As a result, in total for the ILIs, 20 percent of the “at risk” amount was lost. The SSDP period for Nepal (2016/17 – 2020/21) includes 84 results in total and thus has the largest number across the three countries, although not the highest value. All results for its first year were achieved, in which central government-level process indicators predominated. However, from Year 2 (2017/18) a greater number of results required action at the local level and this led to results not being achieved or being seriously delayed. The second year DLI targets were more complex, but also coincided with the concurrent challenges of federalism, with the structures, responsibilities, and funding relationships with the new local government level undergoing major reform. It is likely delays to results were linked to these profound changes to the education sector’s institutional underpinning, which led to adjustments to targets and the DLI framework (see Section 3F).
- **Tanzania** has had two main phases of programming since 2014,⁶⁰ with fewer, more consistent and more quantitative targets than in Nepal or Mozambique. The DLI framework for Tanzania also included more higher-level indicators including outcomes such as learning to which 9 percent of the RBF funding potentially available was linked even though these potentially were more challenging to achieve, and targets for reading and numeracy were not met in Year 4 of the programming. There has been some late adjustment to the level of ambition for some key indicators, in light of radical change in parameters of the stretch vs. feasibility balance with respect to the FFBEF. However, challenges underpinning deteriorating trends are linked to the government’s continued underfunding of basic education in spite of the stated FFBEF, and the DLI framework does not itself address this even with the financial DLIs in place (see Section 3F).

Achievements of higher-level underpinning goals

91. The DLIs link implicitly or explicitly to higher-level goals, and at outcome level the metrics themselves are more direct measures of these goals, particularly the early learning DLIs in place in Tanzania and Nepal. However, most of the DLIs have longer chains to achieve such goals, often seeking to improve sector accountability or PFM, or other inputs and outputs that could be viewed as part of an “education production function”. Bearing in mind the point made above (Section 3A on Design) about the relative weakness of RBF programs’ theories of change, including in terms of links between DLI results and higher-level goals, Table 9 sets out our review of the achievement of the DLIs, interlinked goals, the contribution of RBF, and the sustainability of changes. Some highlights of this analysis include:

⁶⁰ Two of the five RBF funders, GPE and KOICA, had different programming timetables but “joined” the existing RBF program through selecting specific DLIs in discussion with the government.

- For Tanzania’s EP4R, the goals expressed by DLIs included a combination of the ambitious learning and retention goals, goals on equitable distribution of teachers, and some more process, output and even input indicators including on completeness of grant flow to schools, textbook availability, the EMIS system, and more recently on SQA visits. There is some clear contribution of RBF to improvements in these areas. However, even where results on DLIs were good for instance on the data/EMIS DLIs, the achievement of the (implicit) higher-level goal of a better use of data, is “work-in-progress”. In the SQA case, the level of ambition of DLI targets with regard to the pace of the rollout may actually undermine the potential of the system to improve school performance evidence and school performance (if visits are rushed etc.), and it is not clear how improved school performance evidence, if it is achieved, will inform policymaking. Moreover, and most significantly, in several other DLI areas there were challenges in leading to deeper changes, particularly in terms of the fiscal landscape in light of FFBEF, with declining teacher numbers, evidence that, in primary schools especially, the learning environment has deteriorated, and mixed evidence with regard to learning levels, perhaps understandable given the changed profile of pupils with the FFBEF. In this context also, the school grant completeness DLI results were showing a deterioration trend, and the (implicit) higher-level goal of at least maintaining in real terms the value of the grant, was not achieved.
- In Nepal, the SSDP program is still in implementation, although now in its final year (2020/21). The DLI framework covered a large range of technical areas and has faced major implementation challenges with many delays to DLIs. This has led to evolution and changes to the framework particularly for those PFM indicators most affected by federalism. Where process and output DLIs have been achieved there is limited evidence or evaluation of effectiveness, for example on scholarship schemes, Model Schools or new curricula. As a result, it is difficult to see the achievements to date as making the intended strides to improve education quality.
- For the Mozambique PFM4R program, the overall goal was to reform subnational and sectoral PFM, with improvements to both vertical and horizontal accountability. This included to district supervision and the timeliness of ADE payments to schools, both requiring cooperation between MoE, provincial and district administrations. Direct reforms measured by DLIs were achieved, however there are many questions on the quality of supervision, while the “last mile” of ADE in terms of its expenditure at school was not measured by the DLI. Qualitative research, including for this assessment, has found that changes to accountability at school level, including school council functioning, have been limited. The DLIs represented important steps but would, however, require many additional reforms to have a strong chance of achieving the goals of improving the learning environment across Mozambique’s primary schools.
- Mozambique’s GPE DLIs also did not adequately address the quality of teacher and school director in-service training, although the targets for numbers trained were achieved. A DLI target on the evaluation of school directors was met, but it is not clear this has led to any improvements in quality. The biggest tangible improvement may have been on the PTR DLI (see section below), which our quantitative analysis found appeared to lead to improvements in within-province equity in PTRs, although no effect to between-province equity and the broader constraints leading to slow rates of teacher recruitment since the 2015–16 period.
- The use of RBF in construction by KfW in Mozambique saw mixed achievement on indicators; the intended speeding up progress on classroom construction has not taken place.

Table 9 Summary of DLI results, their achievement, and the achievement and sustainability of associated higher-level goals

DLIs	Intended higher-level goals	DLI results	Achievement of higher-level goals	RBF / DLI contribution to results	Sustainability
Tanzania EP4R/BRN (2014-)					
Financial DLIs	- Increased Basic Education funding, changed MoF behavior.	- Good then deteriorating slightly. More challenges on Capitation Grant (CG) release than for budget framework.	- Not led to increased funding or maintaining real terms funding, at both school and sector level. Context of FFBEF and challenging fiscal context.	- Budget Framework DLI as designed cannot deliver deeper change. Possibly contributed to direct CG channeling.	- Relative priority to CGs likely to endure, not likely for other lines in EP4R budget framework once EP4R reward stops.
Data management DLIs	- Increased use of better evidence in planning & policy.	- Good, almost 100% achievement across all years on EMIS data collection + sector report.	- Work-in-progress: constraints on using information in policy & planning remain to some extent.	- Strong contribution of the DLIs, with RBF-associated technical assistance as well.	- Yes, though politically-sensitive nature of data-based policy & planning and absence of third-party verification may affect this.
SQA DLIs	- More effective, bottom-up school oversight; improved school quality.	- Recently rolled out for 2018/19, target nearly achieved (89%).	- Encouraging school-level findings (RISE, 2019e); unclear if highly complex system can be effective in shaping policy.	- Too early to say.	- Unclear if resource-intensive system and “school support” (vs “inspection”) approach can take hold.
Equitable PTR DLIs	- Improved teacher motivation & fairness in learning opportunities.	- Good then deteriorating - Between-LGA DLI initially better than within-LGA DLI, partly reversed from Year 4.	- Work-in-progress: distribution improved at local (across schools) level; not yet at national level (across LGAs).	- Moderate. Greater attention to workforce, but LGA DLI puts pressure at wrong place in the system after FFBEF.	- Unlikely that gains in teacher deployment equity persist if cap on teacher number/ basic education underfunding persist.
Retention DLIs	- Internal efficiency, gender equity, higher (total) learning.	- Girls prim-sec transition targets met; Year 5 (2018/19) target for survival rate partially met.	- Work-in-progress, further progress possible with area-specific data analysis and planning.	- Moderate. Cascade unclear, in terms of strategies LGAs expected to employ.	- Unlikely to maintain retention gains if underfunding/cap on teacher numbers persists.
Early learning DLIs	- Improved (early) learning outcomes.	- Good results in 2015/16 (EGRA/ EGMA), but targets not met in 2017/18.	- Work-in-progress in some aspects, ultimately requires unblocking funding and more focus on result analysis and continuous improvement.	- Moderate. Shifting incentives onto policy learning might be more important.	- Unlikely to maintain if underfunding persists + if adverse results cannot be discussed. - Overcrowding will persist in lower grades + carry up system in cohort.
Nepal SSDP (2016–21)					
Early grade reading	- Improved literacy in early grades + use of multi-lingual education. Directly linked to EGRP (USAID) project.	- Delay and difficulty in moving beyond the 16 districts of EGRP (USAID) from Year 3 (2018/19).	- No results yet, project rollout local, so unlikely to have measurable effect nationally - Slow progress on multi-lingual education.	- Achievement to date dependent on USAID work in the 16 districts. Delays in extending to new districts.	- Unclear legacy of EGRP or NEGRP at this stage - Funding provided for district rollout very little.

DLIs	Intended higher-level goals	DLI results	Achievement of higher-level goals	RBF / DLI contribution to results	Sustainability
Curriculum + school materials	<ul style="list-style-type: none"> - Improved curriculum - Learning materials for math, science, English. 	<ul style="list-style-type: none"> - National Curriculum Framework approved - Activity kits moved to online materials in DLI revision. 	<ul style="list-style-type: none"> - Progress in G1–G3 curriculum with 2019/20 pilot - G9–G12 curriculum taking longer. 	<ul style="list-style-type: none"> - Progress on curriculum but viewed as “politica” by civil servants, a sensitive DLI area. 	<ul style="list-style-type: none"> - Improved curriculum could have long-term benefits, but not clear if any evaluation designed.
Teacher management, accountability, and deployment	<ul style="list-style-type: none"> - More equitable teacher deployment - More accountability for teachers - Subject expertise. 	<ul style="list-style-type: none"> - Delays on subject teachers and training, and time-spent teaching (TST) monitoring - Cabinet approval on redeployment in 2019. 	<ul style="list-style-type: none"> - No change to accountability of teachers - Shortfalls of teachers in certain areas not addressed - No evaluation of training. 	<ul style="list-style-type: none"> - Slow progress. Complex results, particularly in the context of federalism, with teacher accountability a contested area. 	<ul style="list-style-type: none"> - Awaiting Federal Act to clarify roles on teacher management - No system-wide change to training systems - Unclear TST is effective.
Assessment and examination system reforms	<ul style="list-style-type: none"> - Single subject certification for higher secondary - Analysis of NASA. 	<ul style="list-style-type: none"> - Certification targets reached - Challenge on G8 exam done by LGs. 	<ul style="list-style-type: none"> - NASA reports strong on analysis of equity & learning, not clear if policy responses to this. 	<ul style="list-style-type: none"> - Results achieved but challenge to ensure consistent G8 examinations by LGs nationwide. 	<ul style="list-style-type: none"> - Unclear if goal of tackling rote-learning culture addressed - Unclear how results of NASA taken into policy.
Model Schools	<ul style="list-style-type: none"> - Model Schools as flagship secondary schools. 	<ul style="list-style-type: none"> - Delays but “numbers of Model Schools” target of 200 reached. 	<ul style="list-style-type: none"> - Definition of “model” unclear - No evaluation of benefits. 	<ul style="list-style-type: none"> - High government priority aligned with DLI incentive. 	<ul style="list-style-type: none"> - Unclear “finish line” for investment - Very high resource allocation, unclear if sustainable.
Interventions to reduce out of school children (OOSC)	<ul style="list-style-type: none"> - Reducing educational inequalities in the most disadvantaged parts of country. 	<ul style="list-style-type: none"> - DLI targets achieved with reductions of OOSC reported in targeted districts. 	<ul style="list-style-type: none"> - Unclear rigor of estimates of change for OOSC - National-level campaigns may have had success. 	<ul style="list-style-type: none"> - Relatively low spending priority for government - Support from UNICEF and NGOs also provided. 	<ul style="list-style-type: none"> - Unclear. May be increased access inequality as a result of COVID crisis.
New Scholarship Schemes	<ul style="list-style-type: none"> - Secondary enrolment for girls and poor - Consolidated scholarship schemes. 	<ul style="list-style-type: none"> - New scholarships targets achieved - GER targets likely to be met. 	<ul style="list-style-type: none"> - Unclear that any full evaluation of new scholarship schemes carried out. 	<ul style="list-style-type: none"> - New schemes created but no sign of consolidation - Unclear GER improving due to scholarships. 	<ul style="list-style-type: none"> - Unclear if new scholarship schemes can be sustained, or any evaluation of how they fit with existing schemes.
Disaster Risk Reduction (DRR), safe schools	<ul style="list-style-type: none"> - Disaster risk resilient school infrastructure. 	<ul style="list-style-type: none"> - Delays but school block retrofitting targets mostly met. 	<ul style="list-style-type: none"> - Small-scale and only likely to have helped a small minority of schools become safer. 	<ul style="list-style-type: none"> - Intervention backed by government investment and work of NGOs. 	<ul style="list-style-type: none"> - Significant issues with capital works linked to lack of engineers for sign-off.
Governance and fiduciary management	<ul style="list-style-type: none"> - Improved PFM - Performance-based grants to schools. 	<ul style="list-style-type: none"> - Early results achieved. From Year 3, DLIs undertook large revision due to federalism. 	<ul style="list-style-type: none"> - PFM system requires redesign with federalism - Lack of clarity on performance-based grants. 	<ul style="list-style-type: none"> - Initial targets became outdated - Revision process may have formed a new PFM plan. 	<ul style="list-style-type: none"> - System in flux, clear need for new PFM reform approach.
Reliability & transparency of EMIS	<ul style="list-style-type: none"> - Rollout and Improved accuracy of web-based EMIS. 	<ul style="list-style-type: none"> - All achieved to date, “95%+ accuracy” of EMIS on verification. 	<ul style="list-style-type: none"> - Unclear on quality of EMIS, look to be big issues with LG-level data. 	<ul style="list-style-type: none"> - Web-based EMIS put in place, includes TA support. 	<ul style="list-style-type: none"> - Challenges of web-based EMIS far from being resolved.

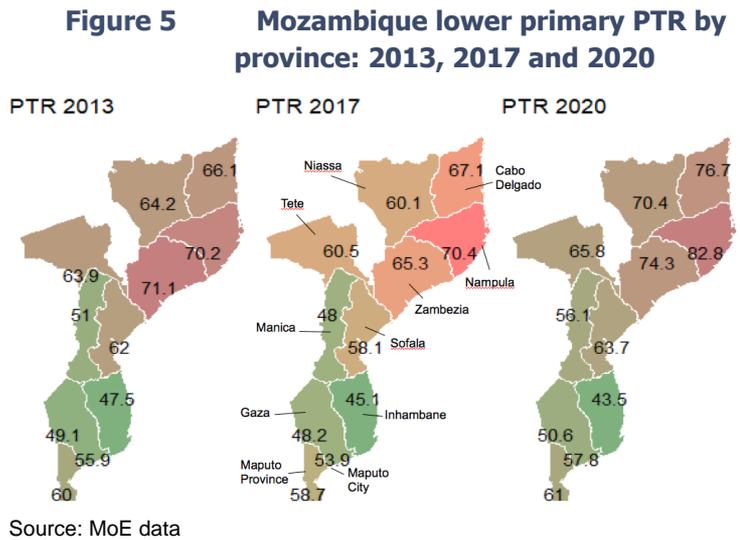
DLIs	Intended higher-level goals	DLI results	Achievement of higher-level goals	RBF / DLI contribution to results	Sustainability
Mozambique, PFM4R (2014–18)					
School councils meet criteria	- Shift in accountability at school, community engagement, more transparency	- Majority of school councils reported to meet criteria, DLI targets met.	- No evidence of a significant shift in local accountability. - National / system-wide change will take many years.	- Unclear. DLIs helped drive process changes, alongside coaches and facilitators.	- More qualitative analysis and understanding of local accountability dynamics needed.
ADE first tranche arrives on time	- Improved vertical accountability, and district PFM inc. reduced risk of corruption/ diversion	- ADE first tranche nearly 100% on time, targets met.	- Reports suggest timely funding to schools, but no evidence of improved learning environment / effective use of funds.	- DLIs helped drive process changes, alongside coaches and facilitators.	- First tranche still on time, more than one year after program.
Schools are supervised regularly, with follow up visits	- ADE well used to address needs of the school - Improved teacher / director attendance - Improved student retention and learning.	- Targets met, more regular and systemic.	- Reported improvement in teacher attendance - Little evidence of improved teaching and learning.	- Reported correlation between supervision and teacher attendance - Quality of supervision weak / variable.	- Systemic / regular supervision sustainable if funding continues - Qualitative change will require more time.
District budget classification		- Target met, SDEJTs become budget management units.	- Improvement in transparency of district spending.	- Some use of PBAs to reward provinces.	- Process changes sustainable, systems now in place.
Mozambique: GPE variable tranche DLIs (2015–17)					
School Director (SD) training + performance evaluated	- School Directors more effective, perform their roles, attend school regularly.	- Targets achieved, for teacher training institutes (IFPs), numbers of school directors to be trained.	- Institutional / political economy barriers still in place - Low impact on behavior of school directors.	- Targets helped push IFP training, but did not contribute to qualitative change.	- Role of IFPs established; systems for evaluation emerging. No evidence of sustainable change.
In-service teacher training	- Teachers in early grades more able to deliver curriculum, more systematic peer support in schools.	- Targets achieved, for IFPs, numbers of teachers with in-service training, linked to cascade model.	- Little evidence of better classroom teaching.	- Targets gave new training strategy momentum, but no qualitative change.	- Role for IFPs more established, but not yet other levels of cascade.
More equitable PTRs across districts	- Improved equity in distribution of qualified teachers; improved environment for learning.	- Reduced number of districts with PTR over 80 partially achieved; real achievement not recognized due to baseline change linked to district reorganization.	- Greater equity within some provinces but not between provinces.	- Correlation, and changes likely to have been incentivized by DLI / target.	- Not sustainable, PTRs have risen, budget constraints on teacher recruitment.
Mozambique: Germany (KfW) use of RBF in construction (2014–17)					
School construction	- Improved and faster pace of school construction.	- Initial targets for 2014 considered not achieved. Classrooms target for 2015/16 not achieved. Full payment made.	- The pace of classroom construction in the incentive period did not increase, falling in the 2014 to 2017 period.	- Considered not effective in addressing bottlenecks to faster classroom construction. Affected by resource constraints.	- Construction of classrooms reduced in recent years compared to RBF period, although picked up in 2018 and 2019.

Notes: DLI results are brief representations of actual DLI results – for more details see Dom et al. (2021); Patch et al. (2021); Holden & Chapagain (2021). Authors' assessment for all other columns.

Comparing approaches to reducing pupil teacher ratios (PTRs) across the three countries

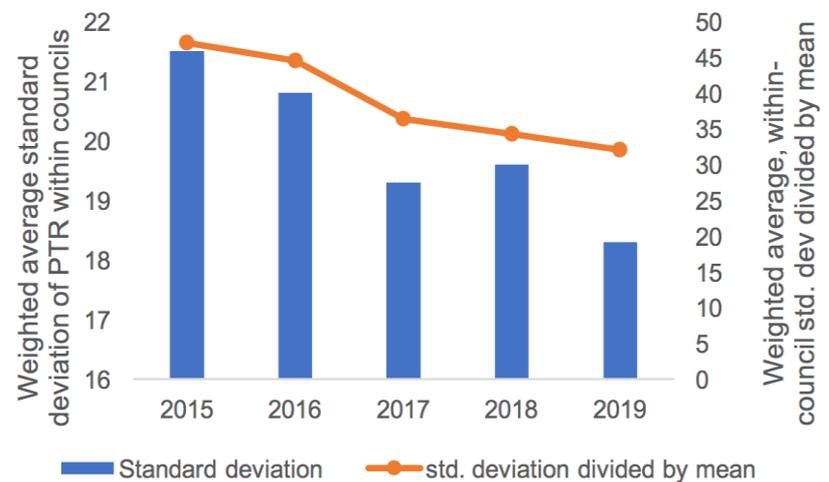
92. The Final Assessment Reports for the three countries have explored the country experiences with teacher deployment and the PTR as a focus of the RBF (see Box 7 in Section 3A).⁶¹ The results show some of the most notable successes of RBF programming as well as some of the most common challenges – in cascading incentives, ensuring equity, and in designing good metrics. The topic of PTRs was chosen as it was present in each of the three countries, and also provides some of the most important quantitative results for analysis.

93. Prior to the use of RBF in Mozambique, although average PTRs had been falling, they remained unequal across the country; for lower primary in 2013 as low as 47 in Inhambane, while above 70 for Nampula and Zambezia provinces. There was also inequality within provinces, and this became the DLI measure for the GPE variable tranche, from a baseline of 12 districts with a PTR for lower primary above 80 in 2014 to eight districts by 2016 and two districts by 2017. The DLI target was only partially reached, and this was largely due to district reorganization in 2013, which increased the number of districts with PTR above 80 from 12 to 16 districts in 2014, although the official GPE DLI baseline was not changed. The reduction in practice was then from 16 districts in 2014 to nine districts in 2017, short of target but a notable achievement. Quantitative analysis for our assessment (see Patch et al., 2021), found that within provinces there was a statistically significant effect of teachers being more equitably distributed towards districts with higher PTRs. This implies the DLI “worked,” and a message may have been passed down from the central to province level. The achievement was not fully recognized via the DLI due to the baseline issue. As Figure 5 shows, between-province inequality persisted in the period since RBF ended in 2017 and worsened by 2020, showing the issues of inequitable teacher deployment have not been sustainably addressed.



94. In Tanzania, a more complex measure of PTR was used, which included inequalities in PTRs within LGAs, as well as between LGAs, making two separate DLIs. The approach was based on the realization that teacher deployment was strongly imbalanced at both levels, and this had major implications in terms of unequal learning opportunities for students. Data analysis undertaken for this assessment show that teacher deployment within LGAs became more equitable by year five of the EP4R. This is shown in Figure 6 above, taken from an analysis by the

Figure 6 Tanzania: Measures of variation in PTRs within LGAs during EP4R period, 2015-19



⁶¹ See final assessment reports – Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021). Quantitative analysis is undertaken and included in annexes to those reports.

program's Technical Assistance Support (TAS) and displaying the 2015–19 comparison of the standard deviation and a coefficient of variation of PTR within LGAs. Our analysis found that for the 25 percent of schools where the PTR decreased between 2015 and 2019, this happened in proportionally more of the schools with a high PTR in 2015; and in most LGAs with high PTRs, the PTR improved most in the most disadvantaged schools. This occurred while the national average PTR was rapidly increasing (from 43 in 2014/15 to 57 by October 2019) due to the cap on teacher recruitment. Thus, LGAs performed well on the goal of improving teacher deployment, including LGAs where the total teacher count was inadequate. However, the metric, the "number of schools in each LGA with a PTR within an acceptable range", no longer measured the equity of change in the context of scarcity of teachers of the later years of the program. This led to losses of RBF rewards for LGAs. In addition, the structure of DLI payments did not consider differential challenges or costs of achieving targets for LGAs, including greater costs for rural areas. In spite of the difficulties, interviews at LGA level suggest there continues to be genuine attention to teacher deployment. The ability to pay for teacher transfers, as LGAs were instructed by the center to use part of the EP4R rewards for this, did make a difference. But there were concerns that LGAs' commitment might waver if nothing was changing, as they were "unfairly punished" (in the words of an education official).

95. In the case of Nepal, for years there has been variation in PTRs across the country – as low as 19 in Gandaki Province and as high as 45 in Province 2 for secondary level in 2019, with much higher variation if only permanent teachers are considered.⁶² In part, this is a reflection of wide variation in population density between the mountainous, hilly and Terai regions. Analysis by the World Bank notes that teacher hiring and allocation remains a contested area across levels of government, particularly since federalism and "this is especially true around deciding who will hire temporary and permanent teachers; and determining teacher allocation within LGs".⁶³ At the heart of the issue is the acute shortage of tenured teachers and the question of their equitable deployment in the school system. The SSDP DLI included initial process indicators covering deployment via a "teacher rationalization and deployment plan" in 2017.⁶⁴ The deployment plan set out district-by-district data on surplus/deficit teachers, based on quite complex criteria for the number of grades running at the school and the number of students. The process resulted in the identification of an overall deficit of over 74,500 teachers and a surplus of 29,000 teachers, suggesting an overall shortage of around 44,000 teachers in the system. The DLI wording for the Year 3 target was then "*Number of teachers to be redeployed* reduced by 25 percent of the baseline", and by "60 percent of the baseline" for the Year 5 target. Progress was slow, with the Cabinet only endorsing the redeployment plan in August 2019. The government have now shared directives to local level, stating that redeployment between LGs, districts and provinces is "under implementation".⁶⁵ However, no data is reported, and it is unclear how progress relates to the original deployment plan. Most importantly, PTR inequality persists, both overall in terms of total teacher numbers, and in relation to better qualified teachers.

96. The comparison of the three countries use of RBF to tackle teacher deployment shows the potential benefits in the clarity of a direct measure such as the PTR as used in Mozambique and Tanzania. While achieving changes in PTR is complicated and driven by the fiscal environment, incentivizing this indicator appeared to be successful according to quantitative analysis conducted for this assessment. In contrast, Nepal's approach of a static and centralized "deployment plan" has struggled to mobilize action, though this is linked to the issues of federalism as for so many other DLIs in the SSDP period. One development partner stakeholder in Nepal outlined: "teacher deployment is a major issue. There are no incentives currently, and it is a major political problem. Poor and remote areas are underserved, it is a knotty and first order problem that is not being tackled by the DLIs and RBF." It is possible that a more direct measure such as PTR may have been more successful in giving a clearer signal. That said, the story of Tanzania also suggests that a clear signal may become ineffective or

⁶² MoEST Nepal (2019)

⁶³ Sabarwal et al. (2019)

⁶⁴ MoEST Nepal (2017)

⁶⁵ MoEST Nepal (2020)

generate discouragement rather than acting as an incentive, if the agents in charge of acting on the signal do not have the power to remove the obstacles to achieving the signalled results.

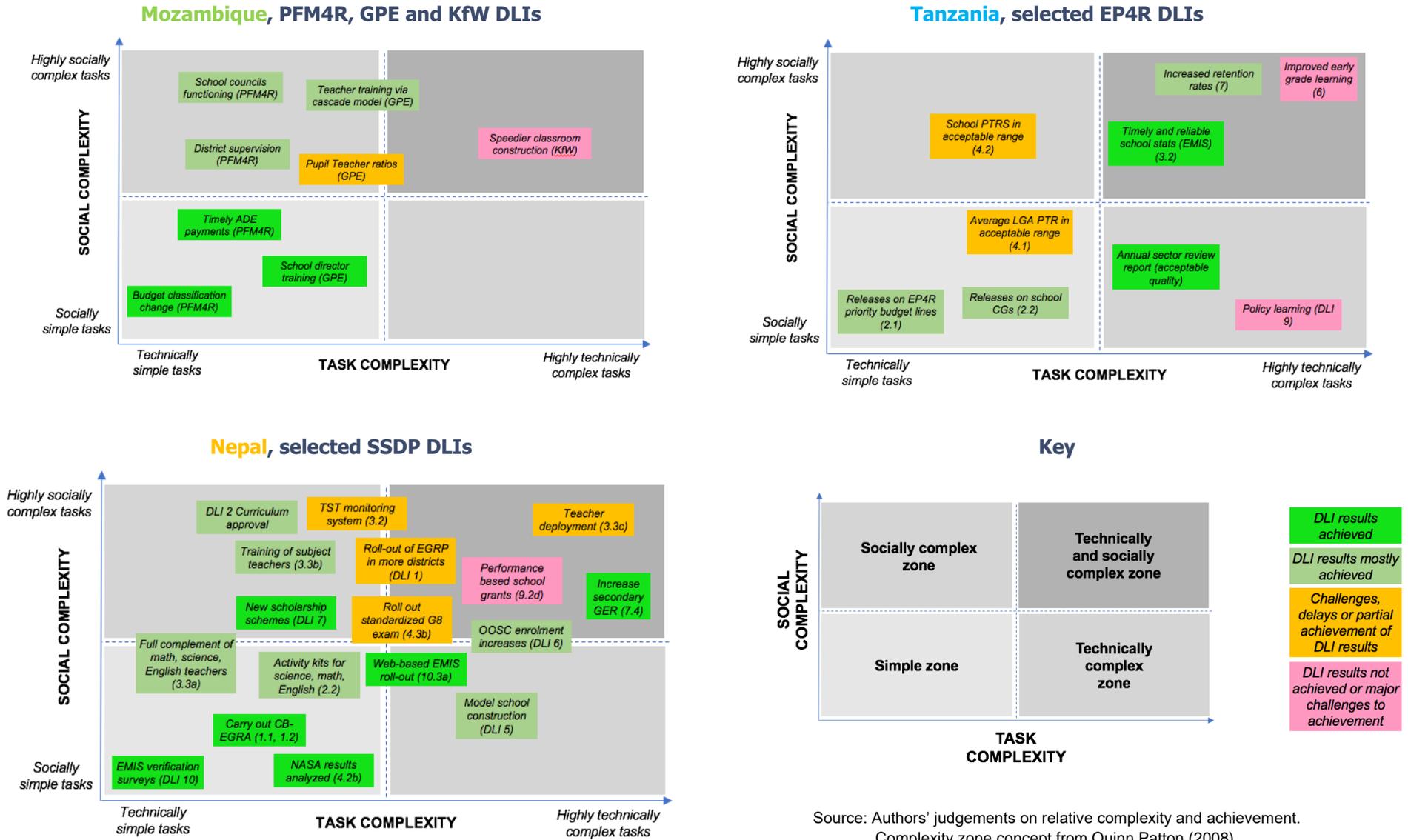
Analysis by different types of indicators and their complexity

97. The teacher deployment challenges show that where metrics are well designed, progress can be made even with a challenging fiscal and political economy environment – even if the design does not always take that environment into account. The analysis also points to the degree of complexity as one driver for whether results are achieved. This can be framed in terms of both the *social complexity*, the number of actors involved and the political complexity in their coordination; and *technical complexity*, the degree of difficulty in the tasks in achieving the results. Figure 7 shows for each of the three countries a framing of selected DLIs based on these concepts of complexity:

- In **Nepal**, facing a large number of process and output DLIs in a changing context can be extremely challenging. Federalism led to changes to the organization of the education system (see Chapter 2 and Section 3D), with LGs becoming the main tier for supporting schools in service delivery. As a result of the changes including among many other challenges, flux in the civil service and reporting responsibilities, the more involved the local level was in the delivery of a DLI, the more complex it has been to both report and implement. Central level process indicators were simpler although these were mainly early on in the SSDP. Most DLIs beyond the first year were process or outputs that required LGs to deliver, with significant social complexity particularly given federalism.
- In **Mozambique**, the school grants (ADE) and budget classification DLIs were relatively technically simple, although the achievement of ADE grant timeliness required coordination of the MoF, MoE, province, and districts, and therefore included social complexity, although there is a relatively top-down cascade between these levels (see Section 3D). This compared to socially complex DLIs within the PFM4R such as the functioning of school councils, and more technically complex indicators for the German (KfW) use of RBF to speed up classroom construction. Many of the DLIs in Mozambique included relatively simple metrics, for example, the GPE DLIs on number of school directors and numbers of teachers trained at IFPs. These were not very technically or socially complex partly as the indicators only captured aspects of the intended programming – for example, the teacher training DLI did not capture the more complex “cascade” by which those taught at IFPs would then train teachers at school level.
- In **Tanzania**, DLIs included more outcomes such as early grade learning results, which are inherently socially and technically complex, with a wide range of actions, and coordination between actors, required to achieve them, including both students and parents. *Financial DLIs* are in principle socially simple as they involve mainly the MoF to release funds, and the MoE and PO-RALG, to provide the information necessary to calculate the releases.⁶⁶ For *bringing average LGA-level PTRs to an acceptable level (DLI 4.1)*, transferring teachers across LGAs involves some technical and administrative complexity in that the teacher’s file must be moved from one to another “employer” (LGA), and politically this can be difficult. Within LGAs, *bringing school PTRs to an acceptable level (DLI 4.2)* requires a larger number of agents to be convinced of the importance of doing so, although the process itself should be technically simpler. The “*data*” DLIs are technically complex requiring data collection, validation and publication processes, while EMIS collection (DLI 3.1) is more socially complex with more actors involved, than for the production of sector review reports of acceptable quality (DLI 3.2). The *policy learning DLI 9* is technically complex to use well, and whilst it is “demand-driven” it is often a challenge for government to proactively identify where they might require expertise to assist.

⁶⁶ Technically, *releases on CG (DLI 2.2)* is more complex than *releasing on a small number of priority budget lines (DLI 2.1)* – it requires computing school enrolments, CG amounts and school bank accounts for tens of thousands of schools, making lists/LGA, MoF and Central Bank to correctly use the list etc.

Figure 7 Complexity zone and selected DLIs within the RBF programming in the three countries



Source: Authors' judgements on relative complexity and achievement. Complexity zone concept from Quinn Patton (2008).

Is there a pattern linking complexity to the results on DLIs?

98. The social and technical complexity framework provides an insight into how “stretch” can be conceptualized in practice, and how this may relate to achievement. The Figure 7 above uses color codes to represent the level of achievement of the DLIs. This assessment is done ex-post, and therefore cannot be said to be without bias. However, this analysis, even indicative, highlights interesting contrast between the three countries in relation to the strength of the link between complexity and results.

99. The analysis suggests that the most distinctive pattern on achievement was in Nepal. This is particularly the case for socially complex DLIs, as above, linking to the complexities of working with LGs following the move to federalism and the raft of new legislation in 2017–18 (see Box 2 in Section 2.3). The few outcome indicators within the SSDP in Nepal (improvements on early grade reading learning scores for DLI 1; the gross enrolment ratio, GER, for DLI 7; reduction in OOSC for DLI 6) are inherently complex to achieve, but mainly come late in the program (Year 4 or 5). During the period to Year 3 of the SSDP, the rollout of initiatives at the new LG level have been socially complex and difficult to deliver, including the implementation of a standardized Grade 8 examination (DLI 4.3b), rollout of the EGRP in new areas (DLI 1), implementation of TST monitoring in schools (DLI 3.2), and a new grants system including performance based grants in schools (9.2d). In contrast, where tasks were carried out largely by central level institutions such as the MoE and its supporting agencies, DLI performance has been better, for example carrying out classroom based (CB)-EGRA studies (DLI 1.1, 1.2), and the analysis of NASA study results (DLI 4.2b).

100. In Tanzania, in contrast, the evidence on the link between higher complexity and lower achievement on results is mixed. It is the case that results on learning DLIs, the most complex results in the RBF programming, were no longer achieved in 2017. But some of the DLIs requiring a lot of subnational actors and actions, which had a degree of social complexity, have not been less likely to be achieved, including DLIs on data and equitable teacher deployment. This may be in part because the program had a more sophisticated design to ensure subnational involvement and payments (see Box 10 in Section 3C). The challenges for the DLI on more equitable PTRs within LGAs (DLI 4.2) were most linked to the magnitude of the teacher shortfall (see above), and by the end of the study period most LGAs did achieve a more equitable deployment of the teachers they have. The socially simpler DLIs have not done so well or are deteriorating, even the socially and technically simple releases on EP4R priority budget lines (DLI 2.1). This suggests the need to look at a wider range of potential explanatory factors, notably whether the “correct” agents are targeted by the DLI as incentive, and the availability of sufficient resources (financial and others) for the agent(s) to achieve results, a factor over which agents in charge of an action may not have full control. Later in this document Section 3C looks at the financial flows from RBF, while Section 3D looks at some of the important aspects of cascade that may also drive the effectiveness of RBF. It is important to also recall here that, as noted earlier, in the case of Tanzania the underfunding of the FFBEP acted as a stretch-increasing factor in a way which is hard to capture in a static framework.

101. The complexity framework provides a potential viewpoint to why in Mozambique, performance was better in terms of the proportion of DLIs achieved. The bottlenecks and challenges targeted by DLIs were arguably less complex than in Tanzania or Nepal, and therefore can be considered to have less stretch, in terms of social or technical complexity. Another factor that Figure 7 points to, is that there were fewer DLIs and associated result targets in Mozambique than in Tanzania and Nepal. The PFM4R had a strongly supportive program structure with coaches and facilitators, that played a large role in helping to achieve DLIs, and with only four DLIs to achieve, good results perhaps should be expected. The GPE DLIs had fewer supportive structures in place, although were simpler with the exception of the PTR. The German (KfW) use of RBF in construction also came with accompanying technical assistance, but for many reasons did not seem sufficient to deal with the many complexities in that area. The Mozambican DLIs were generally simple quantitative indicators, which were (perhaps necessarily) imperfect measures of the complex changes and goals that were intended. But the fact they did not fully measure system changes is important to consider when assessing the achievement of the DLIs.

102. In conclusion, whilst the social and technical complexity framework outlined above appears to have some explanatory power, the analysis also highlights the difficulty of defining all the potentially influential dimensions of complexity, particularly the challenge of adequately “factoring in” complex political economy issues. The cases of Tanzania and Nepal also highlight that the social/technical complexity of any DLI is not static, in turn suggesting the potential relevance of regularly reassessing stretch in the life of an RBF program, especially in situations of large contextual changes like in these countries.

Sustainability of changes

103. Even though the programming period for both Nepal and Tanzania is still under way, with the RBF now going back some years it is possible to look at the sustainability of changes in these two countries as well as for Mozambique. For the earlier period of RBF in Nepal (2013–15), the SSRP ILIs (bank payment for teacher salaries & students’ scholarships, timeliness of textbook provision) were not taken forward into the SSDP period. There was also an apparent absence of evaluation of this earlier RBF experience, which may have considered for example the processes of verification (see Section 3E), in which payments were not made in the first year due to the rejection of one verification report. In addition, the SSRP period showed that payments could be lost, and when scaled up to the SSDP stage with much more finance “at risk” this may have suggested there were challenges to come. We found no documentary evidence of consideration of these risks.

104. A number of process and output changes represented simple descriptions of change that captured only an aspect of the reforms and/or goals targeted. For example, in Mozambique’s PFM4R, the DLI on supervision was a measure of schools being visited once in the first half of the school year, and again in the second half of the school year; the school councils DLI was a “tick box” measure of whether elections were held, meeting minutes were kept, and the amount of ADE received was publicly displayed. In these cases, the DLIs can be viewed as “stepping stones” to system changes but without being guarantees of quality of the supervision system in place or accountability mechanism represented by school councils. District supervision has continued to operate beyond the PFM4R period, and with help from other development partners, has seen additional investment (for example in the use of tablets and sending supervision data to province and central MoE levels). The PFM4R DLI on the quality of supervision was therefore a necessary but not sufficient condition for a higher quality supervision system with its expected benefits in terms of improved school performance. A similar comment can be made with regard to the SQA system in Tanzania, which appears to be fairly strongly owned so could continue to be implemented even without the incentive of the SQA rollout DLI, but with no guarantee that this would lead to sustainable “higher level” change – and in addition, its financial sustainability may also be an issue.

105. The early learning DLIs in Tanzania tell a different story, in which sustainability is far from clear in more than one way. The same issue arises with regard to financial sustainability of this additional learning assessment system. Moreover, it is clear that the national exam agency has not (yet?) bought into the skill-based learning assessment approach incentivized through the DLIs and continues to favor a curriculum testing approach. More fundamentally, in a society which continues to be dominated by attention to end-of-cycle examination results, it is not clear that early grade learning would continue to be a top-most focus of education officials’ attention if the RBF rewards were no longer present, and particularly, if early learning targets continue to be hard-to-reach or mixed – which is likely to be the case if basic education underfunding is not reversed.

106. While each program has interacted with teachers, they have all done so in a relatively narrow way in comparison with the scope and magnitude of issues affecting teacher management and effectiveness in the three countries. For example, none have tackled the pre-service teacher training system. In-service teacher training has been included in the Mozambique GPE programming, although this was still only capturing a part of the change needed, for example, in-service training at the top of a cascade (training the trainers model), but not delivery of the peer training at school level. In Tanzania, the DLI on in-service teacher training, focusing on training teachers on the new early grade curriculum, was discontinued at redesign. No further attention was given to this issue even though the need for in-service teacher training, in the FFBEF-era, was arguably even

starker with the emergence of a high number of very large classes, and there was no government-led in-service teacher training system in place.

107. As seen above, each country sought to more equitably deploy teachers, but whilst deployment faces major challenges for getting teachers to go to more remote and rural areas, there were no direct interventions linked to RBF to address these challenges. More broadly, there was limited focus on the classroom or teacher motivation. The SDI studies of 2014 in Mozambique and Tanzania showed issues with teacher absenteeism, but outside of these studies, which were not directly part of RBF programming, there was no systematic effort to measure whether it improved over time.⁶⁷ In Tanzania, the financial DLI succeeded in securing better releases of budget allocations serving to clear teacher salary arrears; and a new teacher deployment strategy was developed (foundational DLI at redesign) and initial steps taken to implement it, although arguably, the strategy was obsolete as soon as approved considering the magnitude of the teacher shortfall.

108. In Nepal, the SSDP program is still operational. There have been challenges in implementation and, for that reason, it may be premature to assess sustainability for a number of DLI results. The flux in the system due to federalism affects nearly every DLI result area, with the country's long-awaited "Federal Education Act" expected to help clear up responsibilities and functions of government tiers. This includes, for example, teacher management and deployment, one of many areas that is somewhat contested and complex due to the many types of teacher in Nepal and questions of how much decision-making power the new tier of local governments should have. As has been set out, this limits the effectiveness of deployment and addressing the challenge of inequitable PTRs across the country.

109. In all three countries, and perhaps most acutely in Tanzania, the extent to which government is likely to adequately fund basic education is the factor most influential in both a) the sustainability of the results achieved in relation to a number of DLIs (for instance, progress with regard to some of the dimensions of a more equitable teacher deployment), and b) the possibility of making further progress, or simply maintaining the gains earlier achieved and/or addressing emerging results with deteriorating trends (for example, the early learning levels and retention rates in Tanzania or growing inequity in provincial PTRs in Mozambique). The financing environment is therefore critical to understand the context of RBF, and this is discussed in more detail in the next section.

⁶⁷ There were non-systematic assessments. In Tanzania, OPM (2019a and 2019b) assessments as part of the continuous EQUIP-T M&E – associated to the EQUIP-T project focus on in-service teacher training in the areas the project was active. In Mozambique a follow-up SDI was undertaken in 2018, although with methodological challenges in the comparison to 2014. [World Bank (2015a) and Bassi et al. (2019)].

C. Financing and payments

110. In each of the three countries, the choice to pursue RBF programming was predominantly donor-driven.⁶⁸ As a financing modality, RBF generated “hard” conditionality for donor disbursements to government. The nature of the RBF incentive, if it is credible, is that there is a threat of non-payment. Depending on performance on a pre-agreed contractual set of targets (the “DLI framework” combined with bilateral “financing agreements”) financial disbursements may be lower than originally allocated, and at the outset the exact amounts are, by definition, uncertain. The mechanics of RBF with reporting and verification, also mean that payments should only be made after results are credibly verified.⁶⁹ RBF flows also interact with finance flows through government systems, from central to local level, and to schools; these may vary with regard to how well they align with government PFM systems.

111. This section discusses RBF funding in terms of volume and uses, in the context of the sector budget and total sector aid, in each country; the financing risks associated to the use of RBF; the financial management set-up of the different RBF programs; and whether and how RBF financial incentives flowed between government levels. The programs studied differ in all these respects. The Tanzania RBF is the largest in absolute terms, but the Nepal RBF represented a higher proportion of the total sector aid, making it riskier for government, especially in light of the rollout of federalism. The Nepal SSDP and Mozambique GPE and KfW programs used pre-existing aid pooling systems largely aligned with government PFM and fund flow systems, with safeguards. No aid pooling arrangements pre-existed the RBF programming in Tanzania, and the program itself became a large aid pooling device, though less well-aligned with government systems than in the other two countries. Only the Tanzania RBF featured a successfully implemented system of downward cascading of RBF financial incentives.

RBF and education sector financing

112. The RBF period covered in this assessment coincided with a decline or stagnation in real-terms education spending by government in all three countries. In Nepal and Tanzania there was also a decline in the share of spending to the sector allocated from the government budget, and in Tanzania, with the large FFBEP enrolment surge, an ensuing decrease of per pre-primary and primary pupil expenditure by 25 percent over the 2013–19 program period. The Tanzania RBF program included financial DLIs with one aimed to secure government releases on a few “priority budget lines” in the sector; as designed this did not engage with the issue of the underfunding of basic education. In Nepal, the SSDP included a “Global Covenant” alongside the main programming, a form of conditionality that non-RBF JFPs to the program also used. This includes a condition for each year’s reporting that budget allocation consistently increased in real terms compared to previous year, and while the year-on-year increases in real expenditure were met, they were only just, and in Year 2 this relied on a measure of equalization grant top-up, that is, allocation which had been at the discretion of provinces to give rather than central government. As a share of the national budget, education fell to under 11 percent in 2019/20, suggesting that the covenant is not succeeding if its goal is to improve the share of education spending allocated by central government.⁷⁰

113. Table 10 below summarizes a number of features related to RBF funding volumes and uses in the country contexts. These are discussed further below.

⁶⁸ In Tanzania, RBF funders led by the World Bank, spent considerable efforts in ensuring good understanding and adherence with the “logic” of RBF as a modality, including at ministerial level, and the government BRN process provided a “home” for RBF.

⁶⁹ While this is the main definition of RBF, advances using “zero-DLIs” (already achieved or very easy to achieve targets) may be used. Flexibility can be applied to retrospectively change targets and ensure disbursement (see Section 3F). Additional processes to report on and verify results and agree on the outcome of this process, can also have the effect of delaying financial flows. “Rollover” provisions may be included in the design, with planned disbursements for year X (based on results from year X minus 1), potentially made in year X+1, year X+2, or beyond, depending on when results are achieved.

⁷⁰ There were two other Global Covenant conditions: one was to submit the audited SSDP accounts within nine months of the end of the fiscal year, and the other was for the trimester Financial Management Report to be submitted within 45 days of trimester end date. Both conditions were met for all the years of the SSDP we assessed.

Table 10 RBF in sector financing across the three countries

	Mozambique	Nepal	Tanzania
Government sector financing	Relatively high share of total budget & GDP but stagnated in real terms 2013–18 (hence declining in real terms per student), and tight constraints on teacher recruitment, and investment.	Long term decline in education as share of GDP and budget (14% in 2017/18, 11% by 2019/20), stagnating real terms amount. Private spend on education higher than in other countries given size of private sector.	Declined in 2013–19 for basic education, overall and per student (~25% real-terms reduction for pre- and primary students). Strict cap on teacher recruitment drives this decline.
Sector aid dependency	High. Pooled funding FASE is main source of non-salary spending, but strongly fluctuated year-on-year during 2014 to 2018 (10-30% of total government + aid sector funding).	JFPs pooled funding (included RBF) for fully costed plan represented average 20% of SSRP total, falling to 10% of SSDP total (both including salaries); also smaller annual average aid volume for SSDP.	Fairly high for non-salary spending, though possibly less than in Mozambique. Sector aid rose over 2013–2019 (largely RBF, including large post-secondary skills project).
Other major non-RBF instruments	Pooled funding under FASE (set up in 2003; 10 donors in 2014–18), included GPE VT and KfW RBF. Some important bilateral projects (including USAID, GiZ, UNICEF) in selected provinces.	Pooled funding of both SSRP and SSDP through Joint Financing Agreement, including RBF & non-RBF funding. Norway, Japan & UNICEF not using RBF.	No pre-existing basket/ pooled funding (unlike in health sector). Several large single-donor projects, mostly with specific geographical focus.
RBF in total aid	12% of potential & actual donor disbursements.	Minor in SSRP phase (10% total aid funding in two RBF years, 3% for entire SSRP period). Up to 81% RBF in the SSDP JFA; much riskier.	Single largest sector instrument, 36% sector aid disbursements in 2017. Other sectors and post-secondary had very substantial RBFs.
Total potential available RBF	USD 58 million Of which for the three uses of RBF: GPE VT: USD 17 million PFM4R: USD 24 million KfW: USD 17 million.	SSRP (until 2015): USD 25 million, one funder (World Bank). SSDP (from 2016): with USAID joining, USD 399 million, six funders.	Start size (2013): USD 257 million, three funders. End size (2020): USD 515 million, five funders.
Other basic education aid by RBF donors	WB non-RBF ESSP part (through FASE). GPE fixed tranche (70% total GPE), through ESSP thus FASE. KfW broader contribution to FASE and bilateral funded TA (for construction).	All WB & AfDB funding use RBF. EU RBF 60% total EU funding, Finland 41%, USAID joined, GPE added funding to initially small pilot, in SSDP Year 3.	WB, Sida: occasional small-scale TA. DFID: TAS (within-program); large EQUIP-T project (poorest Regions). GPE: country-wide Sida-managed projects (LANES1 & 2). RBF 54% of LANES 2. KOICA: parallel small project.
Limitations on RBF funds use	To be spent on education. FASE funds spent on activities in MoE's annual plan including at subnational level; excludes salaries; includes essential funding such as school grant. Further virtual earmarking on eligible expenditures for KfW and GPE VT, but not hard conditionality.	To be spent on education but "Chinese wall" between fund disbursed and uses by government. Only USAID (very small RBF) is virtually earmarking ex ante but without monitoring.	Ring-fencing to sector, with disbursements on MoE-held account, jointly decided by MoE/ PO-RALG. To be spent on non-salary education costs.
Actual use of RBF funds	Not clear due to limitations in FASE reporting (GPE, KfW). PFM4R funds not disbursed through FASE, so not clear if even fully used in sector.	RBF/JFP funds are commingled, RBF funds any of the agreed SSDP expenditures.	Part of RBF rewards spent on post-basic education. Significant portion spent on infrastructure (including basic and post-basic).

Additionality of RBF funds

114. Claims to aggregate funding additionality are more difficult to make for RBF than under sector budget support (SBS). For SBS the amount of spending, rather than its effects, is typically the main indicator, or at least a very important one. Perceptions of additionality were positive in Tanzania, helped by an increasing and regular financing profile. They were more negative in Mozambique, as RBF funding ended up compensating increased volatility in non-RBF FASE aid funding following the hidden loans scandal and the suspension of GBS. Additionality is less clear in Nepal; total annual aid funding for the SSDP was considerably smaller than in the SSRP period, performance to date has led to delays in fund flow, however donors have provided funds to post-earthquake reconstruction of schools, with spending managed by the National Reconstruction Agency, which is not captured in the education sector plan.

115. In Tanzania, government officials found RBF highly valuable: the funding was seen as additional to the government's core budget, it was flexibly usable at the MoE/PO-RALG level, and disbursed in a relatively timely and reliable fashion (see below). The crunch of basic education government funding, which escalated over the program period especially after 2015, was unrelated to the presence of the RBF aid flow – although it represented a substantial proportion of sector aid financing, it was too “invisible” for the MoF to have displaced government funds. Rather, the main factor driving the government funding decline was the teacher recruitment cap. That said, our analysis shows that a significant portion of RBF funds was actually used for post-basic education. This, whilst in line with the RBF program design to not earmark funds beyond the sector, does raise the question as to why development partners never questioned this development, considering the program’s intended focus on basic education.

116. In Mozambique the flow of donor funds to the sector was volatile during the RBF period, with a particularly low year for FASE funding contributions in 2015, the year the hidden loans scandal led to GBS suspension, and also the first effective payment year for KfW, GPE and the PFM4R. The MoE’s annual report for that year clearly indicates that in such a context, the expected additional RBF funding and incentive was lost: “at the moment, these funds contribute to mitigate the negative impact of the reduction of FASE funds ... the program “lost” its original philosophy of encouraging/rewarding good performance at various levels (even the school)”.⁷¹ There was a key difference in GPE and KfW RBF funds which went through FASE (and are visible in FASE accounts), and the PFM4R allocated funds, in the latter case as set out in Box 8, it is not clear they were used in full for education.

Box 8 **Mozambique – The unclear additionality of PFM4R financing**

As a multi-sector program based in the MoF (with a PFM focus on health and education), PFM4R funds were not channeled through the FASE, but appeared as “PFM4R” project in the MoE’s section on investment expenditure in the budget. At design stage, the program was envisaged as a co-financing operation with a USD 80 million contribution from the government, out of total anticipated expenditure (across health and education sectors) of USD 131 million.

In practice, the actual expenditure of the program was not as anticipated. While all disbursements were made by the World Bank by early 2018, beyond a reported USD 16 million expenditure on a “capacity window”, performance-based allocations (PBAs), program coordination and verification, it has been hard to be clear on what the PFM4R money was spent, and indeed whether the sector allocation was all spent on education (even after excluding the proportion of disbursements for the health sector).¹ It is also not clear that there was any specific government contribution to funding activities of the program, and reporting tended just to include general government education expenditure as its contribution (i.e. all of its spending), so in practice it is not clear whether any additional finance was spent and if so, on what.

¹ World Bank (2019b).

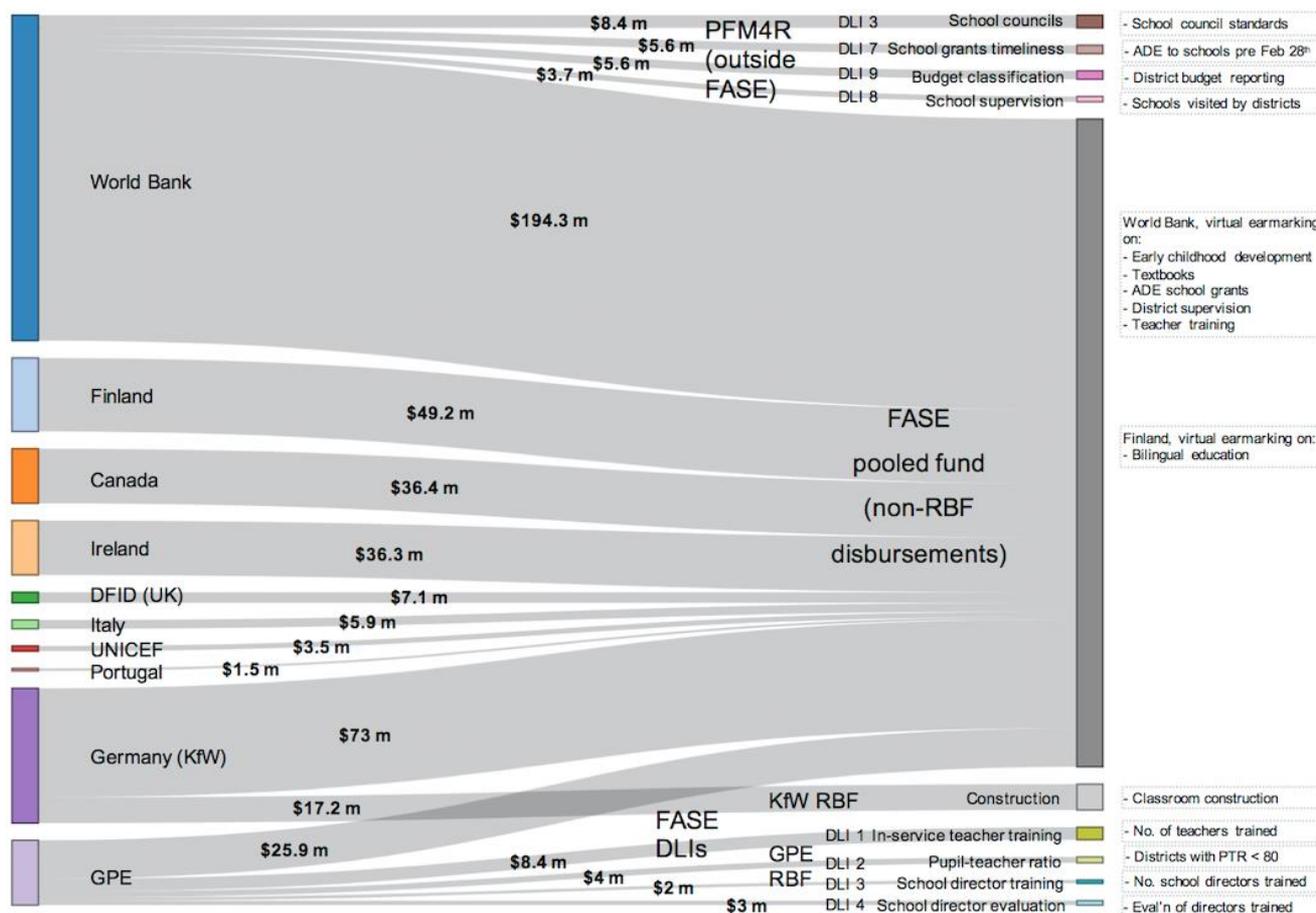
⁷¹ MINEDH (2016).

117. Thus, in Mozambique the three RBF programs were separately programmed and the PFM4R, as a sector PFM program using RBF also covering health, did not use the education sector aid pooling mechanism. In Nepal, the DLI framework, representing 81 percent of the total aid funding for the SSDP, was common but in reality, fairly distinct for each funder and, as might be expected, bilateral agreements between donor and government would take precedence over the Joint Financing Agreement (JFA). In Tanzania, the RBF program acted as its own pooling mechanism, although donor fragmentation in terms of funding different mixes of DLIs in the common framework increased in the course of implementation, and future programming might result in even stronger fragmentation.

118. Below, we set out the financial contributions of each RBF funder to the range of DLIs of the RBF programs studied. In Mozambique the figure also includes non-RBF funding to education through FASE, which indicates the relatively small size of RBF funding in the total sector aid funding:

- Mozambique** - Table 10 outlines the contributions through the PFM4R, which are separate from the pooled fund (FASE). It includes the FASE funding for education, separating out the larger non-RBF financing coming from the World Bank, Finland, Canada, Ireland and more, with World Bank and Finland using some virtual earmarking; and the RBF funding from GPE and KfW, also through FASE. In total, the three RBF funders provided a total of USD 58 million potential disbursements across the three programs, with individual DLI values, ranging from USD 2 million for the DLI on school director training in the GPE RBF program, to USD 17.2 million for classroom construction which was split in two rounds of funding by Germany (KfW).

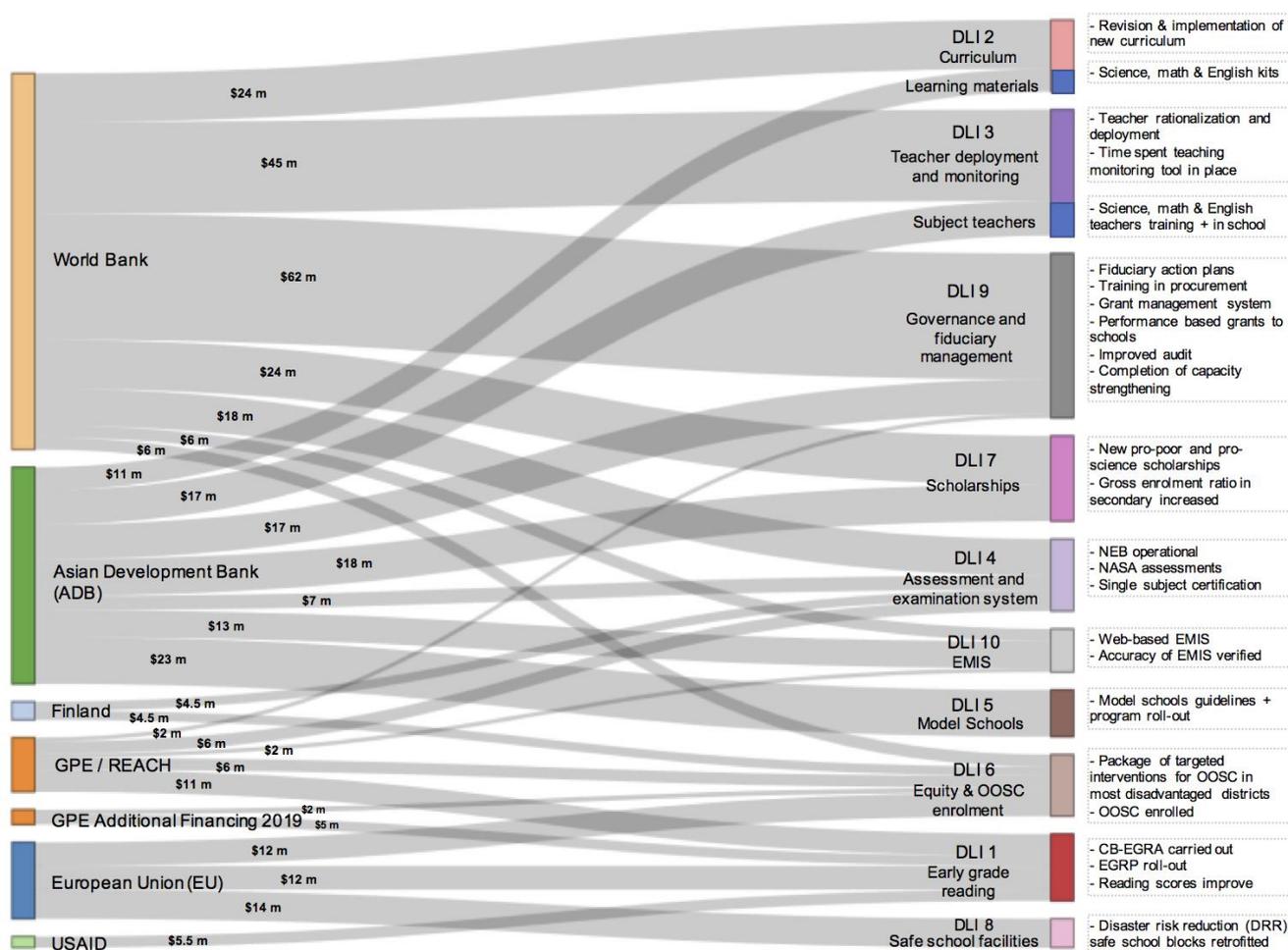
Figure 8 Mozambique education, 2014–18, RBF potential disbursements and non-RBF actuals



Source: FASE Financial Statements including EY (2018), KPMG (2019), World Bank (2014, 2015b, 2019b), KfW (2017), Orlowski (2016).

- Nepal** – The SSDP period use of RBF (since 2016) has seen a more complicated structure of DLIs. Figure 9 outlines the RBF contribution of each joint finance partner (JFP) linking finance to DLIs for the period. The diagram shows transfers for each of the ten DLI areas, which include 84 individual DLI results across the five years (individual targets often covering very different processes, outputs and outcomes within the same DLI area). The largest DLI areas focus on governance and fiduciary management (DLI 9) and teachers (DLI 3). Total external finance for the SSDP is made up of USD 399 million in potential DLI disbursements, and USD 91 million of non-RBF funding.⁷² The two largest DLI partners are the World Bank (46 percent of the total RBF value) and the ADB (31 percent), both providing concessional loans, while the largest grant provider is the EU (9 percent). The average DLI result value varies by donor, ranging from USD 1.1 million per result for Finland, to USD 7.1 million per result for the World Bank. The ADB has more individual result targets (38) than the World Bank (26), with a lower average value of USD 3.3 million.

Figure 9 Nepal SSDP value by donor and DLI, 2016–21 (allocations not actuals)



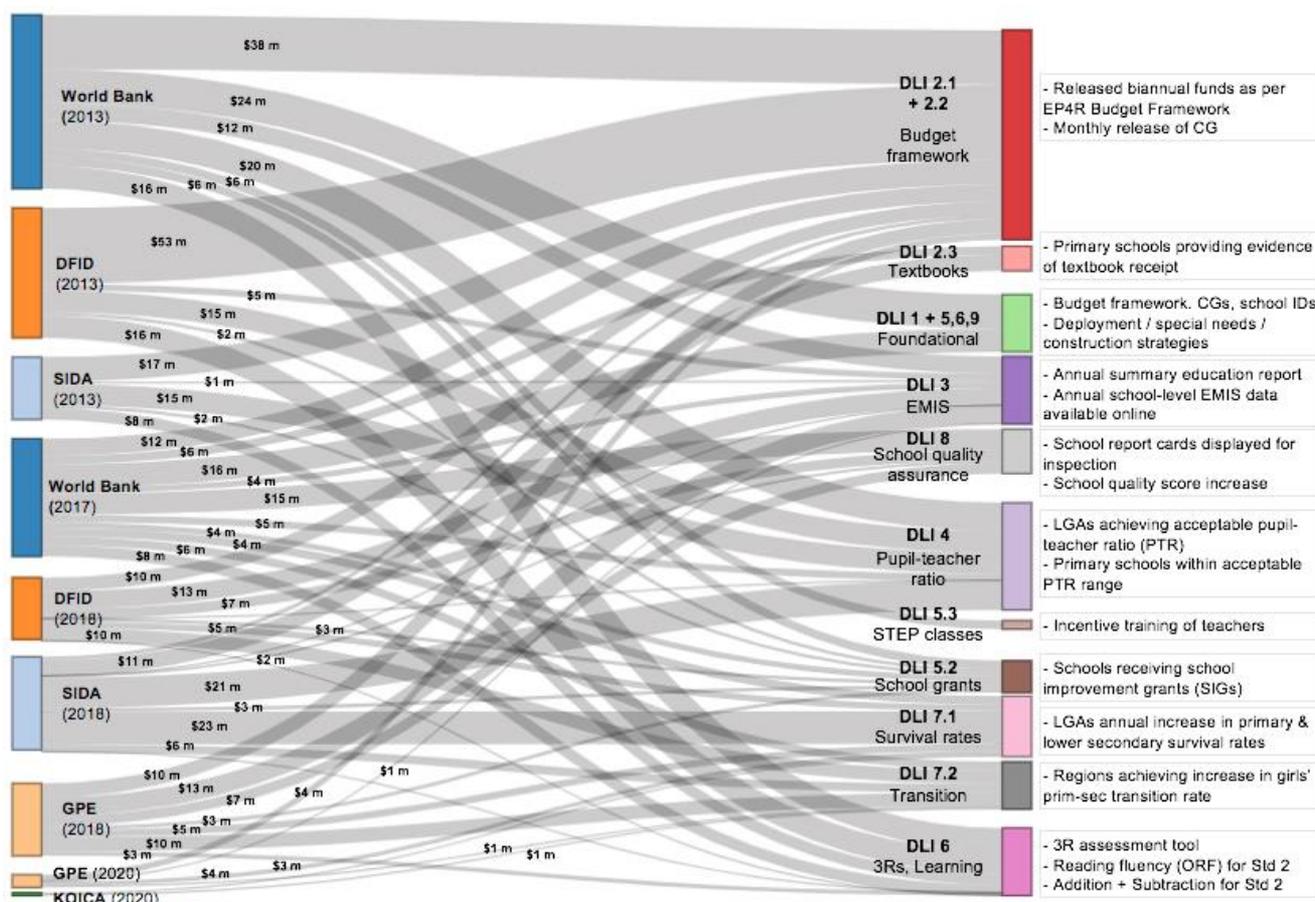
Source: SSDP Joint Protocol updated to October 2019. Exchange rate used of €1 = USD 1.12 for currency conversions.

- Tanzania** – Figure 10 below illustrates RBF funders’ allocations across the DLIs of the common DLI framework, as of early 2020. The program size by then (USD 515 million) was almost double the initial size, and five donors contributed, against three initially. The World Bank is the largest donor (almost 40 percent of the total), followed by DFID (27 percent), Sida (21 percent) and GPE (12

⁷² The SSDP also involves three “non-DLI JFPs” not shown in the figure, which are Norway (USD 36.6 million over the five years of the program), UNICEF (USD 2.5 million) and JICA (USD 13.6 million). The EU and Finland fixed tranches are worth USD 25.3 million and USD 12.8 million respectively, these are also not tied to DLIs.

percent), with KOICA experimenting with just USD 2 million. Like in Nepal, DLI individual values also significantly differed. Of note is the strikingly large proportion (a third) of the total RBF funding potentially available allocated to the “financial DLIs” (DLIs 2), in spite of the rather weak design highlighted above. Another notable feature is the fairly large proportion of the funding (9 percent) allocated to the highly complex “early learning” DLIs 6.

Figure 10 Tanzania RBF 2014/15-2019/20, EP4R by donor (allocations, not actuals)



Source: Up-to-date as of June 2020. MoEST & PO-RALG (2020) provides implied exchange rates, authors' calculations

RBF disbursements

119. As noted above, there are various reasons for disbursements to be lower than initially planned in RBF – including underperformance in achieving the DLI results, and also delays in reporting, verification, and donor decision-making or disbursement processing. They transmit differently to payment issues and have different risk financing effects. In relation to underperformance in achieving results, one can distinguish:

- *Financial constraints* – for instance, the continuous crunch on basic education funding in the three countries, which may affect all DLIs on a program, as achieving results depends on activities that must be financed and RBF payments follow.
- *Governance shocks* – such as for instance the Tanzania presidential change, and the Nepal constitutional reconfiguration, which may affect performance against different DLIs asymmetrically.
- *Technical constraints* – where DLI targets are not achieved, either due to its complexity (see Section 3B), a lack of response of relevant agents (see Section 3D), or challenges in measurement and verification (see Section 3E). Note, if there is stretch there will always be at least some non-achievement expected at the outset of programming.

120. As seen with the examples, there were cases of each of these situations across the three countries, hence reasons why DLI targets might not be fully achieved, discussed in Section 3B. Table 11 summarizes the financial implications that ensued from each country’s performance against planned DLI results, also indicating whether there were pre-agreed ways of using unspent RBF funds owing to underperformance on results, and how this was done in practice, as well as how donors managed the disbursement decision-making process. As seen above, RBF as implemented did not add much to sector financing transparency.

Table 11 RBF disbursement and financial risk in the three countries

	Mozambique	Nepal	Tanzania
Disbursement management	RBF funds through Treasury (direct for PFM4R; FASE set-up for GPE & KfW). RBF programs separately managed with regard to disbursement decisions.	Some DLIs were shared, but all JFPs operated under one Joint Financing Agreement (JFA) including non-RBF and RBF JFPs. Similar mechanism to used in previous SSRP period.	US dollar denominated virtual pooled funding first. Post 2017, multicurrency quasi-pool. Mostly joint disbursement design and joint decisions, although different mixes of incentivized results.
‘Rules’ for unspent RBF funding	PFM4R and GPE VT - carried over for unachieved (or partially achieved) results – said to keep staff motivated.	Rollover of funds possible for most JFPs under SSDP, less so for smaller bilateral donors than World Bank and ADB.	Lost except where partners “ex post” agree to reallocate on future DLI results.
Actual unspent RBF funding	Funds lost for GPE VT PTR (USD 3 million or 18% of the total). KfW first tranche all lost (USD 6 million or 33% of the total) for fairly unclear reasons (see Section 3B). Total loss 16% across the three RBF programs.	SSRP WB use of ILIs (2013–15): undisbursed Y1 funding carried over if target met in Y2. USD 20 million out of USD 25 million disbursed in total. SSDP: Rollover for several funders. By Y3, shortfall estimated at USD 36 million (15% of total). Restructure and rollover seeking to address shortfall.	Annual disbursements ranged from 69% to 85% of annual potentially available RBF funding; overall, by Y5 80% of the first five years of potentially available RBF was disbursed. Several restructures may lead to high final spent.
Financing risk (predictability and volatility)	Raised as issue given high aid dependency and FASE funding critical expenditure (including school grant), but low as RBF small. RBF “losses” also occurred in context of larger ongoing annual underspends from FASE, hence not “material”.	Risk significant, given size of RBF in total aid funding. Started to materialize from Y2 with DLI targets moving up the results chain and federalism (MoE lost control of delivery/link to key subnational agents). Risk prospect rising with federalism rollout unfolding.	Not problematic at national level, also not fully tested through additional funding and restructuring. More problematic at LGA level; annual fluctuations for approx. 40% of the LGAs; declining payments over the period for another 15%.

RBF financing risks

121. A key concern with the use of RBF is in the potential lack of predictability and volatility in funding flows to the government, with negative effects on the implementation of planned programs and reforms.⁷³ Our analysis suggests that in all three countries, financing risks associated with RBF moderately materialized, although have been mitigated by rollover, flexibility or redesign thus far. Unspent money was sometimes made available at other junctures – in Nepal and Mozambique through explicit rollover provisions as seen in Table 11 above, and in Tanzania through donor program restructuring. In Tanzania, the program redesign in 2017 cushioned volatility, while flexibility in changing DLI targets is in the process of doing the same in Nepal. In Mozambique, volatility of flows to the FASE sector pooled fund have been high in recent years, and the additional funds from the PFM4R

⁷³ Financing risks are one in a range of types of risks associated to the use of RBF. These other risks are discussed in Section 3G below. This includes an assessment of the extent to which attention was paid to RBF-specific risks at design stage and throughout implementation.

may have acted countercyclically to that. That said, higher future unpredictability and volatility is a prospect particularly in Nepal, given the rollout of federalism (see further detail in Sections 3B and 3D).

- In **Mozambique**, the financial risks associated with RBF were discussed in terms of a trade-off: potential accountability strengthening vs risk to predictability, less flexibility, added complexity. The risk was seen as undesirable, given the dependence of the MoE on the FASE for most non-salary spending required to implement the SDP annual activity plans. However, RBF was fairly small-scale. Moreover, by coincidence, the PFM4R RBF started as donor funding to FASE was being reduced. The RBF “losses” were also not very large and took place in a context of much larger underspend from FASE generally. While KfW did not pay the first RBF tranche due to non-achievement, they paid the second even though it was not clear that targets were achieved – at the risk of undermining the credibility of the instrument but with the intention to mitigate the effect of a funding shortfall on the broader program. As a result, in the 2014–2018 period, RBF was perceived as having mitigated the worst effects of the strong variations in total FASE funding (but therefore lost its incentive character, as noted earlier). That said, as aid dependency remains high, any increase in total RBF volume and/or in terms of proportion of FASE funding will increase the risks commensurately.

- In **Nepal**, RBF represented a significant risk given its size in total sector aid (81 percent of the JFP funding, representing 10 percent of the total sector funding including salaries). By Year 4, however, the shortfall (USD 39 million) was still fairly small compared to estimates of total RBF potentially available over the four years (around 15 percent). However, with a growing number of DLI results relying on action by the newly established LGs, the prospect of this shortfall rising over time was increasing. While revision and flexibility on DLIs were being used to mitigate financing risks (see Section 3F), it remains to be seen whether the agreed revisions will have taken sufficient account of the radical change that the federal system entails.

- In **Tanzania**, the risk was fairly significant, considering the EP4R size in aid funding. The discretionary nature of RBF funds, which unlike for projects, could be used as the MoE/PO-RALG saw fit, made them all the more valuable. That said, as RBF funds have been used for important but non-recurring items (construction work), this may have lessened the risk associated with volatile fund flows. The financing risks were not fully tested thus far. EP4R disbursed fairly high and annually increasing financing. At national level, the redesign with additional funding for additional DLIs and more additional funding from new RBF funders compensated the gradual decrease in performance for some DLIs (see Section 3B). There were a few instances of significant delays in disbursement (one from donors after the first review, leading to minor exchange rate losses and one from the MoF, that led to some strengthening of accountability for RBF funds), but overall, by 2019, RBF was still seen as reliable funding. Donors including the World Bank restructured their funding several times, apparently keen to make as much funding potentially available as possible. This may have slightly weakened the credibility of RBF as a mechanism, although this did not appear to be the case for individual DLIs. In contrast with this national-level picture, year-to-year volatility at the LGA level was high, with RBF funding for many LGAs varying quite strongly over the years – which, as far as we know, was not discussed.

122. The above highlights a paradox; that is, that “RBF losses” (in terms of RBF disbursements lower than allocations) seem to have been proportionally highest in Tanzania by the end of year five of the program (with one year left) than in the other two countries, yet government officials’ perceptions appear to have been more favorable than in Nepal and Mozambique. As Box 9 discusses, in Tanzania this arose from the fact that RBF actually compared favorably with other aid instruments that government officials had known previously. However, risk perceptions link to the balance between stretch and feasibility, and these perceptions, and the extent of risk materialization, may change as results become harder to achieve, like in Nepal (although for different reasons, see Section 3B).

Box 9 Tanzania's experience of disbursement compared to the recent past

RBF fund flows in Tanzania were 80 percent of the maximum available over the first five years of program activity. Only one DLI result did not pay out at all in one year (2017) – the numeracy learning assessment result, measured every two years and which required a minimum threshold performance in an external assessment. Annual disbursements ranged from 69 percent (first year review) to 85 percent of the potentially available maximum and the total amount disbursed rose each year. Funds were disbursed twice a year as the financial DLIs (government budget releases on selected budget lines including school capitation grants) were measured every six months. There was no prolonged cessation of RBF funding.

This was better execution performance than the prior round of sector budget support, which included an unsuccessful performance tranche. It also compared favorably with the experience of projects – notably, and often mentioned, a large World Bank secondary education project that closed after several extensions, and whose outstanding financing was the major part of the World Bank additional RBF financing in 2017. It also compared favorably with the GPE-financed and Sida-managed first project, one reason why the MoE insisted on the variable tranche of the second GPE funding round to go well beyond the minimum 30 percent specified in the GPE model (GPE RBF funding represents 54 percent of this second grant). Generally, weak project disbursement performance appears to have been a result of the difficulty for government officials at various levels to adhere to donor project procedures.

123. The above suggests that the links between performance on results, actual RBF financing, risks and risk perceptions are quite complex. The information about result performance communicated by the financial rewards is noisier on more complex programs. For instance, deteriorating performance over the years in terms of results on the DLIs in Tanzania contrasts with the increasing financing profile (due as just noted to additional funding from Year 4 onwards, and funders restructuring their programs to allow use of earlier unspent RBF allocations). At LGA level, in spite of annual RBF workshops to inform LGAs, many of them were unlikely to be able to distinguish between good and bad performances on individual aspects across the six DLI results that they were responsible for achieving and earning RBF rewards for. Non-RBF projects present in a few regions may have helped deepening this understanding, but this was neither country-wide, nor systematic. A similar pattern may now be under way in Nepal, where flexibility to results targets (see Section 3F) and use of rollover provisions will allow disbursements to be maintained despite struggling performance. The PFM4R in Mozambique had a more substantial technical assistance (TA) component, which may have helped communicate expectations and understanding of performance, but again the simpler program design and DLIs meant it was simpler to achieve disbursement, while less was also at stake in financial terms.

124. The extent to which RBF risks of unpredictability and/or volatility actually materialize in the course of implementation clearly influences risk perceptions by the government. However, past, current and future context-specific considerations also matter – including, past experiences of other modalities to which RBF instruments are compared (Tanzania), current experiences of non-RBF aid funding shortfalls (Mozambique), and perceptions that the balance between stretch and feasibility is or might be changing, with effects on the future risk profile of RBF (all three countries). In Nepal and Tanzania, the risk profile had indeed evolved in the life of the RBF programs (federalism rollout in Nepal, underfunded FFBEF in Tanzania) whilst in Mozambique this was still at prospect stage (depending on the size of future RBF programming).

RBF financial management and alignment with country systems

125. Among the three study countries, there were various degrees of alignment of RBF with government fund flows and financial accountability systems. At first glance, RBF and non-RBF financial modalities were most “divergent” in Tanzania. In Tanzania, EP4R set up a separate MoE-controlled bank account system, following concerns of poor budget credibility and limited prior MoF commitment to release funds budgeted for education; it was felt that the MoF might not release RBF funding as additional/earmarked for education, and to address this, RBF funds were paid into a separate MoE-held and MoE-managed account. In Mozambique and Nepal, instituting a separate RBF financial system was not seen to be needed, and RBF funds followed pre-existing arrangements for education aid funding (FASE in Mozambique for the GPE and KfW; JFA with the MoE in Nepal).

The PFM4R in Mozambique, with the MoF as the lead agent, did not flow through FASE and was explicitly delivered via government financial systems, featuring as a “project” in the MoE budget.

126. RBF funding flowed to agents other than the MoE in the three countries, though only in Tanzania was this designed as a cascading of RBF financial incentives (see below). However, the RBF cascading fund flow, from the MoE to LGAs, was not aligned with government fund flows (MoF to LGAs) although reporting was combined on RBF and government transfers. In Mozambique and Nepal RBF funding flowed to subnational levels because it was commingled with government transfers – in Mozambique through the FASE system for the GPE and KfW and as part of treasury transfers for the PFM4R funds, and in Nepal through the LG conditional grants. These fund flows were aligned with government systems, but not directly associated with the performance of individual subnational agents. Mozambique’s PFM4R included elements of performance-based allocations (PBAs) within the program design, though they did not work in practice as intended (see below).

Downstream allocations from RBF

127. In the detailed RBF mapping undertaken across the three countries, there were some downstream allocations from RBF to subnational levels envisaged in each case. However, this took place in very different ways, and with different results. Only Tanzania had a successful subnational cascade of financing – and compared to other uses of RBF in other sectors in the country, this was quite circumscribed. More details are set out in Table 12, though this included:

- Formally cascading reward payments to subnational government was fully designed and implemented only in Tanzania (see Box 10 below). In Mozambique, these were planned in design although through PBAs to subnational levels and to schools, the design was not specific and it was not used in a systematic way, in part due to the MoF reluctance during implementation. Attributing RBF rewards to local governments for their performance on identified DLIs was not evoked in Nepal, though government subnational funding structures were fundamentally changed as a result of federalism, with most now transferred through conditional grants to LGs in the new federal set-up.
- There was some use of performance-based school grants in all three countries; very small in Tanzania (2 percent of all schools); piloted though not under any of the RBF programs in Mozambique; with a DLI on performance grants in Nepal, though with unclear implementation status, in part again due to federalism reshuffling of roles and responsibilities, including with regard to school funding.

Table 12 Subnational cascading of RBF payments

Cascade design	Mozambique	Nepal	Tanzania
Financial incentive cascading	Planned in PFM4R through use of PBAs; did not happen to the intended extent, no clear PBA design at outset. PBAs to school level budgeted in the program were not spent. No financial cascading planned for the GPE VT or KfW RBF.	None except pilot school performance-based grant (part of DLI 9 in the SSDP and funded by the WB). The design made in Year 1 of SSDP involved a complex model (>40 criteria to assess for school eligibility), unclear extent to which LGs used the model to select schools for performance-based grants, uncertainty on number of schools reached.	One-third of potential funding and actual disbursements flowed from MoE to LGAs for results achieved on “LGA DLIs”. In addition, RBF funding for school infrastructure from national-level rewards (as decided by MoE/PO-RALG), comprised most of LGA EP4R financing in at least one year. LGAs aware of the possibility of these funds, acting as additional incentive.
Interaction with government PFM system	PFM4R part of broader subnational/service delivery/sector PFM program. Aimed to test use of (RBF-financed) PBAs in government system at all sector levels. Not used as planned.	‘Chinese wall’ between JFP RBF disbursed funds and use of money by government. JFP including RBF funding, commingled with government conditional grant, flows to LGs,	Program circumvented poor budget credibility bottleneck, through non-aligned account holding & fund management by MoE/PO-RALG. (RBF-funded) School Improvement Grant is the only use of PBA in

Cascade design	Mozambique	Nepal	Tanzania
	Performance-based school grants piloted outside RBF program, plan to scale up in spite of poorly documented results. FASE funds (including RBF) flowed to provinces, districts and schools from Treasury, to be spent on activities in MoE's annual plan of activities.	but not directly linked to LGs performance or lack of. Major changes to the fiduciary and governance DLI 9 to reflect the changes of federalism which fundamentally changed the plans at the outset of the SSDP to reform grant systems, procurement and other aspects of PFM (see Section 3F).	government system, concerning 2% of schools each year. No engagement with inter-government fiscal transfer system (funds flow MoF-LGAs) in spite of RBF cascading down to LGAs.

128. Thus, whilst some RBF funding was flowing to subnational levels in Mozambique (as part of the FASE funds transferred for activities in the MoE's annual plan) and Nepal (through government conditional grants to LGs including budget lines for SSDP activities), only in Tanzania were some DLIs specifically assigned to LGAs by design, with rewards paid to LGAs depending on their individual performance. Box 10 sets out more details of this design, where about one-third of the RBF rewards were assigned to LGAs and accrued to them in terms of actual disbursements, though this is significantly less than in some other contemporary RBF programs focusing on basic services in Tanzania. Most LGAs thus annually received less than USD 1 (one US dollar) of EP4R funds per child in basic education; and as noted above, RBF fund flows were volatile for many LGAs.

Box 10 The use of subnational RBF payments in EP4R design in Tanzania

As of 2020, there are 184 LGAs in mainland Tanzania. One-third of the value of RBF in the EP4R was programmed to be routed directly to LGAs, and outside of the MoE, LGAs were the only other directly rewarded entity. The feasibility of direct rewards to LGAs was underpinned by the availability and reliability of EMIS data (post-2016) and the credibility of award decisions with the central government. Donors, while keen to disburse, largely stuck to the contract, and made joint decisions on almost all results. The MoE and PO-RALG respected this design, and the flows due to LGAs were received by them. Payments remained low relative to total per-student government financing, but may have comprised virtually all of the truly discretionary education financing available to LGAs.

How rewards were calculated

There was a range of DLI allocation designs with various implications and no clear story about why a particular incentive design was chosen. Some DLIs classified LGAs into four classes of performance, with a rivalrous design (the more other LGAs fail, the more an individual LGA's success is worth); other DLI designs were not rivalrous. Some DLIs (added from 2017) paid LGAs for improving their own baseline performance. Payments did not account for poverty or remoteness, commonly considered a desirable feature of Inter-Governmental Fiscal Transfer systems and practiced in the health sector basket fund. Nor were there area-specific adjustments – for example, the national problem of girl transition was incentivized even in regions with lower male students' transition. Some payments were scaled, according to the number of schools or students, but most were not. This meant that the per-pupil/school amount, a reasonable proxy for an LGA's effort, had little relation to the assessed performance.

Use of RBF proceeds

The MoE/PO-RALG issued guidelines on how LGAs should use their RBF rewards on activities aimed at achieving the LGA DLIs, and school infrastructure. The MoE and PO-RALG sent additional funds for construction from "their" share of the RBF proceeds. Most often this money was transferred to school accounts directly, using a new system introduced post-2016 to facilitate direct health and education facility grants. Indications are that this resulted in lower unit costs than under centrally executed projects.

Information, autonomy and finance

In practice, LGA offices act more as deconcentrated arms of national ministries/PO-RALG than as locally-accountable administrations, so that there was a range of sensitivity of outcomes to LGA effort. Many LGAs clearly tried but were disadvantaged, by the design of allocation formulae or by central underfunding of basic education constraining their ability to hit targets. But satisfaction appears high; most interviewed LGA staff understood what was being targeted and perceived the EP4R disbursement as additional. However, few could clearly articulate what was done differently as a result of the emphasis placed on more complex targets (retention, transition). The dominant

incentive was to focus on the presidential priority of access, and not the EP4R focus of learning. At the same time, while many construction projects were funded, this was insufficient relative to the increase in access post-FFBE, represented a relatively small proportion of overall school construction, and was not exceptionally more intensive than earlier programs.

One concern is the limited time there was for payments to be explained and results to be discussed, so that LGAs were insufficiently informed on their own performance and routes to future improvement on up to eight annual results. Six years into the EP4R, the government has recently started to set up more formal knowledge-sharing arrangements between LGAs. Another concern is that program design stopped short of increasing autonomy at LGAs; guidelines on use of resources, including in some cases which facilities were to be upgraded using EP4R proceeds, were set by the center. Timing of the payments was also an issue for alignment with the LGA budget process.

The clear delineation of national and local DLIs had a clear downside: national DLIs were not meaningful for LGAs even when LGAs were needed to achieve the results. Notably, underperformance on learning was hurtful to central government, but had no repercussions for LGAs. An alternative approach would have been a single pool with predetermined percentage splits, as used in a contemporary water RBF design,¹ increasing the transmission of targets to LGAs with shared responsibility for the goal, while getting LGAs to think about education quality more holistically; though this approach would have its own demerits.

¹ In the World Bank water RBF, the reward pool is split so that 90 percent goes to LGAs and 10 percent to central agencies.

129. In a context in which most LGA-level government funding goes on teacher salaries, the additional nature of RBF funds was also highly valued at that level. The MoE/PO-RALG imposed restrictions on LGAs' use of the funds. However, this did not emerge as a concern in LGA-level interviews, likely due to the fact that this earmarking was less tight than that imposed by the MoF for the government's core budget, and directed EP4R funds onto activities which LGAs felt were useful, including school infrastructure development.

130. The wider lesson could be that even in a cohesive, relatively developmentalist state, directly incentivizing agents mattered. However, the incentives from the EP4R did not propagate upwards – to the powerful central agencies – nor really downwards beyond the LGAs, that is, to schools and Ward Education Officers. The exception was PO-RALG, thanks to a handful of powerful, professional civil servants with good career incentives and a reordering of accountability of PO-RALG post-2016. The lack of propagation was partly because construction spending, the choice of central government for spending most of the proceeds of central government and LGA DLI results, is not as information-diffusing as, say, information campaigns or increased LGA-school oversight – for which funds from government budget were likely very scarce.

131. In Mozambique, an important design difference between the PFM4R and the GPE and German (KfW) uses of RBF was in the use of downstream financial incentives. The GPE and KfW variable tranche payments were paid into FASE. There may have been specific departments with lead roles in achieving DLIs, but the reward was recognized as funding for the overall plan / annual activity plan. The PFM4R on the other hand had intended to use performance based allocations (PBAs) at each level of the system, to utilize and cascade the RBF incentive effect. Staff members at the central and province level had expected individual payments, which were vetoed by the MoF, and it seems that incentive payments were only made as in-kind prizes for the top three performing provinces and districts against the DLIs. The greater part of the intended PBA budget was instead used to finance district operations, workshops or other capacity development support. The intended use of PBAs at school level did not happen under the program, although performance-based school grants have subsequently been piloted in selected provinces and districts, with World Bank non-RBF funding. In general, PBAs were harder to operationalize than had been anticipated.

132. Nepal did not have an equivalent financial cascade for RBF incentives, with the effective incentive all at the top level, in terms of the MoE and CEHRD. The design and implementation of a performance-based grants to schools scheme was a DLI target. However, this was not linked to the other RBF programming of the SSDP in any clear way. Financial incentives were though important, in all three countries at the national level, and at LGA level in Tanzania as noted. In Tanzania, in addition, 2 percent of schools received improvement grants,

supposedly to the most-improved and best schools ranked according to high stakes assessments (exam results). However, this was awarded mainly to elite institutions in a manner that was not always clear, via a semi-autonomous central agency which otherwise appears invisible in the sector and took a high share of the cost of the activity as administration costs. A further 20 percent of schools may have been recognized in non-monetary terms (for example, awards in public ceremonies officiated by Minister, or Permanent Secretary). Recent research in Tanzania suggests the publication of exam league tables assists the lowest performing schools but results in weaker students being prevented from sitting exams.⁷⁴ Moreover, schools react mostly to the local reputational incentive of high exam results, which casts doubt over whether rewarding the top 2 percent is useful.

133. The following Section 3D looks at other aspects of accountability cascades and relationships between different tiers of government, as well as between government and development partners.

⁷⁴ Cilliers et al. (2018)

D. Coordination, capacity and accountability

134. The theory of RBF usually places it in the context of a *principal*, who pays for results, and an *agent*, who undertakes activities to achieve results and receives funds if they are achieved. In practice, in most RBF designs the *agent* is made up of a set of institutions within government, including the MoE, MoF, and subnational tiers of government, while the *principal* may include a range of development partners, themselves complex institutions. Collaboration between principal and agent can make the theory's assumptions too simplistic; development partners *work together* and with government to plan, to evaluate and to achieve results. Joint sector forums offer the means for this collaboration, while joint programs may also involve technical support and capacity building in which partners are directly attempting to assist government to improve education systems and achieve results. Such mechanisms mean that the RBF incentive is always only one factor at play in a moving set of other factors that determine how institutions interrelate, cooperate and hold one another accountable for results.

135. As institutional relationships evolve over time, and in the absence of measurable counterfactuals, it can be difficult to isolate the changes RBF makes. This assessment has involved looking closely at the existing set of accountability relationships in each of the three countries, including for setting targets and developing operational plans, and for the financing, monitoring, reporting, and information channels by and through which institutions work together. For RBF to succeed, incentives should work with and/or strengthen relationships. Our research involved interviews with stakeholders from across the government system, as well as development partners, other sector stakeholders such as NGOs and civil society organizations (CSOs), and subnational government stakeholders as well as head teachers of schools. Detailed lists of stakeholders consulted can be found in the final country reports for this assessment.⁷⁵

Institutional relationships in the three countries

136. The three countries considered have important differences in institutional set-ups, in terms of relationships between ministries within central government, and between central/federal levels and subnational administrations and government levels, and how they relate to schools. As set out in Table 13, including for the major changes that have happened in Nepal with federalism, these differences are important in terms of how RBF was designed and functioned in each case:

- **Mozambique** can be characterized as having the most centralized government set-up among the three, with district education structures managed by and accountable to the MoE for implementation of education policy and plans, while also being accountable to and part of the provincial administration. RBF program design, in particular for the PFM4R, sought to strengthen this vertical accountability through improved supervision and fund flow. It also looked at horizontal accountability from schools and their directors to communities and their involvement in school councils, and the RBF design explicitly recognized that these latter structures have been weak. The design was not able to address the added complexity of parallel accountability of school directors, appointed by provincial governors.
- **Tanzania**, in principle, has decentralized, autonomous LGAs, and horizontal accountability to elected local councils. In practice coordination and supervision by administrative Regions as well as PO-RALG often amounts to "directives". Technical oversight is provided by the MoE through a centrally-accountable inspection and SQA structure, and funding of LGAs mostly comes via tightly earmarked conditional grants. RBF program design tacitly followed the ongoing centralizing trend at play in government, strengthening LG vertical accountability without engaging with local horizontal accountability, and engaging relatively weakly at school level.
- **Nepal** has seen a complete overhaul of accountability functions since 2017. The shift to federalism included decentralization of education delivery to 753 LGs, from a previous tier of 75 Districts. Parts

⁷⁵ Final reports: Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021).

of the old system were disbanded including a network of over 1,000 Resource Centers (RCs) staffed by Resource Persons (RPs). Legislation has taken a radical interpretation of the 2015 Constitution in devolving powers to the local level, which has limited vertical accountability mechanisms during the SSDP period, though MoE control is being strengthened, including through earmarked conditional grants. The RBF program was designed before the exact nature of changes was known and this has caused major implementation challenges (see Section 3B).

Table 13 Accountability functions for the three countries

Function	Mozambique	Nepal (pre-federalism)	Nepal (post-federalism)	Tanzania
Government tiers	MoE 10 Provinces (including Maputo City); 111 Districts; 52 Municipalities.	MoE / DoE 75 DEOs	MoE / CEHRD 7 Provinces 77 EDCUs 753 LGs.	PO-RALG / MoE 184 LGAs.
Delegation	Deconcentrated provincial and district administration delegated by MoE for basic education management & program delivery.	Deconcentrated DEOs supported by 1,050 RCs and network of RPs. Delegated to for basic education management & program delivery.	Constitutionally-established LGs with decentralized responsibility for basic and secondary education; contested degree of devolution/ autonomy; EDCUs "support" LGs but also role to be defined. Provinces responsible for teacher training.	De jure autonomous LGAs coordinated by PO-RALG; de facto basic education management & program delivery delegated by PO-RALG(Ed)/MoE to LG education offices.
Finance flows	Government (inc. PFM4R) and FASE (inc. GPE) funds from MoF to MoE, province & district education administrations.	Funds (Government & JFP including RBF) from MoE to districts & RCs.	Government & JFP including RBF funds through conditional grants, with budget lines & earmarked allocations for SSDP program delivery by LGs.	Government funds from MoF to LGAs through conditional grants. RBF funds from MoE (not MoF) to LGAs, tightly earmarked too.
Reporting	Administrative reporting to MoE; financial reporting for government (and PFM4R) funds to MoF; for FASE funds via FASE Controller in MoE and joint technical Working Group for FM. Financial reporting via government PFM system (e-Sistafe).	DEOs & RCs report to MoE; line-by-line financial reporting.	LGs report to Ministry of Federal Affairs; MoE seeking to re-establish reporting through EDCUs; ongoing rollout of LG financial management/ reporting system. Financial reporting via <i>Sutra</i> treasury system. EMIS bypassing LGs is more "stable", though with challenges.	LGAs report to MoF through PO-RALG-managed system for government & RBF funds. Local education officials report to PO-RALG/Ed. Technical oversight by MoE deconcentrated inspection/SQA.
Motivation	District and provincial staff appointed by / under provincial governor and administration.	Local actors are central civil servants (promotion opportunities to central level).	LGs locally accountable. Redeployment of civil servants from formerly unitary administration ongoing; reluctance due to seniority & lack of promotion issues.	LGAs locally accountable; variable strength of LGA councils. LGA officers centrally deployed on approval from local executive director; opportunities to be deployed to higher levels.

137. RBF programs can be analyzed with respect to three institutional levels, and relationships, that are key to the education sector, namely: a) the relationship between sector agencies and central agencies (mainly though not exclusively the MoF); b) the relationship of central agencies with subnational institutions / local government administrations with regard to education; c) the relationship with schools. At both subnational and school levels

there are usually meant to be horizontal accountability mechanisms in place, with communities having a role in oversight and school management through School management committees (SMCs).

Relations between central government agencies

138. Both Mozambique and Tanzania RBF program included engagement with the MoF. They did so in quite different ways and with different results. In the case of the Mozambique PFM4R program, achieving the DLI results on the school grant timeliness and district financial management systems clearly required collaboration between the MoE and MoF. The financial incentive was also clear for the MoF as achieving results brought RBF funding into the government coffers. Moreover, the MoF was the central agent in the PFM4R, and the program had a Program Coordination Team located in that ministry, which along with the MoE received technical support. By the end of the program the intended school grant and district financial management systems were in place and functioning, and collaboration between the MoF and MoE did occur, although was challenging. Questions arise on the depth of change and whether the achievements addressed the fundamental sector PFM issues. For the RBF programs which worked through FASE, there was a less obvious engagement with the MoF although the MoE saw the PTR indicator as giving some leverage in budget negotiations linked to teacher recruitment.

139. In Tanzania, the MoE-MoF collaboration was and remained weak. The only way in which the RBF program engaged with the MoF was through the financial DLIs, aiming to secure government budget fund releases for about 5 percent of the education budget (albeit including the important school grant funding), and which in effect made the MoF an agent of the MoE, reversing the government's regular accountability relationship. The incentives for the MoF to "deliver" were rather weak, with RBF funding being strictly earmarked for education. As discussed earlier, results on this undemanding DLI were fair but deteriorating towards the end of the program; and RBF stakeholders largely failed to substantively engage with the MoF (and PO-PSM) on the issue of basic education underfunding and school understaffing. There was a one-off exception of collaboration between the MoE and MoF reaching an agreement and establishing the system for school grant direct channelling – although this was likely also pushed by the President's interest in funds reaching schools given his FFBEF promise/decision.

Relations between central government agencies to subnational levels

140. The three country experiences with RBF saw varying levels of achievement in relation to strengthening the links between central and the local level, and there has been a contrast between what was originally envisaged and what was achieved. There appears to have been greater success in Mozambique, a unitary country with deconcentrated administrations, and least in Nepal, where the rollout of federalism drastically changed the nature of these central-local links. Tanzania has been somewhere in between the other two cases.

141. The most striking story here is in relation to the rollout of federalism in Nepal. Many RBF indicators became redundant and required flexibility to change (see Section 3F), in particular with the removal of the District Education Office (DEO) level, which used to be the main cost center and hub for dissemination of funds and the MoE's education policy and strategies to schools. The decentralization created a new cost center at the 753 LGs taking on the DEO roles. As further outlined in Box 11 below, the rollout of federalism also limited the control of the MoE to direct the implementation for a number of DLIs whilst other possible ways of trying to align LGs and the MoE's incentives were yet to be worked out.

Box 11 Nepal – the limits of a top-down “control” approach under federalism

Following federalism, the main means of control with regard to program delivery by subnational levels has come via the Conditional Grant, introduced in 2017, flowing from the MoF to LGs. The Conditional Grant includes labeled budget lines and amounts as specified in annual workplans and budgets by the MoE, including for DLI-related activities. LGs are not authorized to reallocate funds between budget lines, and implementation details are set out in the Program Implementation Manual (PIM). Initially LGs were not supposed to financially report to the MoE, although the MoE is seeking to re-establish some reporting through the deconcentrated Education Development and Coordination Units.

By November 2019, Year 3 results were not being achieved for many of the DLIs reliant on this control mechanism, which mostly related to ways of supporting schools or teachers. Awareness (hence likely use) of the PIM by LGs may be limited, including in interviews conducted for this research, and this partly is due to staffing and capacity issues with many LGs still not having the allocated position of “Education officer” in place. Whilst budget allocations were being spent as intended, the details of the DLI metrics appear to be at high risk of not being adhered to by LGs – with the reporting not providing information on details and these not being evaluated. For instance, the design of the “Performance-based grants implemented in 400 LGs” DLI includes criteria to be upheld by schools for performance-based grants to be disbursed. In practice, reporting for this DLI by the MoE/CEHRD has become simply the submission of a list of schools, and this is a similar story for many of the SSDP DLIs reported.

With only limited verification, there is a lack of clear evaluation of what took place or of the quality of implementation for many DLIs. Quantitative and qualitative analysis of the DLI on out-of-school children carried out for this research suggests similar “incentive alignment” issues. This is indeed the case for many of the SSDP interventions including those associated with DLIs, and is unlikely to change without a substantial package of technical assistance at LG level. There is also a disconnect between the SSDP operational model and the principles of decentralization set out in the Constitution and via the legislation of federalism in the 2017–18 period, i.e. that LGs should have a degree of autonomy and decision-making power in how to prioritize education spending. There is yet to be a reconciliation with this new reality and the top-down design of the SSDP may conflict with the principles of federalism.

142. As noted earlier, it is only in Tanzania that RBF financial incentives cascaded down to local governments, which received known amounts of program funding for results on clearly identified “LGA DLIs”, disaggregated by local government. This cascading, the prospects for LGAs of getting additional funds for school infrastructure from the central RBF rewards, and PO-RALG/MoE guidelines on the use of local governments’ RBF rewards, ensured strong central-local alignment of incentives. However, the program design mirrored the government set-up of limited *real* autonomy of local governments with a top-down, instrumental approach that ignored their role as distinct decision-making entities, overlooked local institutions and accountability, and engaged at local level strictly to focus on the LGA DLI results. The RBF program led to stronger links between MoE, PO-RALG and LGA education officials but this remained narrowly focused on the LGA DLI; there was, and continued to be by the end of the program life, no institutionalized coordinated intergovernmental planning system in place. This narrow engagement at local level may well have weakened the program’s effectiveness; see Box 12 below.

Box 12 Tanzania - An instrumental approach with insufficient support to key agents

The instrumental approach to LGAs was already a feature of the BRNEd process in 2013: LGAs were “told what to do”. Throughout the RBF program the main mechanism of engagement with LGAs was through annual workshops bringing together MoE, PO-RALG, Regional officials and LGAs to discuss EP4R issues – after the annual review and once LGA reward amounts were confirmed – and through MoE/PO-RALG guidelines on the use of RBF rewards by LGAs. Local executives and councils in effect rubber-stamped the RBF reward budget allocations made (ex post) according to the central guidelines. Although the workshops were reportedly found useful, they focused on explaining the RBF program “mechanics” (DLI metrics and any change, LGA reward amounts, guidelines), with little space for joint discussions of results and of what LGAs could do to achieve better results. In other words, there were no “EP4R implementation plans”, as one government stakeholder noted.

The RBF program also supported a number of other centralizing features, such as a system of centralized posting of new teachers to schools by PO-RALG and the SQA reform which may further strengthen influence from the center in spite of the diversity of challenges faced by students and schools in different contexts. Moreover, the level of institutional support to LGAs did not seem commensurate to the complexity of the results they are supposed to

achieve, nor did it take into account their very diverse situations with regard to resource levels, institutional capacity, and pre-existing educational outcomes. This “light touch” approach of the RBF program contrasted with the more intensive, hands-on support approaches of some other education projects, engaging more broadly and deeply for instance in involving local councils or supporting local-level education planning. Reportedly, such projects directly contributed to the EP4R results, for instance in supporting LG education officers in establishing education-focused coalitions, which could then be mobilised around the results to achieve.

143. In Mozambique, the presence of a central control mechanism, combined with implementation units within the MoE and MoF and coaches and facilitators down to province level, enabled the delivery of the PFM4R DLIs, and this in turn contributed to strengthening key subnational links in the education system. The Education Sector Plan is implemented through three-year operational plans and annual action plans. These are reported to be quite top down, but unlike in Tanzania and Nepal they do exist, and are used to inform province/district plans and spending. Districts play a key role in transmitting and communicating sector policy, plans and results to school level, and connecting schools to the broader system. The RBF focus on strengthening district supervision of schools was therefore appropriate as supervision was identified as an important link, increasingly seen as a step towards monitoring and improving teacher attendance, even though many schools may still not receive any supervisory visits during the average year. At its most successful, including for the PFM4R DLI on the timely disbursement of school grants (*Apoio Directo às Escolas* - ADE), coordination between the different tiers of government was essential for success, as set out in Box 13 below.

Box 13 Mozambique PFM4R: Achieving timely disbursement of direct grants to schools (ADE)

Mozambique has had school grants in place since 2003, and these are still principally donor funded through FASE. There are two tranches of school grants, which are based on per student formulas and are worth approx. USD 1.50 per student. The first tranche comes at the beginning of the school year, the other in the second half. At the outset of the PFM4R, the first tranche due in February was often not disbursed on time and became the DLI. Addressing this required problem-solving down to school level, and coordinated action by multiple agents at different levels, including the MoF, MoE and districts:¹

- 1) Central level: To address delays in budget approval and in availability of donor funding ahead of the new school year, a joint action plan was developed, with the MoF responsible for mobilizing funds to off-set delays in donor disbursement and a multi-sector team in charge of the disbursement process.
- 2) District level: With the district education structures becoming budget holders, they were trained on their new responsibilities and school grant funds were reflected in their budget for better monitoring.
- 3) School level: Through collaboration between the MoF/MoE all schools were registered as legitimate recipient of public funds and many opened a bank account. Monitoring was also improved at that level, with a monitoring tool to track when funds arrive at school.

The PFM4R contributed to this through the financial incentive of the associated DLI as well as the active role of the coaches and facilitators (e.g. facilitating MoF/MoE meetings and the development of joint action plans; revision of manuals and procedures; development of monitoring system, training of district teams).

¹ OPM, 2018.

144. While the MoE provides direction to provincial education departments and the district education services in Mozambique, these subnational levels are part of the provincial administration, under the Provincial Governors. This leads to dual lines of accountability, which may at times undermine sector objectives of the MoE; for instance, school directors are not appointed on merit, and are often not motivated to improve the quality of teaching and learning in their schools. The GPE variable tranche DLI on school director training was an attempt to at least try to improve their skills and motivation, given that conditions for hiring, firing and progression of school directors were seen as too political for RBF at the point of design. As set out below other means were also therefore sought to engage with school-level governance.

Engaging with school-level governance

145. In none of the three countries did the RBF program design entail direct cascading of significant amounts of RBF reward to schools. In a minor way, the EP4R funded the school improvement grant rewarding very few schools for improved exam results, but the introduction at larger scale of performance-based school grants was not considered. Performance-based allocations (PBAs) to schools were originally planned and budgeted for the PFM4R in Mozambique, but never materialized during the program period; this evolved into a school grants pilot since, though not under an RBF program. In Nepal there is a DLI focusing on rolling out school-based performance grants although as just seen, there are serious challenges in ensuring that this is carried out with clarity or rigor, without which the scheme is unlikely to achieve real results in terms of better school governance.

146. RBF programs did however engage with school governance in a number of other ways outside of the use of financial incentives:

- In **Mozambique**, accountability at school level was a direct focus of the PFM4R school council strengthening and supervision DLIs. Other DLIs potentially contributed to school management with training of school directors (as above, the GPE DLI) and the timely school grant funding (PFM4R). As seen in Section 3B, whilst results on these DLIs were achieved there were questions on the extent of achievement of the more qualitative aspects of the intended change (for example, extent of school council functionality, in particular, of change in parents/school director's relationship).
- In **Nepal**, school-related DLIs were all being challenged by the rollout of federalism, including DLIs which directly related to school governance such as performance-based grants to schools against minimum accountability requirements, strengthening school procurement processes, and increasing teachers' time-on-task.
- In **Tanzania**, besides a similar focus on ensuring school grant availability through the financial DLIs, the only school governance DLIs were the SQA system rollout DLI (indirectly) and the DLI on school score improvement, with two out of six SQA school performance measures focusing on school governance. By early 2020, the reliability of SQA school scores, especially in the "soft" areas of school governance and coming from a "school inspection" system, was potentially undermined by the high pressure on system rollout exerted through the DLI targets.

147. In conclusion, RBF programs engaged rather lightly, with uncertain achievements with regard to the goal of improving school governance. This possibly reflects the complexity of such a goal, which is not the culmination of a single, linear results chain. For instance, whilst results were achieved under the RBF programs in Mozambique and Tanzania in terms of improving school grant availability, there was more limited focus on improving the use of these funds at school level and genuinely involving the community in this. As illustrated by the DLIs on school supervision system, the issue of school governance also raises the question of whether and how RBF can effectively engage with complex institutional changes that can be difficult to reliably document through objectively measurable targets.

148. Overall, to address system change down to school level, RBF designs needed to align incentives for a wide range of actors. The financial incentive cascades were much narrower, especially in Nepal and Mozambique, than the range of actors whose action was required to achieve the DLIs; and not all such actors were reached in other ways (such as through guidelines, implementation manuals, training etc.). The clearest "gap" relates to RBF in Tanzania, where achieving retention and learning DLI results depend on supply-side measures incentivized by RBF being aligned with other measures on the demand side (parents, students). In some cases, a DLI requirement might not align with the mandate of a government institution. In Nepal, the new constitution meant that some DLIs would require redesign to reflect changes in the central-local-school relationship, but this was not foreseen.

RBF implementation structures, partners’ dialogue, technical support and capacity building

149. The use of RBF in each country was interlinked with various forms of capacity building and technical assistance (TA). The management of RBF programs relied on mixes of RBF-dedicated and mainstream structures in government, as well as on modalities of reporting on the DLI results and for the RBF dialogue. These “ancillary components” of the RBF programs are first summarized in Table 14, with some important similarities and differences highlighted and discussed below. This shows that quite an extensive array of dialogue, technical assistance and support, and capacity development were associated and interlinked with the RBF programming across all three countries.

Table 14 RBF “ancillary components” in the three countries

Ancillary component	Mozambique	Nepal	Tanzania
Government RBF management modalities	Dedicated PFM4R structures in MoF and MoE. MoE-level structure involved key departments in charge of DLIs, which ensured good understanding of the DLIs. Provinces/districts aware of targets (though mostly not disaggregated). No dedicated structure for the GPE VT. Targets cascaded for training institutes.	No dedicated RBF structures in MoE / CEHRD, although clear de facto structures, including for connected agencies. No dedicated provinces/LGs RBF structures. SSDP/RBF budget lines used conditional grant. Limited awareness of implementation details in PIM.	Dedicated GoT BRN structures, succeeded by EP4R dedicated program structure in MoE & PO-RALG (Coordination Units). CUs linked up with the other units concerned by/leading on DLIs in each agency. LGAs aware of disaggregated LGA DLI targets. No dedicated RBF structure at LGA level.
Reporting on DLIs	DLIs part of broader MoE annual action plan, framework of indicators. Reporting integrated in pre-existing (FASE-associated) sector reporting. Some DLIs continued to be reported on beyond the programs.	DLI reporting integrated in SSDP partnership dialogue (SWAp, JFA), with annual review process (BRM + JRM) generating large amount of time and reporting documentation given JFA complexity. Supported by technical working groups.	Annual EP4R report distinct from sector performance review report, with overlap in evidence bases used in both reports. DLIs differ from sector plan KPIs (some are similar but not identical).
RBF and sector dialogue	Program dialogue embedded in broader sector dialogue, including annual review meetings. Dialogue perceived to be focused on indicators more than substantive content / barriers, possibly too many indicators, not enough attention to quality of change in intended reforms. Sector review becoming more results focused.	Problem-solving nature of SWAp dialogue appreciated. Consensus that RBF-associated focus on DLIs was welcome (from SSRP “jamboree” to greater focus on agreed priorities). However, focus on DLIs seen as bringing a potential risk of project-tied funding (in spite of fungibility of donor funds). “Target obsession”, “lack of visionary planning” also cited.	RBF through “data DLIs” reactivated annual sector reporting. Distinct reports reviewed in separate processes by different audiences (high-level RBF meetings “closed” to non-RBF partners). Funders’ individual decision-making on disbursements. RBF dialogue collaborative but became overly focused on measurement issues; insufficient attention to results and what to do to improve them.
Technical Assistance (TA)	PFM4R included coaching & capacity development (CD) windows. Coaches’ role as facilitation, problem solving; TA at subnational level (provincial facilitators) seen as effective. Separate demand-driven CD window, less effective due to fund flows. PFM4R benefitted from other project TA/CD (e.g. GiZ & USAID re district supervision) or “opportunistic” (CSOs strengthening school	Mix of “on-budget” and “off-budget” TA for SSDP, alongside RBF/non-RBF funding of JFPs. UNICEF (non-RBF funder) TA vital to support SSDP JFA/ RBF processes (huge administration): technical reporting, “glue” between MoE and JFPs (trust). British Council EU/ADB-funded TA providing more technical support (EMIS).	TA key facilitation role, focused on RBF processes; support to DLI-associated system strengthening. Policy analysis/advice constrained (lack of time, sensitivity). Lack of systematic CD (GoT officials drawn on other tasks, often with financial benefits – per diem etc.). Some question on whether TA was “glue” or “screen” between GoT & funders.

Ancillary component	Mozambique	Nepal	Tanzania
	<p>committees). Intended funding for communication on school council role (public awareness campaigns), effectiveness unclear.</p> <p>GPE VT limited TA/CD support within ESSP / FASE. Links with RBF/DLIs not systematically planned for.</p> <p>KfW funded TA to the construction department, not effective (targets not met).</p>	<p>Limited support at local level, although number of NGOs and CSOs support districts and some of most marginalized groups in country (e.g. World Education).</p>	<p>Occasional WB/Sida TA on specific issues. External TA/CD also key – somewhat planned (several projects' support to development/ rollout of early grade teaching reform & learning assessment system) or opportunistic ('on the ground' projects support to in-service teacher training, community engagement indispensable for retention and school supervision DLIs).</p> <p>'CD window' through DLI 9, not well understood.</p>

150. The extent to which RBF programs were embedded in sector/subsector structures and processes was quite different in the three countries. Central-level RBF implementation structures and reporting and dialogue processes were more distinct in Tanzania and for the intersectoral Mozambique PFM4R program. In Tanzania, there was strong coordination within the RBF program, but less so with the broader sector. The Mozambique PFM4R was run as a separate program, but was nonetheless well-mainstreamed, in part due to the work done by coaches / facilitators, but also the strong senior MoE involvement in DLI selection. The GPE program, as part of FASE, was well embedded in the sector dialogue framework in part due to its direct links with the World Bank's ESSP program. For Nepal's SSDP program, structures, reporting and dialogue were embedded in the approaches to joint working between donors and government. If anything, with RBF representing 80 percent of the SSDP funding, the dialogue on the DLIs reportedly came to dominate the joint sector processes, creating some tension for non-RBF funders.

151. In all three countries there was a similar concern arising from the perception that the dialogue was overly focused on DLI measurement and targets; a trend expressed as "target obsession" by a Nepalese government official. This was seen to distract attention from "deeper" issues such as the need for visionary planning (Nepal), the quality of teaching (Mozambique) and the reasons for adverse results (Tanzania), and may have led to insufficient discussion of what to do to raise the quality of change or to improve results. In Tanzania this bore the risk of reinforcing the broader "what gets measured gets attention" trend associated with President Magufuli's "drive for results". A related point is that there was limited focus on and time devoted to policy learning in the RBF program processes; that is, with little research undertaken through the programs, limited attention to third-party research, and the focus on DLI measures and targets crowding out policy discussions. This was perhaps a more critical challenge in Tanzania and Nepal, considering the scale of RBF and its ambitions in these countries.

152. Within government structures, there were varying levels of understanding of the RBF programs and DLIs. In Tanzania, possibly due to the key role of the Coordination Units, the understanding of the RBF program appeared to be limited in some of the central agencies directly responsible for results on DLIs. In contrast, LG education offices were well aware of the RBF "logic"; that is, the conditional nature of RBF funding and their role in achieving results, even though there were no dedicated RBF structures at that level. In Nepal, central-level actors were very aware of DLIs, but subnational actors much less so, although DLI results were integrated into the Program Implementation Manual which would accompany the Conditional Grant as the main mechanism for dissemination of the activities to carry out at LG level, they were not labeled as DLIs and were not truly conditional from the point of view of LGs. In Mozambique, key departments in the MoE were aware of the DLIs within their remit, and the PFM4R coaches/facilitators, while representing a parallel structure, helped build awareness of the program logic and targets at subnational levels, even in the absence of successfully cascaded financial incentives (the intended PBAs were not effective as set out).

153. Technical Assistance and other capacity development has played a central role in most of the RBF programming, though this has differed in terms of design and approach. Tanzania's EP4R and the Mozambique PFM4R RBF programs had technical assistance or capacity development windows alongside RBF funding, with a focus in both to technically support actions necessary to achieve results on the DLIs (for example, system development and strengthening). In Tanzania the "RBF TA" also played a key and arguably indispensable role in directly facilitating RBF reporting and the government and donor dialogue process, as well as providing significant analytical support for the processes in drafting the sector performance review report. Additional technical assistance was provided separately by RBF funders for specific DLI-associated tasks. The RBF-embedded technical assistance was also effective in the PFM4R in Mozambique, facilitating joint work by the MoF and MoE which was not found in Tanzania. Due to the nature of the DLIs the Tanzania technical assistance also had a greater policy focus, although this could be allocated relatively little time; there was some wariness on the part of the MoE of interference technical assistance, which was prevailing more broadly in Tanzania under President Magufuli. For Germany (KfW) long-standing bilateral technical assistance to the MoE's department of construction in Mozambique came before and alongside its RBF contribution to FASE; while not embedded, it was directed towards achieving the DLIs.

154. Other RBF programs had less explicitly dedicated technical assistance although had other support often performing similar functions. In Nepal, UNICEF, not an RBF funder, provided technical assistance support to the broader SSDP processes and has played a central role supporting planning processes, reporting on the DLIs, and the annual processes of sector review (the May BRM and November JRM). There has been separately funded technical assistance, by RBF funders, to support more technical tasks, within the main department for education under the MoE (the CEHRD) and contracted to the British Council; as well as other more ad hoc support to specific technical areas, including to the EMIS for example. In Mozambique the GPE targeted DLI results did not have explicitly focused technical assistance, although it could be funded under FASE. However, the DLIs targeted teacher training institutes (IFPs) for example, which also receive elements of direct project support from a variety of donors provided outside RBF programming and with no link to the DLIs.

155. Other contributions of this kind were unplanned, non-systematic support from other ongoing "on-the-ground" operations. In Mozambique, a broader public service capacity development program, POEMA, which Germany (GiZ) and UNICEF managed in selected provinces and the MoE in others, provided education planning and management training to subnational education officials and was broadly well thought of (especially where GiZ had implemented). CSO managed projects have a track record of working with school councils on local accountability, and awareness raising, and may well have contributed to change achieved through the school accountability DLI. In Tanzania there were projects implemented at subnational/school levels. These kinds of projects were not systematically harnessed for DLI reform areas as part of the programs.

156. Another important feature of, and difference between, RBF programs relates to the role of technical assistance /capacity development at subnational levels. In all RBF programs subnational-level agents were key to achieving the program results. Yet only in the PFM4R in Mozambique was there a dedicated form of support at this level through the provincial facilitators. They were well thought of, helped facilitate the change processes and training, although it is hard to assess how much capacity they transferred. In Nepal the rollout of federalism raised huge challenges with regard to supporting subnational agents, with about ten times more LGs than there were district education offices. At the time of completing the research it was not clear how adequate support to LGs was going to be provided, and many LGs still had not received allocations of sufficient staffing via the civil service adjustments by government.

157. While some use of TA was "embedded" or directly associated with RBF programs, support from other donor funded TA/projects also contributed to the achievement of DLIs and common areas of reform. Some of this work was planned collaboration for national-led activities. In Tanzania, this included funding of the early grade learning assessment system development and rollout by USAID, a non-RBF funder, supporting the learning DLIs; as well as capacity development support for the achievement of the data DLI targets, from a large cross-

sectoral local administration and IT system strengthening project in PO-RALG. In Mozambique, USAID and GiZ (both non-RBF partners) contributed to the development of tablet and online tools for district supervision. In Nepal, there is also a USAID EGRP which directly links to the DLI on early grade reading, and for which USAID joined as a DLI partner to the SSDP in Year 3 of the program.

158. Reliance on technical assistance was raised by the Tanzanian stakeholders as a concern. Indeed, towards the end of the program period it was clear that government structures on their own would not be able to “run” an RBF program of a level of complexity similar to that of the EP4R. The approach in Mozambique was narrower and in some cases achievable without technical assistance. However, reliance on technical assistance was also a feature of the PFM4R, for which intended changes were more substantial/complex and technical assistance was indispensable to make the progress that was made, in the timeframe. Reliance on technical assistance is also a feature of the SSDP within which the Nepal RBF program was embedded. However, whether or not reliance on technical assistance should be a concern is open for discussion. The Tanzanian experience, in particular, suggests that on cost-effectiveness grounds, some of the capacity to run complex aid-financed RBF programs should arguably continue to be outsourced, except, perhaps, when a government would envisage adopting RBF as its regular way of working. Even so, it might continue to be more cost-effective to carry on outsourcing some aspects of RBF (for example, policy-informing studies by independent academics, etc.).

159. In conclusion, technical assistance and capacity development support have played a crucial part in the RBF portfolios examined in this research. However, it could be better planned, notably with a view to strengthening complementarity between the policy and system-strengthening of RBF programs, subnational and local capacity development, and “on the ground” projects that can better support deeper institutional or “rules of the game” changes. This could include finding ways of harnessing the experience of CSOs/NGOs active at the school level. As one Tanzanian stakeholder put it: RBF programs are good at working “at the head level”, they need to be “given legs” through more operational projects.

E. Evidence-based policy and verification

160. The use of RBF provides the potential for improved focus on evidence of results, and this in turn has the potential to improve the policy-making process – *sharpening minds* of policymakers and sector stakeholders. RBF programming interacts with evidence systems in a number of different ways, including the DLIs themselves which can create indicators and metrics that may not exist to that point, and/or improve their measurement. The introduction of verification can provide a new source of data as well as the adoption of greater rigor in measurement. RBF may also interact with EMIS systems in a number of different ways, using evidence from EMIS for metrics themselves, or with DLIs specifically seeking to improve EMIS functioning, efficiency and accuracy.

161. The experience of evidence systems and RBF is one of the more complex areas in this assessment due in part to the number of ways to classify what is “good” data and information. The creation of data is incentivized in different ways, for example enrolment data within EMIS can be used to drive per capita funding, scholarship payments or a number of other transfers such as teacher pay. Many developing countries face the issue of “ghost” pupils and teachers.⁷⁶ Given such issues, it is important to seek a variety of data sources in understanding change within the sector. With RBF, the creation of a metric usually leads to an additional incentive around the information created (see Section 3G on Risks), and as a result, over time, the information within a DLI measure can become less reliable. This is one of the reasons why it is important to assess results in terms of non-incentivized measures as well as incentivized measures.⁷⁷ This is also relevant in that RBF programs usually target a subset of education results, yet other results are important to achieve improved students’ outcomes. Finally, and this applied to the RBF programs studied here, DLIs may not adequately capture whether more qualitative, “thicker” change is achieved,⁷⁸ suggesting the relevance of drawing from a wide range of types and sources of evidence.

Evidence within the RBF programming

162. Evidence systems were a key part of the design of RBF in each country. In Nepal’s SSDP this included the REACH grant-funded DLI on EMIS,⁷⁹ framed as a “foundational” issue as the “verification of IEMIS” (DLI 10) with a study on whether the data in the system matched verified information from a sample of schools. Despite this, in practice, very few of the DLIs in Nepal linked directly to the EMIS system – only the GER for secondary level [DLI 7.4, 7.5]. Instead the SSDP predominantly drew on a number of “ad hoc” DLI reporting systems, outside of mainstream government M&E systems. Some new evidence sources were created with the use of CB-EGRA for DLI 1, including an assessment in each of the first two years, and the Year 5 learning target linked to this. This was interlinked with the National Early Grade Reading Program (NEGRP, 2014/15–2019/20) from the end stage of the SSRP period, itself closely interlinked with the USAID program of very similar name, the Early Grade Reading Program (EGRP). Together this provided the basis for DLI 1 on early grade learning with the USAID assessment forming a baseline for the program. This accompanied the existing learning assessment system of NASA, which had undertaken a range of assessments on Grade 5 and Grade 8 level for Nepali, math, English and science (initial assessments done in 2011–13 period); NASA was an ILI under the SSRP RBF, and also a DLI in the SSDP.

163. More details on comparison of the ways in which DLIs interacted with evidence systems in terms of their source data, across the three countries is presented in Table 15.

⁷⁶ “Ghost” pupils are cases where enrolment statistics are inflated, perhaps due an incentive of per capita funding for schools. “Ghost” teachers are cases where individuals are on the payroll with salary collected but do not attend school and teach.

⁷⁷ Clist (2016): “one may think that the important consideration with a performance measure is that it is highly correlated with the latent variable, meaning that completion would be a good indicator if it were correlated with learning. However, the multitask model shows us that being correlated before a measure is incentivized is not enough: it must remain so afterwards.”

⁷⁸ See Pritchett (2015) for discussion on “thick” vs. “thin” change in education systems.

⁷⁹ Originally envisaged to be within the SSRP but rolled into the SSDP in early 2016.

Table 15 Sources of DLI information in RBF programming

Country and DLI	EMIS	School quality info system	New data-sets	Other data system	Ad hoc data collection
Mozambique (GPE, PFM4R, KfW)					
GPE DLI on PTRs from EMIS data.					
PFM4R DLI on District supervision – school performance data – through / more systematic visits + automated supervision					
PFM4R DLI - Strengthening of subsectoral financial data via result on district budget classification.					
GPE DLI providing new evidence of school director quality.					
KfW DLI – number of classrooms constructed.					
KfW other construction indicators (studies, hiring of technicians).					
GPE DLIs on numbers of teachers and school directors trained, became internal management indicators.					
PFM4R DLI on school grant timeliness reported during PFM4R period.					
PFM4R on DLI on school council functioning for PFM4R period.					
Nepal (SSDP)					
DLI 7 – GER data from EMIS.					
DLI 1 – CB-EGRA.					
DLI 6 – baselines + follow-ups on OOSC in targeted districts.					
DLI 1 – rollout of EGRP.					
DLI 2 – activity-based math, science and English language materials.					
DLI 3 – teacher deployment / TST monitoring / subject teachers in place / subject teachers trained.					
DLI 4 – single subject certification + G8 exams.					
DLI 5 – Model Schools.					
DLI 7 – scholarships from new schemes.					
DLI 8 – DRR retrofitting.					
DLI 9 – performance-based grants.					
DLI 10 verification of EMIS (new data to compare with EMIS).					
Tanzania (EP4R)					
DLI 2.2 School grant completeness, enrolment & school ID data.					
DLI 3.1 / 3.2 EMIS data collection and use of for Annual Reporting					
DLI 4.1 / 4.2 – PTRs, data from EMIS.					
DLI 7.1 / 7.2 – survival and transition rates from EMIS.					
DLI 8.1 / 8.2 – school quality assessment (SQA) government data + RBF and other TA.					
DLI 6.2 / 6.3 – data from EGRA / EGMA; other technical assistance to initially develop EGRA/EGMA.					
DLI 2.1 / 2.2 – Government release of school grant and on priority budget lines, using data from MoF FM system.					
DLI 5.2 – School Improvement Grant - exam result data system.					
DLI 9.1 New commissions granted re policy, planning and innovation.					
DLI 2.3 Textbook availability.					

Source: Authors' assessment based on DLI reporting documentation.

164. In Mozambique, there was less direct focus on EMIS, although the PTR GPE indicator was based on EMIS data. Evidence systems were integral to the PFM4R, as the budgetary classification DLI would improve budget-level reporting, ADE timeliness had not extensively been reported prior to the RBF period, and school supervision by districts created a growing set of potentially useful data on schools. Other DLIs relied on ad hoc data systems,

for example the school councils functioning DLI for the period of the PFM4R, and the GPE DLIs on numbers of teachers and school directors trained at IFPs, although these latter indicators are now part of the internal targets utilized by the MoE for IFPs.

165. In Tanzania, strengthening evidence was a more major focus of the RBF DLI framework. The EP4R engaged significantly with strengthening and using the EMIS, using an improved learning assessment and system, and strengthening and using an improved school quality monitoring system. In contrast, financial data systems were paid marginal attention except through the use of government financial data for the financial DLIs, even though LGA financial management systems are managed by PO-RALG. School-level statistics system DLIs focused on the production and national-level use of data, and specifically on strengthening school-level statistics (timeliness, coverage, reliability) and on the annual sector review report, with the aim of gradually augmenting its analytical content. A strong RBF contribution to positive changes has been found in this assessment, due to both the DLI incentive and accompanying TA. More recently, from Year 5 of the EP4R, a new school quality assurance system has become a DLI, focusing on attempting to change a “top-down inspection” approach to “school community participatory support”, and school quality measures generated by this system. Finally, learning has been a major focus, and the learning assessment system through focus on early learning, with DLIs generating new information on EGRA and EGMA.

The role of verification in evidence systems

166. Verification was an important part of the RBF programming in each country. This was arguably most strongly so in Tanzania, where it developed into a well-specified process that was central to the RBF program management. In Mozambique, the German (KfW) use of RBF in construction did not involve an independent verification agent (IVA) role. In Nepal with the SSDP, some joint financing partners (JFPs) do not require a formal IVA for all DLIs, for example the ADB (second largest provider of external finance to the SSDP by DLIs) does not require this for all its DLIs, instead seeking “endorsement” of the Local Education Group following the annual Joint Review Meeting (JRM) processes. However, in other cases, including for the GPE, World Bank, European Union and Finland, an IVA role is required. In Tanzania, after a first IV explicitly set out to follow a learning process, the annual independent verification of government reporting on the DLI results has evolved into a highly-formalized exercise using well-specified protocols and carrying out a government and funders’ jointly agreed scope of work. RBF funders are seeing this third-party IV as indispensable whilst its value is generally not contested by government officials.

167. The three countries saw a range of approaches to the hiring, contracting and the type of institution that took the IVA role. This included in Mozambique, the national audit institution (*Tribunal Administrativo*) which took the IVA role for the PFM4R. For the GPE variable tranche, the World Bank chose instead to hire a private entity for the IVA role (Ernst & Young). In Nepal, the RBF programming has involved two academic institutions in the IVA role – Kathmandu University School of Education (KUSOED) and Tribhuvan University, Center for Educational Research Innovation and Development (TU-CERID). In Tanzania, a private entity (Oxford Policy Management) was contracted for four years for the independent verification process to support the annual reviews.⁸⁰ This approach provided continuity which allowed the process to accumulate contextual knowledge and build stakeholders’ trust in integrity of the verification process, as well as enabling it to provide useful feedback on ways of further strengthening EMIS. This contrasts with Nepal which saw two different institutions being contracted in two years, with arguably a lower level of rigor in the second process.

168. In all three countries, with the possible exception of the choice of the *Tribunal Administrativo* by the World Bank for the PFM4R in line with the program’s overall focus of strengthening government PFM systems, there do not seem to have been strong rationales or even a discussion of the reasons for the other choices made. For instance, what were the rationales for the change in institutions in Nepal, the choice of a private contractor

⁸⁰ The first IV process was contracted to a different organization. It was very much a lesson learning experience, and considerably helped RBF stakeholders to design the further-specified process of the later years.

for the GPE variable tranche DLIs unlike what was done for the PFM4R, and the same question for Tanzania where other RBF programs in other social sectors were using the national audit institution. Table 16 outlines a number of features of the verification process across the three countries.

Table 16 Verification process used in RBF across the three countries

Verification	Mozambique	Nepal	Tanzania
Contract holder	World Bank commissioned IVA for PFM4R, and on behalf of the GPE for the VT. No IVA for the German (KfW) use of RBF.	Government commissioned IVA. Initially donors expected a multi-year contract, but the Year 1 IVA was replaced and retendered.	Consistently used independent verification by donor-commissioned third-party.
Verification agent	PFM4R, using the <i>Tribunal Administrativo</i> , the country's national audit institution. GPE DLIs, contracting of <i>Ernst & Young</i> , a private entity for IVA role.	Two universities, KUSOED and TU-CERID undertook IVA role, both for SSRP period. KUSOED took all of SSDP Year 1 IVA role. TU-CERID took over for SSDP Year 2 IVA role.	OPM contracted via DFID.
Process to define scope of work	IVAs worked with both World Bank and the MoE for both. <i>Tribunal Administrativo</i> had combined audit role for program.	IVA function not required for all DLIs, only required by WB + EU + Finland. Not required by some other JFPs including ADB for some DLIs.	Highly formalized process, annually jointly (govt/DPs) agreed TOR/scope of work. Agreed & documented verification protocols.
Scope of work	No cross-learning across RBF programs; three different systems; use of third-party for the GPE VT; no verification by KfW RBF. PFM4R process involved field work, relatively limited survey approach used.	Jointly (govt/JFPs) agreed scope of work. Use of IV not consistent among DPs. Large number of reports produced for all the different DLIs; a limited number of which involved field work (e.g. OOSC DLI 6).	Scope included check on GoT data and calculations, and triangulation including through gradually larger school / LGA sample survey, with explicit focus and dedicated reporting on qualitative evidence alongside quantitative data.

169. The experiences of verification have highlighted some important issues on the resourcing of the role, the challenges in how to manage the contract, the reporting to sector stakeholders, and whether and how independence can be maintained. Selected issues found in this assessment include:

- Specifying the IVA scope of work and verification protocols** – In Tanzania, this was very clear, and jointly agreed, which gave strength and credibility to the findings and as discussed in Section 3F, eliminated room for the creative ambiguity which was apparent in the assessment of some of the DLI results in Nepal and Mozambique. However, the Tanzanian experience also suggests the need for RBF stakeholders to consider the trade-off between the costs of increasing the rigor of verification (financial and in terms of time spent discussing measurement and verification protocols), and the relevance of the more rigorous evidence in terms of policymaking and/or evidence-based planning. This is illustrated by the case of the early learning assessment DLIs in Tanzania. In this instance, with the government now being in charge of carrying out the assessment earlier done by a third-party, development partners have insisted on the IVA “verifying” the government assessment methodology. In addition to this being resented as it is perceived as a lack of trust, a number of stakeholders, including technical assistance working with the government agency in charge, highlight that the standards being demanded by the IVA (seen as being pushed by development partners) for the methodology to be deemed acceptable are unnecessarily demanding considering the generally low learning levels being measured. This has the potential of being a “flash point” in the next review, and to further compound the tendency for review and other high-level meetings to exclusively focus on methodology and measurement issues at the detriment of attention to the results and their possible underlying causes.

- **Importance of adequately resourcing the verification process** – Tanzania and Nepal again offer an interesting contrast in this respect. In both cases, the IV budget appears to have been not fully worked out in the initial RBF programming. Whilst in Tanzania this was addressed by DFID stepping in with the necessary funding to properly resource a sophisticated IV process which, reportedly, was critical to ensure DFID headquarters' buy-in, this was not the case in Nepal. Budgets for the IVA role were very small, for instance representing just about 0.02 percent of the overall SSDP DLI disbursements for the Year 1 IV process, whilst the scope of work as outlined in the contract was complex – which may have limited the quality of the verification.
- **The verification process role in strengthening (or undermining) the integrity of government systems and reporting** – In Tanzania verification generally provided evidence that could usefully triangulate and/or complete government reporting (including qualitative evidence, for example, on [deteriorating] timeliness of school grants not measured by the DLI focusing on completeness); and feedback on evidence quality and systems, which was seen as useful by government especially with regard to the EMIS system. In contrast, the case of the OOSC DLI in Nepal (see Box 14 below) illustrates that the IV process may not always strengthen government evidence systems and reporting. In this case, following a first IV process which challenged the government evidence on OOSC enrolment, the government contracted a different IVA for the next year verification.

170. Underpinning several of these points, is the question of the government reaction to evidence “internal to the RBF”, if it brings up adverse/mixed results. As seen above, this led to the government challenging either the evidence itself or the methodology underpinning the evidence. A more extreme reaction was that of the Tanzanian education officials challenging the EGRA/EGMA approach in 2017 when it indicated learning losses whereas no such challenge was raised in 2015 when learning levels were found to have risen, but also eventually deciding not to allow the publication of the results. In this case, the main factor was not the financial loss as DLI targets were missed, but the embarrassment caused by the use of these adverse results by political opponents.

Box 14 Management and tendering of the independent verification agent (IVA)

The role of verification is central in RBF to ensure that results are “credible”. Prominent guidance allows a range of institutions to undertake the role of independent verification agent (IVA), and does not specify how tendering should be done, that is, via the government or by the development partner.¹ For the SSDP, the World Bank required the IVA for all its DLIs, as did the European Union and Finland. Other JFPs have been more flexible, including the ADB, which requires “endorsement” from JFPs at the JRM for some of its DLIs, for example the DLI on Model Schools.

The IVA role was performed by Kathmandu University School of Education (KUSOED) for the Year 1 verification. KUSOED was replaced as IVA after Year 1 by Tribhuvan University (TU-CERID). There was one DLI which KUSOED queried in the Year 1 verification process – the DLI 6 on targeted interventions to bring out-of-school children (OOSC) to school in districts selected according to the Equity Index (DLI 6.1a), and a 20 percent target for the reduction in numbers of OOSC (DLI 6.1b). This involved a government baseline, and a follow-up assessment. In November 2017, the government reported for the five Year 1 districts, a “22.4 percent reduction (target was 20 percent) in the aggregated number of OOSC was reported as achieved”. This would have implied that the government had successfully reached the target enrolled and the full amount for the DLI would be disbursed. There was USD 5.5 million potential disbursement associated with this indicator from the GPE, the EU and Finland.

The result reported by the MoE for Year 1 on the OOSC DLI (6.1b) was the first (and only) to be significantly challenged by the IVA. The exercise could only verify 81 percent of the names the MoE had provided through school register checks. The implication of this was that the target was not fully met, with approximately USD 0.5 million “lost” out of the potential disbursement. For the following year, KUSOED competed for the IVA contract, but were not successful. The story may be circumstantial but it appears that the fact they had questioned government reporting may have been a reason to discontinue the contract.

¹ For example in World Bank (2017b) or GPE (2019). ² KUSOED (2017).

EMIS data and RBF

171. As shown in Table 15 above, the three countries' experience with RBF included varying roles for EMIS. In the case of Mozambique, there was no foundational or ongoing work on EMIS throughout the RBF period. This is in spite of issues arising from EMIS affecting the quality of evidence on DLIs. A clear issue arose with the baseline for the PTR DLI, where a simple change to the district organization in 2013 (splitting and creation of new districts) increased the number of districts in which PTRs were above the threshold. This issue came up in sector reporting, but should have been apparent at the outset given the baseline was from 2014; and the issue led to progress on the measure not being recognized in the DLI achievement. In general PTR derives from enrolment and teacher statistics captured in EMIS, and evidence suggests that both are subject to the challenge of ghosts (ghost pupils and ghost teachers, see Box 15). This directly impacted on the accuracy of the PTR itself as measured by the DLI, whilst also raising questions on the depth of change in school governance as targeted by other PFM4R DLIs: the lack of integrity of the enrolment data directly affected the basis on which the school grant is calculated (DLI on grant timeliness), whilst in relation to the school council DLI it raises questions on their actual effectiveness or integrity.

172. In the case of Nepal, EMIS was specifically included as DLI 10, and presented as a foundational issue at the outset of the SSDP. Two verification exercises included within the DLI subsequently undertaken have been achieved and both found more than 95 percent accuracy of the system, which was the target figure incentivized. However, quantitative analysis we have undertaken using EMIS for this study suggests this may be an optimistic assessment.⁸¹ The data shows large year to year variation and has not aligned accurate back-dated information at the LG level before 2018 (the year the data switched from being reported by 75 Districts to being reported by 753 LGs). These may be understandable additional teething issues of federalism, however they raise questions about the accuracy of EMIS. EMIS data is not used for many DLIs, and while this may just be a design feature, a number of ad hoc data systems have been created for DLI reporting, mainly due to the nature of results targets. The EMIS DLI also included an indicator on "Web-based EMIS ... operational in all LGs in schools" (DLI 10.3a), part of the general move to an electronic system. Interviews suggested that there are challenges with the system, not least that many schools have no internet provision, while the previous network of Resource Persons (RPs) that supported schools in uploading information has been disbanded.

173. In Tanzania, EMIS has been integral to the RBF design, including DLIs on EMIS statistics and with EMIS providing the basis for measuring results on several other DLIs. This focus appears to have been effective – with broader coverage, more disaggregation, improved timeliness, and continued effort to strengthen quality/reliability of the school statistics. The much-expanded scope of analyses found in the annual sector performance review report and relying on the EMIS data is testimony to this – the 2014/15 first annual report made up just 45 pages whereas the 2019 annual report counted 209 pages of dense and detailed analyses (disaggregated at LGA level, by grade or other relevant category). The gender- and grade-specific enrolment data required to calculate the retention DLIs was simply not available before 2016. The first success of the RBF program in this respect was actually to incentivize government to decide on where the school statistics management would be located, in a complicated context of rivalry and duplication between the MoE and PO-RALG.

174. That said, this success story is not exclusively due to the RBF DLIs. Important other factors included the strong senior leadership support provided to a qualified and motivated professional team in PO-RALG (without any direct access to the considerable RBF rewards they were instrumental in earning); and TA, provided both by RBF funders and externally, as well as spillover effects of the massive TA provided by a large cross-sectoral administrative and IT system strengthening project in PO-RALG. As discussed in Section 3B above, this is a work-in-progress in that there is no evidence of significantly increased use of this improved data at LGA level, whereas at national level much of the analysis of the data done for the DLIs, while the annual report was relying on TA.

⁸¹ See Holden and Chapagain (2021), in particular the quantitative annex to the assessment.

Box 15 EMIS, enrolment numbers and ghost pupils in Mozambique

A significant limitation with the PTR as an indicator, is that it is made up of EMIS-generated information. Analysis of enrolment numbers in particular highlights a potential issue with *ghost pupils*, particularly in northern provinces. The proportion of pupils at different ages reported as enrolled, as compared to available population information by age¹ shows many more pupils enrolled than should be the case, particularly in Zambezia province. The data may indicate issues with census information on population, however it is more likely that it represents an issue with enrolment data. Deviation in growth rates of pupils also point to ghost pupils being a more likely explanation.

UNESCO notes possible “double registration” at the start of the school year – “it could be that many children first register with the school, but do not participate or drop out before the end of the year; they register again the following year but are not recorded as repeaters. Statistically they appear twice as a new entrant.”² Given that ADE school grants are linked to a per pupil formula, there is a risk that inflation of numbers could be happening to secure additional resources. In one interview we had with district services (SDEJT), we were told 2,000 ghost pupils had been removed from the data in their district. As this initiative appeared to be linked to their own professional drive as opposed to any top-down directive, we were able to corroborate this story in the district data.³ This suggests the problem may be much more widespread.

¹ MINEDH (2019). ² UNESCO (2018). ³ More detail in Patch et al. (2021).

The scope for performance evaluation alongside RBF

175. There was variation in the ways and means that RBF programming interacted with evidence systems across the three countries. However, a similarity across the three is that whilst the DLIs led to a degree of improvement in terms of quantitative data, inherent in their measurement, and in Tanzania the RBF also strongly contributed to the EMIS system being strengthened, none of the RBF programs engaged with the generation or use of more qualitative evidence or the issue of how evaluation of performance might be strengthened more broadly. Even in the cases of the school supervision and teacher monitoring systems being rolled out and incentivized, DLIs mostly “stop” at measuring the pace of the rollout. A number of points in this respect are discussed below.

176. In Nepal, the SSDP involves a large number of what can be considered projects, many of which were new to the education sector from 2016. As set out in Section 3B, the goals of these initiatives were generally clear, but the pathway to higher-level outcomes was often uncertain. This includes for example the DLI 5 on Model Schools, DLI 7 on new scholarship schemes, DLI 3 on a new time spent teaching (TST) monitoring process, DLI 8 on disaster risk reduction with resilient infrastructure in non-earthquake affected districts, and DLI 9 on performance-based grants. None of these initiatives built in clear evaluation of their benefits or cost-effectiveness. Given that some (Model Schools in particular) became large expenditure items for government, this is then a significant limit in understanding the effectiveness of the SSDP use of RBF itself.

177. In Mozambique, RBF programming and its focus had been stimulated in part as a response to the Service Delivery Indicator survey of 2014, which showed an alarmingly high rate of teacher absence, as well as low skills of teachers in the classroom, in addition to the National Learning Assessment of 2013 which showed very low levels of learning in literacy and numeracy. This led, in part, to the goals and programming linking to accountability reforms, and the focus on PFM systems, on supervision, functioning school councils, training for school directors, and training of teachers. However, once the RBF period was under way, the information presented became narrowly DLI focused, with DLIs formulated as narrow quantitative targets. Whilst it arguably may be challenging to formulate DLIs that would capture more qualitative aspects of the intended changes, broader sector information, available for example from civil society groups, was not absorbed into evaluation to this effect. The independent evaluation process for the PFM4R, carried out by the *Tribunal Administrativo*, was also somewhat restricted in generating more qualitative information, and it seems that the resources allocated and terms of reference also restricted this in the EY case as well.

178. Overall, Tanzania has seen the most coherent design with respect to evidence. There has been some use of improved evidence in policy/strategy development, for example in teacher deployment, and infrastructure development strategies. There was a stronger focus on qualitative field work and reporting within the scope of work of the IVA, although it is not clear this data was used. There was limited engagement of the RBF program, beyond evidence generation, with issues of use and ownership of data and system integration/interface, transparency in the use of evidence, and government communication on evidence. As seen above, for the last two issues this links to political challenges with high stakes data, particularly when evidence is adverse or mixed. In this regard it is also relevant to note the absence of integration of evidence generated by non-RBF, reputable third-party stakeholders. Some of it was challenging indeed for the government (such as evidence on the demotivation of teachers via the formative evaluation process underpinning the EQUIP-T project),⁸² but even evidence that would have caveated adverse DLI results or allowed to put them into a less narrow context, was not integrated – such as the reputable Uwezo surveys that yielded less negative evidence with regard to learning levels. With regard to learning, the RBF focus has also not brought up any changes in quality of evidence from other components of the learning assessment system (that is, exams), in spite of this being a stated intention – for instance in the World Bank programming documentation.

179. In all three countries, there has been some focus on school-level quality information, and some of this has been linked to the idea of performance-based grants. In Mozambique, the World Bank and government have been working for some time on the piloting of such a scheme, which would see schools gaining additional funds for good performance in a number of metrics linked to the supervision system, and this is being scaled up from 2020 with the new RBF programming. In Tanzania, there is an idea of replacing exam result-based school ranking as a basis for the school grant improvement DLI (thus far small in terms of both RBF funding allocation and number of schools reached) by much more comprehensive SQA-generated school performance data, which government also sees as indispensable to expand the use of performance-based school funding, though no programming to this effect has been put in place yet. Nepal has also included a performance-based grant as a DLI within the SSDP, although there is very limited information on what this scheme is, what it entails, or how it will be evaluated.

180. These initiatives all raise the question about the quality of school-based data, and the wisdom of linking financing to this data without a more comprehensive effort at ensuring sufficient quality.

- In Mozambique, the PFM4R focused simply on there being supervision visits. Interviews conducted for this research found that there are major challenges still with this system; many schools do not get visits, while district supervision agents find that they have insufficient resources to undertake visits, particularly for travel expenses to more remote rural schools. The data collected is apparently sent to province and MoE level, but this has not yet become a national database of supervision information, at least in sector reporting. The system is therefore very much nascent, and many accountability challenges are likely to remain in improving its functioning. This creates challenges to link this information to finance, as well as the challenge in how to achieve a consensus for the rational design to “synthesize” school performance data into an indicator(s) as a basis for the school grant allocation system – with the obvious risk of rewarding better-off schools that perform well with even more resources.
- In Tanzania, a similar challenge arises. The SQA framework actually includes a large amount of qualitative evidence, which in itself raises the challenge of standardization, and how to ensure the quality of the qualitative evidence, all the more so that the largest of the two SQA DLIs incentivizes a very fast rollout of the SQA system, not conducive to the quality of assessment made by the SQA officials. There was evidence indeed of formulaic formulations by the SQAs, which suggests that the threat to the integrity of the system in terms of quality is real. Qualitative evidence is also difficult to transform into a score; and there were also diverging views on how the synthetic “school score” should be constructed.

⁸² OPM (2019a) and OPM (2019b).

F. Adaptation and flexibility

181. Programs setting targets over a period of several years are likely to require flexibility to respond to change, and may need to apply adaptive management to respond to new realities and the successes or failures of existing and previous initiatives. The degree of *flexibility* of the mechanism may be a critical factor in how successful it is, including in how evidence-based policy and results-based management are used in practice. This means whether programs are subject to *adaptation* – including sector plans, strategies and policies, and how they respond and adapt to new information. The assessment of the three countries' experiences found that flexibility was required and often applied, but it could have been better supported, and better organized, to be applied more judiciously.

182. In both the Tanzania and the SSDP-era Nepal cases, there was a lack of a continuous lesson-learning process including exercises of anticipation, analysis of a wide range of evidence, discussions of the implications, and formulation of specific ways in which the program might adapt/show flexibility. This was not totally absent but appears to have been "ad hoc" and not systematic. In Tanzania, the high level of ambition of the RBF design (targeting increased learning levels with 9 percent of the RBF funding) arguably means there was a strong case for the program design to support a more systematic policy learning process. This might have been incentivized, through DLIs that would reward policy learning on issues arising from program implementation; at the very least it should have been allocated sufficient space and time in RBF discussions, and resources in terms of expertise for research, third-party research synthesis, analysis, policy advice etc. to feed into the RBF dialogue. In Nepal, an earlier assessment of the impact of federalism on the provision of basic education might have been possible, which would have better supported the flexibility that was required. There was also a lack of mechanisms enabling RBF stakeholders to formalize the way flexibility/adaptation was applied. The Tanzania and Nepal mid-term reviews (MTRs) were useful, but with programs of that complexity within fast-evolving contexts, such stocktakes should arguably have been more frequent. Annual reviews focused too much on DLI reporting to leave sufficient space for these kinds of broader exercises.

183. The period of RBF considered for this report was relatively shorter in Mozambique (2014–18) than in the other two countries. The RBF program in Tanzania (2014–) and in Nepal (for the SSDP from 2016) are both ongoing. However, in all three countries, these have been eventful years with major developments in political economy that have created challenges for the education sector. These include the following:

- In **Mozambique** the 2015-16 "hidden loans scandal" and associated suspension of GBS. The scandal had implications for the amount of development aid coming into the country, but also had implications for the macroeconomy and exchange rate. The depreciation of the currency somewhat cushioned the reduction of aid, however, the environment was still resource-constrained and this was particularly felt in reductions to numbers of teachers recruited year-to-year. The period also included climate shocks and escalating conflict (see Chapter 2).
- In **Nepal**, the Constitution of 2015 and the major *Gorkha* earthquake of the same year, came before the SSDP program period, but the former was a precursor to the "big bang" of federalism in the 2017–18 period in which a raft of new legislation brought in a radical interpretation of decentralization in education provision. After the earthquake many available resources were channeled into reconstruction, including of classrooms and schools, which may have partly constrained the general education budget elsewhere.
- In **Tanzania**, elections in 2015 considerably changed the political landscape, while the introduction of FFBE in 2015/16 brought profound effects on the basic education system as a whole, with enrolment surging at all levels and an ongoing "bulge" effect. The resource constraint led to slower recruitment of teachers at the same time as the enrolment surge, resulting in higher class sizes.

184. The changing context brought a wide set of challenges for RBF and the achievement of the DLIs set out. In all three countries, a degree of fiscal austerity was experienced, with effects on basic education funding. This

was most pronounced in Tanzania, with declining government budget allocations, compared to stagnating funding for education in Nepal and Mozambique. In Tanzania this decline was accompanied and largely driven by a strict cap on teacher recruitment, implemented whilst enrolment was surging due to FFBEF.

185. The rest of this section discusses what happened and did not happen in the three countries, in relation to three somewhat different forms of flexibility/adaptation, namely: a) implementation-related “creative ambiguity” or “de facto” flexibility, and adapting to small implementation bottlenecks; b) flexibility and adaptation to build on early success; c) flexibility and adaptation to take account of large-scale change in context – including with regard to DLI targets. We look at these forms whilst also identifying possible cases of “missing flexibility” – that is, cases where greater flexibility or adaptation might have been useful. This is followed by a discussion of adaptation mechanisms and of the funding implications of the flexibility as it was implemented (or lack thereof) in each country. The section concludes with a discussion of how strategically flexibility and adaptation was applied.

Implementation-related flexibility

186. Implementation-related flexibility was of two types – it could be deliberate adaptation to small implementation bottlenecks, but also allowing creative ambiguity, or implicit, “de facto” flexibility. Table 17 below summarizes what was found in each country in relation to these types of flexibility.

Table 17 Implementation-related flexibility in RBF programs

Adaptation	Mozambique	Nepal	Tanzania
<p>‘Creative ambiguity’ on DLIs</p> <p>‘De facto’ flexibility</p>	<p>Creative ambiguity about way to “count” results on the trained school director evaluated DLI (GPE VT).</p> <p>‘De facto’ flexibility with KfW disbursements made for 2nd tranche whilst not clear on what basis.</p>	<p>Creative ambiguity about “Model School” standards hence number of such schools counted as DLI 5 results, compounded by no IV requirement (ADB).</p> <p>‘De facto’ flexibility granted by one JFP on standards for definition of early grade reading package in DLI 1.</p>	<p>Continuous process aimed to reduce space for creative ambiguity, frequent further specification of GoT measurement/ IV verification protocols (POM, IVA TORs).</p> <p>A few instances of varying DPs’ reactions to DLI results. E.g. DFID & Sida raised disbursements for exceeded learning target (2015); DPs’ rates of disbursement varied for “missed” school score card display target.</p>
<p>Adapting to usually “small” implementation bottlenecks</p>	<p>Adaptations made with TA support, e.g. revising manuals, carrying out training.</p> <p>Lack of flexibility in change in district number not reflected in adapting equitable PTR baseline hence target considered as “missed”.</p>	<p>Priority given to addressing big implementation bottlenecks arising with federalism (see below).</p>	<p>Adapting to change in number of districts.</p> <p>Shift to direct MoF-school grant channeling, with ensuing change in DLI agent (from LGAs to MoE/MoF) in response to bottleneck in CGs reaching schools.</p>

187. Table 17 suggests a strong contrast between Mozambique and Nepal, where “creative ambiguity” was tolerated and “de facto” flexibility applied in several instances, and Tanzania where RBF partners sought to eliminate room for creative ambiguity through a continuous process of further specification of what was measured and how in terms of DLI results. This was associated with, as discussed in Section 3E, a stronger focus in the Tanzanian RBF program design on strengthening government evidence-generating system, and a highly-formalized government reporting and independent verification processes.

188. The Tanzania RBF program was also able to adapt during implementation to an increase in the number of districts in 2016, and to adapt more substantially in response to the “LGA bottleneck” in achieving results on the DLI incentivizing school grants to reach the schools. This led to reformulating the national-level DLI on equitable PTR in 2016 and adopting a system to channel grants to schools directly from the MoF in 2017. In

Mozambique, in contrast, the equitable PTR DLI was not adapted to take account of a similar change in the number of districts affecting the baseline and subsequent performance on the DLI, and as a result, actual success was not recognized and paid against. In this case, it was reportedly agreed between funders and the MoE that it would be too time-consuming and costly to make the change to the baseline; this, however, shows a startling lack of flexibility, while the GPE contend that the change could have been simpler to make.

Adapting to further build on initial success – “Raising the bar”

189. In theory, there could be a case for RBF programs to adapt to further build on initial success in achieving results on DLIs, for instance, if it was felt that initial targets or definitions of results might have been relatively unambitious. In such cases the idea would be to move to “the next step” in the theory of change underpinning the selection of a DLI.

190. There were a few cases of such type of adaptation to “raise the bar” in Tanzania. Notably, the data DLIs were adapted year-on-year, through expanding the scope of school statistics and of the analyses to be included in the annual sector performance review report for these to be considered “satisfactory”. Further adaptation might have been warranted, for example, by orienting these DLIs so as to incentivize the use of the data that was successfully collected, especially by LGAs. This did not occur within the RBF program period, but was recognized as an issue that could be addressed in successor RBF programs. As another case of similar adaptation, it was agreed at redesign stage to include an early grade learning DLI focusing on math skills, additional to the initial early grade reading DLI. This responded to the fact that the reading DLI target had apparently been quite easily achieved in 2015 whilst the learning assessment had also shown that math learning levels stagnated. However, arguably this was ambitious and possibly unwarranted considering that the challenges arising from the FFBEF implementation were by then emerging and known by RBF stakeholders.

191. In contrast, there is one case where RBF funders could have wanted to “raise the bar”, and did not do so – in relation to the “budget framework” financial DLI. This DLI was designed as rewarding government for releasing budget funds on a set of “priority budget lines” which at program design time, formed the BRNEd budget framework. The high-value rewards are proportional to any level of release. This fairly easy-to-reach DLI was left unchanged at redesign, with the “budget framework” of “priority lines” amounting to just 5 percent of the sector budget. Arguably, in light of the well-known rising risk of underfunding of the FFBEF, donors might have wanted to push for a more demanding DLI to try to better mitigate this risk. It is also true, however, that this might have been resisted by the MoF.

192. In Mozambique there was no example of such type of adaptation. For instance, with the DLI on ensuring timely release of the first tranche of the ADE school grant being successfully achieved, there could have been a shift of focus onto the timeliness of the second tranche of school grants. A similar logic could have been applied in shifting from counting school supervision visits to seeking to incentivize improved quality of the supervision, or paying attention to whether teachers were using the training they had received, or building in the DLI target specifications a requirement for qualitative evidence on the functionality of school councils that would allow, for instance, to assess the extent of genuine change in the relationship between parents and school directors. A somewhat similar example relates to the use of Performance-based Allocations (PBAs) at all levels in the education system, which was foreseen in the PFM4R design. Whilst World Bank missions proposed various ways in which to bring PBAs into play, these were not adopted, which suggests reluctance on the side of the MoF vis-à-vis “raising the bar”.

193. That said, it could be considered that such adaptation, which would focus RBF on incentivizing deeper institutional change was not warranted, considering the short lifetime of the RBF programs in Mozambique. In addition, while the PFM4R targets were all “met” with funds disbursed, the progress was challenging in the first and second years of the program, but by the time of the third year it may have been seen as too close to the program end to make any adaptations to program targets. In the case of the data DLIs in Tanzania, where the

RBF program lasted longer, breadth was prioritized over depth. Admittedly it also is harder to design DLIs that adequately capture more qualitative change.

194. In Nepal what was needed was quite different: for many of the SSDP DLIs, federalism as a large-scale contextual change introduced discontinuities in the DLI theories of change, that necessitated a complete revisit on the results that could be targeted and the ways in which results would be sought. By the end of the second year of the SSDP with the changes of federalism the program moved into a “Transition phase”, recognizing the huge shifts occurring to the system. The MTR in early 2019 went beyond the DLI framework but set out challenges across many of the technical areas, and on request (as it was not written into the ToR) the mid-term reviewers made recommendations on changes to DLIs. This led to a major revision to certain but not all DLIs, which took until early 2020 to complete, making DLIs more realistic in the new context, though not “raising the bar”.

Adapting to large-scale change in context

195. In both Nepal and Tanzania, the RBF programs faced large-scale changes to the context that forced some kind of adaptation – in Nepal, the “big bang” of federalism and ensuing rollout of a radically decentralized approach to basic education provision and in Tanzania, the “big bang” implementation of the policy providing fee-free basic education covering twelve years of schooling. This was not the case in Mozambique. In the analysis below, focusing therefore on Nepal and Tanzania, we further distinguish between adaptation through expanding or shifting the focus of the RBF program, and adaptation through redesigning existing DLIs and/or reformulating/adjusting targets. It should first be noted that in both countries, development partners did not consider withdrawing or moving from RBF modalities to traditional financing approaches as a response to these “shocks”. In Nepal the shock was more predictable, and indeed noted (for example, in World Bank programming documentation) as the new Constitution came into force in 2015. Box 16 below briefly presents the process and form of adaptation that occurred for the DLI area that was perhaps most affected by federalism: on governance and fiduciary management (DLI 9).

Box 16 Nepal – Adapting DLI targets to federalism, the case of a governance and fiduciary management DLI

The Nepal SSDP design had a major focus on governance and fiduciary management, with a third of potential RBF disbursements allocated to this area (DLI 9) by two RBF funders, the World Bank and ADB. Over the five years of the program the initial design for the DLI involved 19 targets of which 11 required action at either school or subnational level, and this was increasingly the case from Year 2. For the World Bank, the DLI was conceived as a means of addressing fiduciary risks arising from weaknesses in record keeping, procurement, and financial governance at local levels as well as at the MoE level. The rollout of federalism significantly affected the DLI whilst also heightening the risks, moving from 75 districts as primary cost centers for spending and procurement to 753 LGs as new cost centers, with even more limited financial record keeping and procurement capacity.

The issues with the DLI 9, and more broadly with the SSDP conceived in a centralized context, emerged in Year 2, with a raft of legislation issued in 2017 enacting the radically decentralized approach envisioned in the 2015 Constitution. In 2017/18 a transition plan and road map was developed by a government high-level “SSDP Restructuring Committee”. A third-party review commissioned in view of the SSDP MTR examined the feasibility of SSDP targets in the new context, and also provided recommendations related to adapting various DLI targets. A substantial number of targets were found to be obsolete and needing to be revised, for example targets such as “9.3a Computerized Government Accounting System implemented in all 75 DEOs”, with DEOs no longer relevant. The MTR meeting in May 2019 endorsed the need for these revisions, with new targets added specific to the challenges of the new context, for example “9.4f At least one designated education office appointed in the sanctioned positions in 60% of the LGs.”

The process took further time, involving a number of formal steps such as government requesting the World Bank program restructuring and the approval of the restructured SSDP, as well as negotiations between RBF funders (principally the World Bank and ADB for the DLI 9) and the MoE on the revised targets. The final re-design of World Bank DLIs was complete by early 2020.

196. In Tanzania a number of new DLIs were added at redesign / MTR together with the provision of additional financing (focusing on the development of a range of strategies, on textbook availability, survival and transition rates, SQA rollout and school score improvement, and a very small policy learning DLI); and a few minor targets from the initial design were dropped. The focus on learning was broadly maintained, which was reportedly felt important by senior education officials taken by surprise by the presidential decision to implement the FFBEF in a “big bang” manner. The added DLIs were conceived as responses to a number of risks to quality and learning arising from the implementation of the FFBEF. Whilst these were broadly relevant, a number of considerations might have been warranted, such as an analysis of the relevance of focusing on textbook availability, vs. in-service teacher training to support teachers facing a rising number of very large classes; or the contextual relevance of the highly complex SQA system, for which in addition, achieving the DLI targets would require rushing the rollout process with the risk of insufficient attention being paid to the “quality” of the rollout.

197. The revised targets in Nepal, although more realistic, were nonetheless quite challenging to achieve, considering that a number of governance issues are far from settled. For example, challenging targets such as a performance-based grant system for schools to be put in place, were not removed, but just adapted in terms of the scale of coverage to be reached and perhaps made even more ambitious (target revised to 400 out of 753 LGs by Year 5, compared to previous 7,500 out of 30,000 schools as the original target). Other targets raise mandate issues that are as yet unresolved, for example asking for unaudited financial statements based on expenditure from LGs on SSDP conditional grants, raising the questions of whether the MoE is or will be entitled to get access to this data. In Tanzania there were two notable cases of ambition being lowered as a recognition of larger-than-anticipated effects of the FFBEF, together with the underfunding of the implementation of FFBEF resulting from prioritization decisions made at the highest government level. These cases, concerning the DLIs on equitable teacher deployment and on increasing early grade learning levels, illustrate a number of issues in relation to both the process and the form of flexibility (see Box 17 below).

Box 17 Tanzania – Moving the goal posts on a changed pitch

The DLIs on equitable PTRs are formulated in reference to an acceptable PTR range. In 2017, in line with a revised strategy that RBF had incentivized, it was agreed to raise the upper bound PTR from 50 to 53. This took effect following the review of Year 4. However, with insufficient government recruitment of new teachers for several years, by the time this decision was taken PTRs were already above the revised upper bound in many schools and LGAs. For more than a year, there were discussions between education officials wanting to reformulate the LGA DLI, and development partners wanting to keep the pressure on government to recruit more teachers. Several government redesign proposals were not accepted. In 2020 the government proposed a reformulation that would be equivalent to counting contract teachers to assess the DLI results – a proposal which is reportedly under consideration in DFID’s RBF successor program. On one hand, this might indeed allow deploying a more adequate number of teachers more equitably across schools. On the other hand, it is likely that these teachers would have to be paid locally by LGAs and communities/parents. This, in effect, would also change the targets: PTRs counting only teachers on government payroll could continue to increase.

The DLIs on second year primary pupils’ reading and math skills (DLIs 6) measure results quite narrowly, through two EGRA/ EGMA subcomponents. Results on the DLIs showed progress from 2013 to 2015 but then gains were reversed from 2015 to 2017, with a loss of USD 16.7 million for government. The government, which in parallel to the USAID-funded EGRA/EGMA had implemented their own early grade learning assessment testing for students on the early grade curriculum, challenged the adequacy of the EGRA/EGMA approach as it yielded these adverse results. This at the time was not taken into consideration. In order not to lose the potential RBF rewards, in 2019 government agreed to “add on” the necessary EGRA/EGMA subcomponent measurement protocols to its own assessment approach. Concurrently, at the end of 2019 government and development partners agreed to adjust the DLI targets, recognizing that with the FFBEF the profile of pupils had changed (with a larger number of poorer pupils likely to face more schooling difficulties), resulting in a slower trajectory of improvement.

198. The story of the equitable PTR DLI in Tanzania illustrates an instance of a long-lasting lack of flexibility that was belatedly addressed. Development partners’ initial refusals to consider a redesign of the DLI with a view

to maintaining pressure to “unblock” teacher recruitment was exerted on agents (LGAs) with no power to address this bottleneck. Future RBF programs might use a PTR metric which in effect shifts the costs of new teachers onto LGAs and/or communities. This is likely to raise equity issues related to differential ability of LGAs and communities to pay for contract teachers, and it potentially leaves national level agencies “off the hook”. The challenges show that the very specific definitions of DLI metrics can fundamentally affect the nature of the incentive and the equity and feasibility of delivery.

199. The case of the learning improvement DLIs illustrates the danger of a narrow focus on the metric and targets, that is, here, the two specified EGRA/EGMA subcomponents. This may have diverted attention from the more substantive conversation that was arguably warranted as soon as 2017, as the targeted goal of improved early grade learning seemed to be in jeopardy. Instead, the conversation occurred late and was incomplete. As a result, the next assessment will be using a compromise “hybrid” system, which maintains comparability with earlier assessments of two narrowly defined reading and math skills measured with EGRA/EGMA protocols, at development partners’ insistence. However, this means that the sustainability of the measurement approach is unclear – financially, and because in reality, the MoE and the exam agency continue to favor the curriculum testing approach that they have used to date and they have not bought into the EGRA/EGMA skill-based learning assessment approach.

Adaptation mechanisms – Technical assistance, RBF dialogue

200. This section discusses the extent to which, and in which ways, RBF-associated technical assistance and the RBF-associated dialogue enabled/supported adaptation/flexibility in the program design. Table 18 below summarizes the evidence.

Table 18 Technical assistance (TA), RBF dialogue and adaptation

	Mozambique	Nepal	Tanzania
Adapting in TA/CD	Most flexibility was “de facto”. Funding was rebalanced between windows of PFM4R, away from demand-driven CD which did not work. TA (coaches and facilitators) were recognized as effective.	TA is provided more “externally” (funders for specific DLI projects, overall SSDP TA). TA to adapt DLIs was also mostly provided externally (third-party review of SSDP, bilateral funder/MoE dialogues).	Effective in supporting GoT seeking flexibility on DLIs. Jointly agreed annual work plan left space for pragmatic, responsive adaptation of TA focus – e.g. stronger focus on system development once program implementation infrastructure in place. Planned adaptation towards more systematic CD did not work.
RBF dialogue and adaptation/flexibility	PFM4R dialogue separate from FASE dialogue process and structures. For the GPE VT & KfW, whatever flexibility there was appears to have taken place separately as well.	Third-party consultancy to look at SSDP targets in light of federalism, also looked at DLIs, followed for most affected DLIs (e.g. “governance” DLI area 9) by GoN request to redesign, and WB-led redesign.	EP4R annual review & regular high-level GoT/funders meetings provided a forum for discussions on adaptation. Mid-Term Review important adapting point also worked well.

201. Tanzania stands out with regard to the role of technical assistance in relation to adaptation and flexibility. The RBF-associated technical assistance was instrumental in supporting the government in seeking flexibility, and able to itself show flexibility in its work program to respond to government needs, while retaining the “blessing” of RBF funders. In Mozambique’s PFM4R, coaches and facilitators were reportedly effective and responsive to needs in terms of day-to-day flexibility. However, the program did not function as designed with PBAs or funds for capacity development. There is a lack of documented evidence on why the original design did not work, and what could have been done differently.

202. RBF dialogue processes have been central to the application of flexibility. In each of the three countries, the regular process with annual review meetings and more frequent high-level meetings between senior education officials and development partners provided well-established fora for discussions on adaptation and flexibility, with development partners generally approaching these discussions in a joined-up manner, and government (with technical assistance support) playing a proactive role in the process. This included the government bringing redesign proposals for discussion. In Mozambique, due to its more limited size, RBF was less of a central focus to annual reviews than in the other countries. In Nepal, it can be said that RBF in terms of DLI reporting has come to dominate the annual review processes, in terms of time taken for discussion and the predominance of DLI reporting in the sector review reports (bi-annual with BRMs/JRMs). In Tanzania the annual and RBF reviews were separate processes, though largely drawing on the same evidence basis especially. The RBF review was, in principle, able to draw from the broader sector review which preceded it, although the extent to which this has been the case is unclear.

Flexibility/adaptation in RBF funding

203. There is a question of how adaptation and flexibility as applied in each country played out in relation to the risk to donor funding predictability that the conditional nature of RBF entails. This played out quite differently in the three countries. In Mozambique and Tanzania, in different ways, flexibility as applied facilitated higher disbursement levels than would otherwise have been the case. In Nepal it is too early to judge but it is likely to be following a similar path. Box 18 sets out more detail.

Box 18 Flexibility in RBF programs and financial implications

In Mozambique, for the PFM4R and GPE variable tranche (VT) undisbursed funds could be carried forward where targets were partially achieved, and disbursements could be made in proportion to the level of achievement. Carrying over funds was seen as sustaining the incentive to achieve, and is likely to have helped. These features, together with the “de facto” flexibility allowed in program delivery, meant that with the exception of the equitable PTR DLI for which the government lost funding, all disbursements under PFM4R and GPE VT were made by the end of the programs. For KfW funding, the “de facto” flexibility that was applied for the second funding tranche (see above) was not applied for the first tranche. Reasons for either decision are not fully clear, and RBF disbursements were only two-thirds of the maximum available.

In Tanzania, RBF expanded in response to the FFBEF. Whilst there was no such rule about carrying over funding in case of under- or late-performance and development partners did not always grant the flexibility that government requested or did so late, they appeared to be keen to disburse as much RBF funding as possible. The evidence supporting this includes one case of development partners agreeing to defer for one year the measurement of the SQA rollout DLI; and DFID and Sida disbursing more than allocated for the exceeded learning target in 2015, which “de facto” compensated under-performance and under-disbursements on some other DLIs. Moreover, whilst the World Bank was being seen as less flexible than other RBF funders, in reality the World Bank RBF program was restructured several times. Most notably, an early 2020 restructuring reallocated funds “lost” in earlier years on underperforming DLIs (including the learning DLIs in 2017) onto other DLIs for the remaining year of the program.

In Nepal disbursements could also be rolled over in case of delay in achieving targets. At the end of Year 3 of the SSDP RBF, there was a potential shortfall of USD 36 million against allocation. This was made up by delays to DLIs worth USD 28 million from the World Bank and USD 9 million from the EU, linked to delays in delivery, reporting or verification, as well as some internal delays at the development partners. RBF funding allocated to DLIs affected by the federalism rollout was reallocated to redesigned DLIs, so that full disbursement remains possible, although in some cases this has required development partners to extend the life of their programs with some targets pushed beyond the five-year SSDP period.

204. Tanzania presents an interesting example of a likely political economy effect on flexibility. In early 2020 discussions were ongoing between the MoE and World Bank and DFID to the effect of further reallocating the outstanding undisbursed funding to new DLIs associated with post-COVID school reopening. DLI design work

was fairly advanced but was not taken forward – which may well have been linked to the broader government stand, that COVID was eliminated in Tanzania.⁸³

The strategic use of flexibility

205. With very different program designs, fewer DLIs and a narrower “spread” over the education results chain in Mozambique than in Tanzania and Nepal, as well as implementation contexts including the “game changer” events in Nepal and Tanzania, the type of flexibility that was required and applied also differed among the countries. The main reasons why adaptation and flexibility should have been relevant were therefore quite different. The “forms” of flexibility also differed in relation to whether they might have been or were driven by donors, or sought by government and granted by donors, or if flexibility was applied “de facto”. A summary of the strategic use of flexibility is as follows:

- In **Mozambique** the main need for flexibility was or would have been about evolving toward next steps as “low-hanging fruit” DLI targets were met; that is, building on early success. There was little incentive for government to seek this, and donors did not seek it – possibly thinking the time was too short for this in the prevailing government capacity context. Creative ambiguity and de facto flexibility were also applied in instances.
- In **Tanzania** the main reason for adaptation was the FFBEP compounded by government underfunding. This “game changer” raised challenges to achieve DLI results after initial success in the pre-FFBEP era. In contrast with Mozambique, education officials in the MoE and PO-RALG had strong incentives to seek flexibility. Development partners had to assess the relevance of granting it. With support from technical assistance, government sought this type of flexibility with strong arguments, which generally led to fairly heated discussions. Flexibility was not always timely (for example with regard to learning and PTR DLI design/target), and was more reactive than proactive. Arguably, the prolonged lack of flexibility by development partners in spite of it being repeatedly sought by government was unhelpful; and the adaptation now being considered is problematic (see Box 18 above). Looking forward, in the cases of these two DLIs and many others, there unquestionably is a need for a more systematic process of “tracing” the FFBEP effects across the sector, and reflecting sufficiently frequently on the implications with regard to RBF programs, as the FFBEP bulge travels upward in the system.
- In **Nepal** the “game changer” was federalism. The MTR led to redesign of DLIs in a process aimed to align indicators to the new reality. Stakeholders knew at the outset of the SSDP that major change was going to unfold, though the details and timing were not known. Looking forward, it is likely to take several more years for decentralization to “bed down”. In such an evolving context, “pause, think and revise” moments are likely to be frequently required even now that the redesign process is complete.

206. Importantly, there is a similar story, in all three countries, with regard to RBF programs lacking sufficient attention to policy learning, which would provide a better foundation for adaptation. While the Tanzania’s EP4R included a DLI (“number of new commissions granted to support policy, planning and innovation”), this has not successfully catalyzed a strong process of policy learning. In all three countries, this “gap” can be partly associated with the fact that the dialogue was dominated by discussions of DLI designs, targets and results, at the detriment of attention to causes for the results to be what they were, and whether programs were “on the right track” with regard to longer term goals (see Section 3D and the next section on risks, Section 3G).

⁸³ See Financial Times, 4 May 2020 “Tanzania hiding true number of COVID deaths, opposition says”, David Pilling.

G. Risks and unintended consequences

207. RBF programming and use as a financing modality creates risks that are not generally present in traditional financing modalities.⁸⁴ Risks mostly relate to the incentive effects created by linking high-stakes incentives to results, and can be either mitigated or exacerbated by specific design decisions. Information risks – that is, whether goals have *really been* achieved or only *appear to have been* (sometimes known as *gaming*) – arise because payments on reported results create a potential perverse incentive in reporting. This may include, in the extreme, *cheating*, in which neither the goal nor the metric are achieved despite being reported so; or less extreme, *fudging*, in which the metric may be achieved, but this does not mean the goal has been achieved (perhaps due to imperfections in the metric used to measure). Risks also include “*cherry-picking*”, where incentives lead to suboptimal outcomes in equity. Other behavioral responses to the RBF incentive include *demotivation*, where RBF “crowds out” other incentives and accountability mechanisms, for example, peer-to-peer accountability and intrinsic motivation. The RBF incentive may also lead to *diversion*, in which important goals are deprioritized. Finally, there are risks relating to the flow of finance linked to the RBF mechanism, including non-payment or under-disbursement when results are not or only partially achieved, which may have knock-on effects on the program and broader sector plan implementation and results.

208. This section outlines how RBF risks were assessed and the extent to which they materialized in the RBF programs studied. This is followed by sections successively discussing the different types of risks. Note that financing risks associated to the RBF mechanism are discussed in Section 3C above.

Risk assessment at the outset of RBF programming

209. Development partners going into programs in each of the three countries at the outset assessed a number of risks, often related to government systems, issues of PFM, governance, political economy, and technical capacity; these risks would be present regardless of the use of RBF.⁸⁵ However, the risks that are specific to RBF were not extensively set out in the design phase (for example, in the World Bank’s Project Appraisal Documents, PADs) and have not attracted much attention in reporting and/or partners’ dialogue during implementation. There are a few exceptions, for instance, the PAD for the PFM4R highlighted that for RBF there is a need for robust data systems for credible results, and noted some concerns on the “completeness, timeliness and integrity of information”. Cherry-picking was mentioned as a concern for one LGA DLI in Tanzania, during implementation when the DLI became harder to reach (see below on whether this materialized). Choices on some other DLI designs were risky in terms of likelihood of success and the risks of gaming, and it is unclear whether these risks were fully explored at design, at least according to program documentation available to us.

How risks materialized in RBF programming

210. Table 19 summarizes some of the key risks emerging from our research in the three countries. The wide variety of DLIs and program designs means there have been a range of experiences with regard to the scale and materialization of the potential risks of RBF. Mozambique had relatively few DLIs and they were mostly process or output indicators across the RBF programs. On the other extreme, Nepal’s SSDP RBF program targets 84 results across ten DLI areas over the five years of the program, with a mixture of mainly process and output and some outcome indicators. The Tanzanian RBF design, with sixteen recurrent DLIs from process to outcomes and sophisticated metrics for several quantitative DLIs incentivizing improvement over time, was also complex. An overall consideration, for the Nepal and Tanzania RBF designs, is the risk that the sheer number of DLIs and results targets could be overwhelming and in itself a diversion.

⁸⁴ Although modalities can be expressed more as a spectrum or continuum than having binary distinctions – see Section 3H.

⁸⁵ Risks mentioned included, in Tanzania, the risk associated with low funding for infrastructure; in Nepal, the fiduciary risks associated with the lack of accountability mechanisms and the low capacity in planning, budgeting and monitoring; in Mozambique, the lack of trained personnel at school and school council levels.

Table 19 How RBF risks have materialized across the three countries

Risk	Mozambique	Nepal	Tanzania
Gaming, cheating and fudging	<ul style="list-style-type: none"> - Potential of fudging linked to the nature of simple “tick box” indicators, suggesting success when real change more complex (e.g. school council / accountability shift). - One instance of divergence in IVA and gov reporting (school councils). - Inaccuracies in data systems (e.g. ghost pupils / teachers). Not from RBF but affecting data reliability for some DLIs. 	<ul style="list-style-type: none"> - Unintentional fudging for complex-to-measure DLIs lacking specificity, e.g. OOSC enrolment). - Lack of clear qualitative evaluation on indicators with very simple reporting of complex areas (new scholarship schemes, Model Schools, school performance grants). - Challenges to reporting on a number of output indicators, linked to federalism change to local-level reporting channels and responsibility. 	<ul style="list-style-type: none"> - Fudging risks for SQA DLIs: “number of schools” and informational display insufficient to capture quality of school visits and intended community involvement. Diverging interpretation of metric requirements led to disagreement between GoT and donors and an isolated case of divergence in donor recognition of performance. - Other information risks not detected. The IV did not find systematic / widespread discrepancies between EMIS school statistics and school actual situation, or with financial data.
Cherry-picking	<ul style="list-style-type: none"> - Risk of school supervision to focus on more easily reached schools. Most indicators focused on “complete primary schools”, meaning others not covered (those offering only lower primary). 	<ul style="list-style-type: none"> - Model Schools, only a minority of secondary schools (200 via DLI), receiving disproportionate amount of budget. 	<ul style="list-style-type: none"> - Evidence that exam result-based school ranking associated with weaker students not sitting, to raise exam pass rates. Risk pre-existed RBF but not discussed. - Concerns that LGAs might cherry-pick schools for the PTR DLI did not materialize.
Diversion risks	<ul style="list-style-type: none"> - Simple process DLIs and focus on targets may have diverted from more fundamental and complex issues (e.g. power dynamics school council / director, political appointment of school directors, teachers’ use of training). - Strong focus on four very specific PFM DLIs may have weakened focus on broader PFM. 	<ul style="list-style-type: none"> - Focus on DLIs took precedence over monitoring of other SSDP results within results framework or KPIs, at least within dialogue process (BRM/JRM). - RBF “target fixation” may have deflected partners from focusing on “visionary planning” or addressing the many practical challenges created by federalism. 	<ul style="list-style-type: none"> - Narrow financial DLI design diverted attention from broader issue of GoT basic education underfunding. - Narrow focus on equitable PTR insufficient to address broader issues of low teacher motivation, identified as key bottleneck at design. - RBF dialogue increasingly dominated by focus on DLI results and measurement issues, diverting attention from broader goals.
Motivation risks	<ul style="list-style-type: none"> - Expectations of individuals in the education system to receive incentives / rewards via the PFM4R not met. - Motivation of school directors to improve school quality, reported as low and not addressed by RBF. 	<ul style="list-style-type: none"> - Some frustration in MoE with RBF in relation to federalism, given it was a hard-to-address externality undermining achievements of results. - Other MoE frustration around conditionality linked to loans and the overall proportion of financing coming via RBF, again at a time of major change. 	<ul style="list-style-type: none"> - (High stakes) DLI for increased learning may have led to demotivation with disappointing results, and led to lack of transparency. - Top-down approach to LGAs, reinforcing centralizing trends, may have further weakened local accountability. No evidence but not discussed. - Lack of awareness of intended motivations at the LGA level due to light-touch on communication, possibly stemming from too little TA.

Risks relating to information, the quality of indicators and metrics

211. Information risks may not always be intentional but can relate to the selection and design of indicators, particularly for the more complex DLIs. In Nepal for instance, the Model Schools DLI lacked the necessary specificity, meaning reporting of success was hard to challenge, while as an ADB DLI, no independent verification was required. Nepal’s DLI for out-of-school children (OOSC) did not define “targeted interventions” and data collection and verification were based on questionable methodologies, leading to the risk of reporting apparent

success that may not have been a realistic assessment. It is notable that in Nepal, when the IVA raised questions over the OOSC results, this was not accepted or acted on and the government selected a different IVA in the following year. In Mozambique, the apparent success in meeting most of the DLIs may have in part been that the measures were not good proxies of real change, with indicators allowing for a lack of clarity (or fudging) in reported results. Moreover, it is not clear that the IVA scope allowed for the capability of detecting instances of inflated reporting by school directors or even districts.

212. Information risks were more limited in Tanzania due to a combination of high-quality DFID-procured IVA, and the continued commitment of government officials to take verification and information seriously, along with a successful, timely and apparently genuine strengthening of the EMIS, itself supported by appropriate DLIs, RBF and non-RBF TA, and the intrinsic motivation of a professional IT team in PO-RALG in charge of the EMIS (and other LG information systems). In contrast, in Mozambique, interviews and analysis of the data raised questions over problematic EMIS data, linked to over-inflated reporting of school enrolments and the presence of ghost teachers in the system. This was not a consequence of RBF, but it seems to have been an issue with a lack of attention in recent years, and RBF did not seek to tackle it or improve EMIS more generally.

Risks related to diversion and motivation

213. Considering the complexity of the RBF programming in Tanzania and Nepal, there were instances of diversion risks given the prevailing contexts of finite resources, in terms of finance, attention and capacity. It is not possible to determine whether these diversions negatively affected learning more than would have occurred if other outcomes had been prioritized. However, in Tanzania, this might have been the case in the instance of RBF focusing on textbooks: a different prioritization, for example, could have brought a much-needed focus on in-service teacher training, found to be more effective in the international literature,⁸⁶ and which in Tanzania in a context of large class sizes, would have potentially been more relevant. In Nepal despite the large number of DLI results, very few (even if achieved) would make themselves felt to the average primary school classroom, with many quality interventions focused on various subsets of the country's 30,000 public schools; with limited focus on in-service teacher training or addressing the fundamental challenge as expressed by many sector stakeholders, of the continued dominance of "rote learning" teaching methods.

214. In Tanzania, RBF may have compounded a trend of focusing attention on what is or can be measured, over other reforms. However, positive diversion could be detected as the RBF kept some focus on quality at a time when access became the top political priority. There is evidence in Nepal that DLIs have become the major focus of sector reporting, at least in terms of the biannual BRM/JRM processes. During this research, most stakeholders agreed that the broader range of results framework indicators in the SSDP have in practice been downgraded as priorities if they were not included in the DLIs. In Mozambique, the RBF DLIs were important within the sector review and planning process, but did not obviously divert attention from other priorities, which were also funded under FASE without a link to RBF. This is likely to be linked to the number of DLIs being less than in Nepal or Tanzania.

215. More broadly, a key risk in all three countries was with a focus on the DLIs, losing sight of higher goals, or even of the DLI results themselves due to measurement issues:

- In **Nepal** it was felt that "target fixation" deflected sectoral stakeholders from focusing on results and in particular the changes that would bring about improved teaching and learning. This was not for any particular DLIs but rather the dominance that the full set of DLIs brought to sector dialogue processes and annual planning discussions. Once combined with delays to earlier targets and planning for subsequent years' targets, sector reporting in Year 3 of the program included reviews of progress of around half of the SSDP's 84 DLI results, with significant amounts of detail and the

⁸⁶ Recent evidence in Evans and Acosta (2020)

process of IV to also manage. Together this was a major burden and some senior government officials felt that this meant “visionary planning” was missing.

- In **Mozambique**, targets were sensible as first steps in a process of reform but the focus on DLIs may have reduced space for broader dialogue and setting out a longer-term reform agenda. For the PFM4R, the DLIs targeted important bottlenecks, but did not in themselves change the sectoral PFM issues in service delivery that the program set out to address. For example, achieving the checklist of criteria in establishing school councils was important but did not fundamentally change accountability relationships between the school director and the community.
- In **Tanzania** a similar argument could be formed around a rushed deployment of a new school quality assessment, where thousands of schools were to be assessed under a quite different inspection framework in the first months after the program redesign. More broadly, examples discussed elsewhere in this report (notably, for the learning DLIs) indicate that target fixation was also found, at the detriment of more substantive discussions on what results indicated and what needed to be done to improve them.

216. As noted earlier (Section 3C), whilst RBF financial rewards were an important incentive, the intrinsic motivation of meeting the targets also played a role. Hence the motivation risks were real in cases where targets were “missed”, especially if agents perceived the non-payment “sanction” as unfair because achieving the targets was out of reach. The main motivation risks captured in this research are around complex and hard-to-achieve indicators, and are outlined in Box 19 below. Cases of these risks are found mainly in Nepal and Tanzania, as Mozambique RBF designs were less ambitious. However, on a more operational level, in Mozambique, individual actors within the MoE had expected incentives to reward their effort, which were not forthcoming, likely vetoed by the MoF. Raised expectations clearly had a demotivating effect, though stakeholders interviewed stated clearly that it did not affect their work.

Box 19 Risk to agents’ motivation – Donors failing to share the risks

In Nepal, particularly by Year 3 of the SSDP, a number of targets were extremely difficult to reach in the context of federalism. There was a degree of frustration expressed in interviews with the time taken to apply flexibility related to the more difficult DLIs, and that the effects of federalism were an externality over which the MoE has no control. The ability to change DLI targets is important to overcome potential inertia arising from this frustration. This was taking place at the time of the research, but it is not clear that the MoE and RBF funders fully converged in their appreciation of the risks entailed by the new federal set-up. It is not clear that the scope of work of the MTR consultancy commissioned to review the feasibility of the SSDP targets, including DLI-incentivized ones in the federal context, was commensurate to the importance of this government reform, going well beyond the education sector.

In Tanzania, the EP4R design had high-stakes DLIs to reward improvements in assessed early years learning. But in a context in which adverse results are politically difficult to handle, this may have contributed to demotivation and certainly reduced transparency around the learning assessment results. Moreover, whilst the EP4R covered many areas of the results chain, important omissions around teacher training, a lack of TA time to communicate a map or recipe book for progress to LGAs, and spotlights placed on interventions which were laudable but might not have sustained early learning gains in the midst of a huge enrolment expansion, all of which donors shared responsibility for, may have contributed to a lack of progress and understanding around the learning assessment DLI.

RBF funders arguably could have also “done more” on the underfunding of education, in particular, caps on teacher recruitment. The latter made it impossible for the PTR DLIs, as designed, to perform, hence “unfairly” penalizing LGAs. Yet (see Section 3F), RBF funders refused to adapt the DLI for two years of deteriorating performance. More broadly, underfunding was arguably at a level affecting most DLIs (per pupil spent declined by 25 percent between 2013 and 2019). Donor interaction with high levels of the political leadership was reduced as it became more difficult after 2015, but this also coincided with less donor effort due to a combination of staff departures, increasingly thinly stretched donor staff, and decreased donor risk appetite for political dialogue. This meant that responsibility and risks fell back on the MoE and PO-RALG technical leaders, whose motivation was essential as they worked diligently to make EP4R work, but for whom these big “roadblocks” were “above their pay grade”, and cross-sectoral.

Risks related to equity

217. The risks of RBF in relation to equity have emerged in various forms. Cherry-picking on specific DLIs is only one of these forms of risks and it turned out to be fairly small, compared to others. Risks to equity existed at program design level, at the level of specific DLIs which could reinforce existing inequalities, and in terms of ignoring differential ability of same-type agents to achieve DLI targets, for reasons not (fully) under their control. Overall, these issues meant the DLI frameworks and RBF designs were not addressing a fundamental system challenge. In Tanzania, subsequent donor interventions are picking up on equity concerns more directly, although at the time of writing this report detailed designs were not yet known and it is therefore not clear the extent to which they would do better in this respect.

218. At program level, there was no overall appraisal of the potential effects of the changes incentivized by DLIs on equity as a crosscutting issue, even though in all three countries there were significant disparities in key development indicators, including educational outcomes – across geographical areas, ethnic groups, wealth quintiles and gender, depending on the country. There were DLIs addressing some equity aspects, more so in Nepal than in Tanzania and Mozambique, although in Nepal this somewhat greater focus was undermined by other DLIs designed in such a way they were highly likely to compound pre-existing inequalities. This was also the case in Mozambique. Whereas in Tanzania, the design of several LGA DLIs meant that per student or per school amounts of earned RBF funding varied across LGAs having performed at the same level – potentially reinforcing the known lack of equity in government intergovernmental fiscal transfer system.

219. It is notable that in the context of Mozambique, where sector analyses and discussion recognize the significant inequity between the better resourced south and the central and northern provinces, the only direct focus on equity was through the PTR DLI (number of districts with a PTR for lower primary below 80) – which, while an important measure in itself, distracted attention from the larger inequality between provinces, which had long been a problem and was not addressed during the RBF period. The challenges of school governance are treated uniformly across the country by DLIs, when in reality there are large differences between provinces and districts which may require more targeted approaches. And in a context where gender equity is an enormous challenge, there was no focus given to this through the DLIs. In Tanzania the focus on gender was more explicit, through the requirement of disaggregating all EMIS statistics and NECTA results by gender in the DLI-incentivized sector performance annual report, and the primary-to-secondary transition DLI which specifically targets improvement of girls' transition rates. However, this DLI fails to take account of the fact that in some areas, girls outperformed boys in transitioning.

220. In Mozambique and Nepal there were DLIs that had direct implications in terms of undermining equity/reinforcing inequalities. In Mozambique, three of the four PFM4R DLIs were designed to focus on complete primary schools only, following the logic that these would be better able to adopt the incentivized reforms. However, at the start of the program, this accounted for less than half of primary schools (serving around 60 percent of the school population). As such this could be seen as a direct diversion of resources from more disadvantaged schools – for example, small, incomplete schools in rural areas; it was indeed mentioned as a material risk. A parallel can be drawn with Tanzania, where urban and rural areas have different deployment costs for shifting teachers between schools and achieving the related DLI result payment.

221. In Nepal, OOSC and pro-poor scholarship DLIs targeted greater equity, including gender equity in the scholarship DLI; however they represented a comparatively small, though not negligible financial incentive. However, they appear to be one of the least resourced components in LG conditional grant transfers. In addition, the pro-poor scholarship DLI merely added a scheme, with no attention paid (contrary to the stated intention in JFP documents) to rationalize existing schemes. Counteracting the possible pro-equity effects of these DLIs, a range of DLIs aiming to bring additional resources to subsets of schools are likely to compound preexisting resource disparities across schools, and potentially to concentrate high levels of resources in a small proportion of schools as some DLIs are designed to overlap, as detailed in Box 20.

Box 20 Nepal – Equity and counteracting DLIs

There are two DLIs within the SSDP framework with explicit equity dimensions: DLI 6 on OOSC and DLI 7 with a pro-poor scholarship scheme for secondary level. However, other DLIs within the SSDP also have equity implications. Notably, many interventions do not operate country-wide but specify achievements for subsets of schools or localities. The DLI 6 is an exception in targeting interventions at those areas geographically with most challenges in relation to poverty and educational attainment. Other DLIs set out target numbers of schools to which more resources should be flowing (including mathematics, science and English activity kits and teachers, construction of school blocks, procurement training), with the most significant focusing on 200 secondary schools to become Model Schools (DLI 5). All these DLIs have been responded to, including provision within annual activity plans and LG conditional grants. The total budget allocations for these activities make up a significant proportion of the overall funds particularly for investment, and this makes a difference to the overall equity of sector expenditure.

The Model Schools budget allocation comes to around USD 134,000 per year per selected school, a very large amount when the average non-salary allocation per school is in the region of USD 2,400. Other budget lines are also likely to link to the Model School intervention – including on establishment of libraries, ICT labs, and science labs, all targeting a subset of secondary schools. If these resources also go to Model Schools (which some documentation specifies), they would each get an additional USD 17,000. The DLI protocol also specifies that there should be a cross-over in schools receiving a “full complement of mathematics/science and English subject teachers” with schools provided with activity based kits for math, science and English language; again, potentially further benefiting the same, or a similar set of schools. This provides a potential inequity in the system, in part incentivized by RBF. At the extreme, it is possible that 1,000 of secondary schools (16 percent of all such schools), are capturing approximately 80 percent of the non-salary budget for schools at that level. Thus, the equity risk is that the combination of DLIs are incentivizing the consolidation of resources in a minority of schools, and this does not appear to have been discussed.

222. In Tanzania, in addition to the points raised above, as discussed in Section 3C many of the payments for LGA DLIs were not commensurate to the “size” of the basic education sector (for example, number of schools or of students) and/or of the required effort (for example, much higher costs of teacher transfers in rural vs urban areas). This could be very different across LGAs – so that LGAs performing at the same level and getting the same payment, would get very different amounts per student and/or would be able to fund different levels of activity (for example, different numbers of teachers could be deployed).

223. Very little attention was paid to the equity issues arising from rewarding performance without accounting for differential ability to perform among same-level agents (capacity, own resources, local communities’ perception of education etc.). This had direct financial implications in Tanzania, where RBF rewards accrued to LGAs proportionally to their individual performance. Indeed, a number of local officials interviewed for this study thought that the competition among LGAs should be made fairer, for instance by categorizing LGAs based on characteristics that can either facilitate or constrain capacity to meet targets, and scaling targets or rewards accordingly. The risk entailed was also not monitored, a considerable omission on a half-billion dollar program. That said, even in Mozambique and Nepal where RBF financial incentives do not cascade to subnational agents this issue of differential ability matters, because if left unaddressed, the lesser ability of some of the agents undermines the overall ability to reach the DLI targets – in addition to leaving unaddressed the inequalities in educational outcomes that are likely to be associated with these variations in management performance.

224. One form of equity risk is “cherry picking”. This is the case of risks where it may be easier to meet a target by working with groups that are most likely to hit a target result, or are the least costly to meet the target. The review has not noted significant material cherry picking risks. However, it is noteworthy that such risks did not appear to be discussed at the design phase nor monitored during implementation:

- In Tanzania, as noted in Box 21 below, independent and robust research has identified that students have been pushed out of schools in order to inflate school exam pass rates. This seems to have been driven by pre-existing local reputational incentives rather than RBF payments, but it is noteworthy that this risk was not discussed by RBF stakeholders.

- Also in Tanzania, there was concern among government stakeholders that LGAs might cherry-pick in response to the “equitable PTR DLI” (measured by the within-LGA number of schools within an acceptable PTR range); redeploying teachers to less understaffed schools to bring them into the acceptable range and leaving out most underserved schools. Analysis undertaken internally by the TAS five years into implementation and also under this research suggests LGAs remained true to the intention of the DLI, although the risk could still materialize as LGAs do not have the power to take the actions required to return to better performance on the DLI.
- In Mozambique and Tanzania, it is possible that the DLI for school supervision/quality assessment led to district teams/SQA officials focusing on schools closer to district/LGA headquarters in order to meet the target, missing those schools further away and potentially in most need of support.

Box 21 Risks of incentivizing school performance in Tanzania – cherry picking and equity

The RBFs we examine did not, relative to interventions in other sectors, target local-level facilities themselves (in education, schools) as the unit of incentivization. However, on a small scale, this was pursued in Tanzania. There are several risks around school performance incentivization. Some are not due to RBF or any donor programming – the most obvious being the incentive to over-report enrolment to obtain more teachers and higher capitation grants. Where RBF IV processes are effective and the scope includes a focus on such issues, RBF might reduce risk compared to the education system without donors.

Risks can be ascribed to any intervention associated with applying high-powered incentives to systems that require a more organic form of accountability. Exam performance is one of these, as performance can be artificially boosted, in low-accountability systems, in several ways. In Tanzania, the EP4R included a DLI incentivizing schools to raise their rank in a national exam result-based “league”, quite small-value and targeting just 2 percent of the total number of schools at the top of the performance distribution, but building on historically high levels of attention to exam results in society as a whole. Third-party research by a reputable team (academically led and funded by the RISE program)¹ found schools cherry-picked, “pushing out” weaker students to have more students pass the national exams. They also found that local reputational incentives were the main driver of this behavior, and pre-existed the RBF-associated national school ranking and performance-based school grant system, which in any case was a very small-scale financial scheme. However, this risk was not mentioned and barely discussed by RBF stakeholders, in spite of its plausibility. The MoE was keen to expand the use of performance-based grants to schools, but it was agreed that this would require replacing the exam-based ranking by a more comprehensive measure of school performance, which the SQA system was intended to make possible.

¹ Cilliers et al. (2018)

225. The Tanzania RBF design also included an example of contradicting incentive signals, with a risk of RBF “missing the point”. The program combined a focus on early learning – through high value DLIs representing 9 percent of RBF funding – at the same time as incentivizing schools to increase their end-of-cycle exam results score – through the smaller School Improvement Grant (SIG) DLI just discussed. For early learning to really matter for government, it needs to be seen as important by the society more broadly; yet in Tanzania, as just discussed (and likely more broadly), parents are usually worried about leaving exams, which is too late in their child’s life to worry about whether the teacher is teaching etc. A notable absence from the programming we examined is incentivizing government-to-community communication for the mindset change required to interrogate teacher efforts on early learning, and generally the weak emphasis on “information strategies” (giving information to society on education benefits, costs, and quality), in spite of their high effectiveness according to international literature.⁸⁷

⁸⁷ See for instance, Evans and Acosta (2020), and Angrist et al. (2020). The only example of information strategy incentivized in the RBF programs we studied was the Tanzanian small-scale SIG DLI rewarding schools at the top of or most improving in a “school league” based on exams results. The reliability of the findings was not considered high by RBF funders.

H. Cost-effectiveness

226. The principal-agent theories for RBF center on the idea that results are more likely to be achieved due to the *agent* focusing effort and resource allocation, or the *agent* having greater discretion in how the incentivized results are achieved. RBF could also bring system-level benefits, such as improved results-based planning, better use of evidence in policy-making, improved coordination of institutions over time, and sustained attention to key priorities. This assessment has looked at each of these theories, and Chapter 4 presents some conclusions, with each presenting potential pathways through which RBF may generate *benefits* in terms of results or system-level improvements. In practice, it is difficult to robustly measure these benefits given the complexity of attribution, and both benefits and costs are also complex to quantify or monetize. As such, only a high-level appraisal of benefits and costs is possible in practice. This section sets out some of the main benefits and associated costs of RBF in the three countries assessed.

RBF compared to alternatives

227. There are many possible typologies of aid; the purpose of this assessment is not to review them, however the question of cost-effectiveness is a relative concept and in that sense it is important to note key differences that RBF brings compared to other modalities.⁸⁸ In this section we use “approach” to define a way of structuring the dialogue between donors and government (for example, Sector-Wide Approach, Program-Based Approach), “modality” to define a range of ways of providing aid sharing a number of important characteristics, and “instrument” or “design” when we talk about a specific design or way of using a modality. We consider modalities as including “budget support” (distinguishing between general and sector budget support), stand-alone projects, and pooled/basket funding.⁸⁹ Aid-financed RBF came to be considered as a distinct modality when the World Bank rolled out its PforR “lending instrument”,⁹⁰ although several donor agencies had been using forms of result-based aid financing before that. A simple illustration of some of the key characteristics is shown in Table 20.

Table 20 Characteristics of different aid modalities

Aid Modality	Conditionality	Earmarking	Accountability
General budget support (GBS)	Macro & budget	None or notional	Government systems
Sector budget support (SBS)	Sectoral	Notional	Government systems
Basket arrangements	Sectoral	Real within sector	Blend of government and donor systems
Projects using government systems	Sector (and project)	Real to project	Blend of government and donor systems
Projects using parallel systems	Limited (due to low government ownership)	Total (real)	Donor
Projects through NGOs	Limited (due to low government ownership)	Total (real)	Donor / NGO
Aid-financed RBF programming	Sector results – disbursement linked indicators (DLIs)	Within sector	Blend of government and donor systems. Added function of verification agent.

Source: RBF programming added to the table used in Bandstein (2007). Conditionality applies to government levels. For earmarking, if the control over the external resources is ex post, the term “virtual” or “notional” earmarking is often used. If they are ex ante, earmarking is referred to as “real”.

⁸⁸ See for example Tilley and Tavakoli (2012). When it comes to comparison, our TOR used the term “traditional financing”. In this section we consider it to refer to other forms of financing, perhaps in particular, project and basket fund/pooled financing.

⁸⁹ See Bandstein (2007) and Development Cooperation Ireland and Mokoro (2005).

⁹⁰ Note that the World Bank PforR is a “lending instrument” in the same way as IPF and Development Policy Lending, which are the World Bank’s specific ways of using the RBF, project, and budget support “modalities”.

228. Modalities are not discrete categories, but they range along various continuums, for example, the degree of government autonomy in spending; the extent and scope of conditionality; the rigor of reporting requirements; the degree of alignment with government financial reporting and accountability systems; the presence and forms of additional safeguards; and the degree and role of technical assistance. RBF as a modality uses a form of conditionality which seeks to use government systems for accountability, hence in this manner is closer to budget support than project and basket/pooled funding. The degree of autonomy conferred by RBF is supposed to be greater than with projects, although in practice this may also vary widely – as indeed is the case across the RBF designs studied. Conditionality in the form of indicators has been part of the donor “mix” for some time; the EU has had a long history in using a variable tranche for its contribution towards SBS, including disbursements based on indicators. In that sense, RBF with its emphasis on results/indicators, is not a completely *new* innovation.⁹¹

229. At one extreme of aid modalities is *project aid*, with less use of recipient government systems, using parallel systems and the donor often taking the lead in design and appraisal of the inputs to be provided, while using its own disbursement and accounting procedures. *Budget support* on the other extreme, would totally use government systems with respect to allocation, procurement, accounting and audit processes. While GBS and SBS are non-earmarked support to the partner government budget, the assessments, dialogue and conditionalities are linked either to overall or to sector issues; SBS can also be virtually earmarked to the sector.⁹² Basket arrangements or sector program support, are more holistic in design and use SWAp processes. They have predominated in Mozambique with FASE, as well as in Nepal during the SSRP period. In basket arrangements donors pool their resources using a special account either managed by one of the participating donors or by the respective line ministries. The use of RBF programming can utilize these existing structures or establish separate structures, and therefore may be very similar to basket or sector program financing. The two main exceptions are in the use of DLIs, and the addition of the verification function. In terms of earmarking of expenditure, RBF transfers are fungible so long as they are within the sector, although some exceptions may be applied (for example, via expenditure categories), however this may not differ much from other sector program modalities.

230. In each of the three countries, RBF emerged after periods in which either GBS or SBS predominated, then faded away, at least in relation to supporting the education sector. In all three, RBF did not replace other modalities in the sector (for example, project and pooled funding), as both RBF and non-RBF donors used other means of supporting education sectors alongside their RBF programs. Stand alone projects existed in the three countries, for example for USAID with their early grade reading (EGR) programs, which predominated in their sector engagement even in Nepal where they contributed RBF financing.⁹³ DFID in Tanzania had a stand-alone project of roughly the same size as its RBF program. One notable difference between countries is the existence of a pooled funding instrument in Mozambique (FASE, established in 2003) and Nepal (successive Joint Financing Agreements, JFAs, established concurrently with the RBF design processes), which was not the case in Tanzania where the RBF program was itself the pooling mechanism (for five donors, in its final stage). As a result, donors such as GPE and the World Bank used both RBF programs and non-RBF financing through the basket fund, to support education in Mozambique (in the GPE’s case, entirely through FASE), whereas in Tanzania GPE used RBF and stand-alone projects (managed by Sida).

231. Note that overall, in all three countries, and most countries, de facto there is a “portfolio approach” to aid effectiveness in operation. This means that RBF is used along with other modalities as a mix, and potentially with the view that they may complement each other, based on their comparative advantages (benefits). This is explicitly the case for some donors such as the GPE with their 70 percent fixed tranche and 30 percent variable

⁹¹ The World Bank’s IPF modality for example evolved to one known as “IPF+DLIs”, such as used in the final years of the SSRP period in Nepal, before the PforR modality was developed as a harmonized instrument for sector financing using RBF. However, in a number of cases the World Bank has used RBF instruments then reverted to IPF with DLIs, like in the case of Tanzania, where one of the successor programs of the EP4R is a secondary education IPF with DLIs.

⁹² Bandstein (2007)

⁹³ In Tanzania, Tuosome Pamoja (“Read Together”) (2016-21), USD 67 million program; in Nepal, the EGRP (2015-20), USD 54 million program; both contracted to RTI International. In Mozambique, Vamos Ler! (“Let’s Read”) (2016-21), USD 73.5 million program, contracted to Creative Associates International, Inc. with sub-partners including World Education Inc.

tranche model; and a similar approach by the EU. In each of the three countries assessed a mixture of modalities has been in operation as has been shown.

Benefits of RBF compared to other financing modalities

232. The context in which RBF emerged had similarities and differences across countries. In all three, there was a highly visible “learning crisis”, indicated by relatively recent evidence prior to the initial RBF program design (exam results, the Uwezo learning assessment, and SDI indicators of quality in Tanzania; similarly, SDI quality indicators and the 2016 national learning assessment in Mozambique; and various ongoing NASA assessments in Nepal). Other contextual similarities included that GBS was suspended in 2016 in Mozambique, while SBS and GBS faded away in Tanzania between 2013 and 2015, and in Nepal the SBS structures of the SSRP remained in place but RBF became the main financing modality with 80 percent of donor funding linked to it for the SSDP period.

233. Thus, RBF came at a point in which pre-existing sector financing and dialogue/engagement modalities were not felt to have made sufficient progress (“leading nowhere”, in Tanzania; “inertia” in Mozambique). In Nepal, the SWAp dialogue was valued, but the focus on clearly labeled priorities within the SSDP (with DLIs) was welcome, as a move away from the “jamboree” nature of dialogue in the prior, SSRP era. Tanzania seems to have maintained a meaningful sector dialogue throughout the life of the program. The disbandment of the BRNEd structures post-2016 had both advantages (mainstreaming responsibilities in regular government structures more clearly) and disadvantages (disappearance of higher-level cross-sectoral dialogue on results).⁹⁴ It also ushered in a period in which the RBF dialogue also emerged as quite distinct from the broader sector dialogue, whilst being connected to it through its focus on strengthening government data systems, and in terms of timetabling the sector and RBF annual reviews.

234. As set out in Section 3B on results, an RBF program can be evaluated on the performance of DLIs themselves (‘narrow success’) or on the goals that the DLIs seek to achieve (‘broader success’). As discussed, the two types of success will likely be highly aligned when DLIs are strong measures of sector goals; and/or when DLIs are located on clear pathways towards these goals. The three countries assessed all had some notable successes in terms of DLIs, but the broader success of the programs was in general more limited. The key question is the degree to which RBF itself contributed to the success, and how much of the important results were due to RBF and may have been more difficult to achieve if a different modality had been used.

235. The way that benefits can be viewed may vary depending on the degree of alignment between government and the priorities set out in DLIs. Box 22 below highlights that – as was the case for budget support – RBF “worked better” where there was a reasonable degree of alignment of priorities between government and RBF funders, and when this could be somewhat maintained in spite of large contextual changes. It is noteworthy that whilst this “fits” with the theory that RBF may strengthen alignment of all actors, it raises a question on the explanatory power of the Principal-Agent contractual relationship paradigm, possibly highlighting that, as was the case with budget support conditionality, RBF financial incentives and contractual relationships are unlikely to successfully support change if change is “pushed through”.⁹⁵ Examples in the box (related to Nepal and Tanzania) also illustrate the complexity of maintaining priority alignment throughout time and contextual changes.

⁹⁴ See Todd and Attfield (2017), and Dom et al. (2021)

⁹⁵ There is a vast literature on this point. See, for instance, DANIDA (2014).

Box 22 Priority alignment and its effects on RBF “success”

The nature of alignment between the donor as *principal* and the government as *agent* is central to the success of RBF; and a broad categorization of alignment can be set out as follows for the three countries:¹

Shared priority and strategic approach: The Tanzanian BRNEd/EP4R at the outset had very strong government and donor alignment, with the donors seeking to strengthen the government’s strong results agenda. For this period of strong alignment the program was able to highlight some extremely important priorities, most notably early grade learning, and make it a flagship issue for the sector. While this became more mixed over time, it demonstrated the importance of strategic alignment. The Tanzania case also suggests that the strong initial alignment of priority to learning/quality is likely to have contributed to the fact that this initial priority remained more of a focus for attention than would otherwise have been the case following the shift in the presidential-level “top priority” towards expanding access from 2015/16.

Shared priority and search for solutions: The PFM4R program in Mozambique had a very long gestation period and grew out of the need to build on the government’s “PFM Vision”, with a problem-solving approach. The program brought together the MoF, MoE, and subnational provinces and districts, to enable the specific targeted DLIs to be achieved, with the support of facilitators and coaches as a form of program-embedded technical assistance. In Tanzania the BRN lab that gave birth to the government BRNEd plan and the first design RBF program supporting it was similarly solution-oriented (based on the initial shared recognition of the magnitude of the “learning crisis”) and drawing to do so on a much wider range of stakeholders than usual, features which were seen as a welcome change. The redesign was also solution-oriented but arguably with a less robust identification of the most challenging bottlenecks.

Diverging priority and strategic approach: Nepal’s SSDP period has faced a more challenging context, with the changes of federalism coinciding with the introduction of large-scale RBF. In practice this meant that even if there had been strong strategic alignment of education priorities at the outset in 2016, by 2017/18 (Year 2 of the SSDP), the MoE was placed into a period of tremendous change in civil service structures and in institutional structures with respect to subnational government, which raised a huge challenge as to how to ensure alignment with 753 local governments likely to strongly defend their newly-established, if still vague and untested, constitutional powers with respect to education service delivery.

Low government priority: There were some DLIs that became a lower government priority over time. This was a greater risk in Nepal given that the five-year period incentivized 84 results, much larger than used in Tanzania, even though fewer resources were available. This was linked to the number of donors involved in RBF, but also meant it was more likely some DLIs would be relatively low priorities for government. In cases of low priority, the DLI might still be “achieved” but it was more likely to be done so via a “tick box” or more superficial measure of change, and also have fewer resources allocated.

In Tanzania, the extent to which early learning, with two high value DLIs, is a genuine government priority, is not clear. It certainly was a shared priority at the outset. It appears to have remained so at redesign, when an additional learning DLI was included. However, the intensive efforts focusing on early grade teaching/learning in the 2013–15 period (curriculum reform, country-wide teacher training) were not maintained, whilst they would have been all the more needed with the FFBEp-associated enrolment surge. As assessed through the DLIs, early learning levels plummeted from 2015 to 2017, whilst end-of-cycle exam results continued to rise (after the historical low point of 2012). In a society where exam results have historically mattered greatly to how parents view the value of education, it is unclear that early learning would continue to be a focus of attention if it were not for the fact that it is measured under an RBF contract.

¹ This uses a categorization set out in DANIDA (2014).

236. A key advantage of RBF is that it can be used to extend donor reach into a wide range of policy areas and bundle them together into (a) coherent pathway(s) to the agreed goal, through the agreed range of DLIs across policy areas and levels in the education results chain; and (b) it can reach areas that other instruments cannot. The boundary of the program can be redrawn well outside of cost areas that donors can directly finance – for instance all countries incentivized placement of teachers in some way, and plausibly incentivized the government to act differently. However, as above, alignment remains critical, and where a policy area is politicized and outside (or perceived to be outside) the direct control of ministry bureaucrats, for example curriculum reform

in Nepal, then the RBF is unlikely to produce tangible benefits, but may cause friction, reporting challenges and ultimately non-disbursement or disbursement delay.

237. For donors wishing to incentivize reforms, an advantage of RBF is that these are identified and agreed at the outset, rather than being “developed on the way”, as generally the case with for instance advisory technical assistance projects, with the risk of weak government ownership; and reform implementation through government does not require a large national coverage project which might not be affordable. However, this clearly puts the onus on the RBF program design process, which must enable strong ownership and full understanding of the changes required by the government. This, for instance, was the great value, recognized by all stakeholders, of the “lab” process at the initial design stage of the EP4R in Tanzania, which was not fully replicated at redesign stage. For instance, our review of teacher deployment and PTR DLIs (see Section 3B) found that these relatively ambitious systemic incentives could be effective, including to province level in Mozambique, and to LGA level in Tanzania, and our quantitative review found statistically significant effects in this regard.⁹⁶ The scale of improvements in equity of teacher allocation was, almost by definition, limited by the overall resource envelope for teachers and teacher salaries, and the RBF period was resource constrained for both countries, however the effect was perhaps even more impressive in this context.

238. Comparing programs’ effectiveness in relation to this reform, across the three countries, we found that the progress in Mozambique and Tanzania was potentially linked to the clarity of the indicator and the EMIS system being used for data (making it harder to game in the short-term). This contrasts with the SSDP DLI on teacher deployment in Nepal, which was based on a more opaque measure (government baseline followed by an indicator reducing the “number of teachers to be redeployed”). That said, for clear signals to “work”, the agents targeted by the signal also have to have enough “power” to do what is needed to respond to the signals. In Nepal, another reason why progress was more limited is likely to be that with the shift to a decentralized basic education management set-up, roles and responsibilities with regard to teacher deployment became unclear, and contested – hence making it more challenging to align all actors towards the goal. In Tanzania there was a growing risk that LGAs would stop responding to the signal as the constraints imposed by the centrally-imposed cap on teacher recruitment were increasingly undermining their efforts.

239. The programs brought other tangible benefits. For the PFM4R in Mozambique, a number of the PFM bottlenecks targeted related to more efficient and effective use of (mainly) FASE funds. This included ensuring that school grants got to schools on time (ADE), and giving some focus on accompanying accountability mechanisms that might help to ensure those funds were well-spent as well as to other parts of school management and quality – including the school supervision, and school council DLIs to some extent. A complementary DLI from the GPE on training of school directors was potentially less successful in improving their functioning and accountability for that key role. To the extent that money “got to schools” and accountability mechanisms at least established some additional checks and balances, there is a chance that the “leverage” of the PFM4R funds was higher than the USD 25 million or so spent on the education sector. This was linked to the technical assistance that had a significant role in “delivery” for the program. The KfW attempt to improve progress in classroom construction, and the GPE in-service teacher training system, had less clear benefits, even though DLIs were achieved in the latter case. At the least, there is very little information on how quality has improved and whether the teacher training “cascaded” as intended. Some reviews have made a link between the accountability reforms and improved teacher attendance with evidence from the SDI 2018 survey, however our review found the evidence not to be strong enough to support a clear conclusion on change or contribution in that case.

240. In Tanzania the RBF program brought a significant and likely technically sustainable improvement in the school statistics collection and management system (EMIS). This is important beyond reporting on the program results and directly fed into the production of much more comprehensive and analytical annual sector

⁹⁶ See Final reports for Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021).

performance review reports. Whilst other factors contributed to this success, the incentive effect of the DLIs appear to have been useful, especially at LGA level, whereas at the level of the IT team in charge at national level, their being in charge of earning a substantial RBF amount is likely to have raised their profile – a powerful motivation for a professional team. The DLIs were high value, hence represented a substantial investment, but this approach seems to have worked better than support through year-long, often expensive, technical assistance projects, which rarely achieve sustainability.

241. More broadly, Tanzania's RBF program provided a wide subsector coverage (12-year basic education), in terms of targeted policy areas, at a low "unit cost" (USD spent per student-year over the program six-year period), and with country-wide outreach. It arguably yielded a double dividend – policy influence through the DLI design and agreed targets, and (much needed) infrastructure development, as the major proportion of RBF rewards have been used to this effect (moreover, using an approach reportedly more cost-effective than others in use in the sector). With regard to policy influence, the program plausibly influenced more policy per dollar than earlier Budget Support. However, this cannot be generalized as a characteristic of RBF compared to BS: in this instance, this greater influence of RBF is likely to have derived from context-specific features such as the greater extent of priority alignment /government ownership achieved through the RBF program design, led on the government side by the MoE with high-level support to the BRN process, than with the earlier budget support operations led by the MoF with reportedly weak engagement with sector processes and government actors.⁹⁷

242. The Tanzania RBF experience also suggests the potential benefits of a sector aid "portfolio approach" to aid effectiveness – as stakeholders highlighted the complementarity between RBF working "at head level" on policies and systems, and "on-the-ground" projects that "gave legs" to RBF, with project teams supporting LGAs and teachers in undertaking actions aimed at achieving the results. This was not as systematic as could have been the case, and the upward channeling of grounded lessons from these projects to inform the RBF dialogue and policy/system developments was also not as effective as it could have been, but this experience is still worth noting – including in relation to the role of the RBF program in raising aid effectiveness as a whole, discussed below.

243. In Nepal, the RBF experience has to date been much more challenging than in Tanzania or Mozambique. As set out numerous times above, this was linked with the experience of federalism. The program moved into a Transitional Phase from the second year, which meant the DLI framework was recognized, with only two years gone, to be unrealistic in the new context. The challenges were manifold, from greater turnaround of civil service positions at the center; set up of new institutional structures and reporting requirements; less top-down control on the implementation of policy and strategy at the local and school level; contested and conflicting interpretations in terms of what degree of devolution and decentralization was to be put in place and which government tier had decision-making authority; redesigns to government systems of financial transfers and PFM; and delays in recruiting education officers in LGs meaning the "human infrastructure" of education service delivery was (and still is for many LGs) simply missing for many areas of the country.

244. Development partners have shown flexibility to deal with these complex factors, and while the use of RBF may have made some degree of flexibility possible particularly through the use of rollover provisions, the "target fixation" inherent in the RBF "contract" (DLI framework, Joint Protocol, and bilateral financing agreements) meant that at the very least extensive conversations would be required to negotiate what new targets could and should be. As discussed below, it can be argued that there was an opportunity cost to this process in terms of the time taken for technical and senior officials in dealing with the DLIs at a time that more fundamental shifts were occurring. Changes to the framework, in particular to DLI 9 on governance and fiduciary management, now include more helpful and logical targets for the new federal reality and to help incentivize the speed of transition.

⁹⁷ EC et al. (2013a) and EC et al. (2013b)

Interaction with government funding of education

245. The use of RBF had varying ambitions to address sector financing challenges. Overall, education financing by government was weakly influenced by the shift to RBF as it was driven by major non-RBF factors, with which the programs studied did not engage. Note that other designs engaging more substantively with this issue would have been technically possible. It is not clear that the relevance and political feasibility of doing so was discussed at design stage, in spite of the size and ambition of the programs especially in Nepal and Tanzania.

- The most salient aspect of education financing in **Mozambique** is that government takes on “recurrent” expenditures, while donors take on a high and increasing share of expenditure, classified as “*investment*”, up from 82 percent in 2015 to 94 percent in 2019. In addition, many budget lines called *investment* are actually recurrent expenditures on goods and services – school grants, textbooks, funds for district supervision and teacher training institutions etc. It has been noted for several years that the situation is viewed as unsatisfactory,⁹⁸ however with the suspension of GBS in 2016, sector finance is more constrained compared to access and enrolment than at any point in the past ten to fifteen years, as demonstrated by lower numbers of teachers hired each year, and growing PTRs (see Section 3B).
- In **Tanzania**, the government budget as a whole declined in real terms; and whole-of-government budget priorities shifted towards large non-education infrastructure projects at the same time as fee-free basic education was enacted as a policy. As low budget credibility was a bigger issue than the level of budget allocations, in focusing on release of government resources for the BRNEd priority actions, the narrow financial DLI design made sense at the time of the initial design. However, this was not changed at redesign, when the government basic education underfunding trend was already clear, cutting off the possibility of engaging with this issue. RBF may have contributed to more effective government spending through CG allocations reaching schools more fully and timely, and through more equitable teacher deployment across schools. This is, however, relatively marginal in comparison with a gap between primary and tertiary per student spent increasing from 1:14 to 1:24 in the course of the program period,⁹⁹ and the likelihood that government underfunding partly shifted costs onto parents/communities (and possibly LGAs) – notably with regard to infrastructure and “contract” teachers to mitigate the acute shortfall in government-paid teachers.
- In **Nepal**, a tighter fiscal space followed the 2015 earthquake, with the major damage requiring a major reconstruction effort. This included reconstruction of schools, put under the National Reconstruction Agency rather than the MoE. This is likely to have partly caused the shrinking proportion of the government budget going to basic education. The large-scale use of RBF for SSDP period, also came at the same time as a shrinking contribution in relative and absolute terms from development partners.

246. In each country, RBF is unlikely to have “displaced” government funding, though for different reasons. In Mozambique, RBF as used to date, has been relatively small-scale and involved a minority of donor contributions to FASE. However, it could be said that FASE as a whole is displacing government funding as there is no timetable for gradually augmenting government funding of non-salary costs. In Tanzania, RBF was larger but still small compared to government funding; and as it did not finance critical expenditures, and did not challenge basic education underfunding, it was mainly “invisible” for the MoF. In Nepal, the main issue in the SSDP period has been delays to RBF transfers linked to delays in DLI achievement and ongoing revisions to the DLI framework; this may have constrained sector budgets, although the exact extent of this is unclear. In both

⁹⁸ For example, in Orlowski (2016).

⁹⁹ This is, assuming students’ loans are mostly not recovered. Loan recovery has historically been very low, and whilst by 2019/20 there were signs of improvement in the recent past years, it is unclear the extent to which this will survive the 2020 election period – in which attention to vocal constituencies such as tertiary students’ families is likely to rise.

Tanzania and Nepal, the question of aid (as a whole) ultimately being substituted by government funding was also relevant, considering the trends in government funding just outlined.

247. In terms of aid and sector working, the Tanzania RBF “crowded in” donor funds into a comparatively cost-effective instrument. There was reasonable, even if only partly planned, complementarity with other technical assistance and on-the-ground/hands-on projects, “reasonable quality” RBF dialogue, and RBF enriched the broader sector dialogue through better data availability. In these ways, RBF plausibly raised the cost-effectiveness of the overall basic education aid portfolio. In Nepal, a single DLI framework and long-established SWAp meant close donor working. While it was clear this allowed close working around DLI areas, there were still strong bilateral priorities based on which DLIs the donor “owned”, including in the provision of TA, particularly aimed at achieving DLIs. In Mozambique, as fewer development partners have used RBF, other partners were less involved in the specifics of DLI monitoring, and this and the relatively distinct focus of the three programs, meant DLI reporting was more of a bilateral matter – to the extent that some of the reporting on the German (KfW) use of RBF in construction is at times unclear, in terms of the weighting and reporting of specific indicators.

Costs of RBF

248. Compared to other financing modalities, RBF can have some additional transaction costs; in theory it is the government’s (as the RBF *agent*) responsibility to “show their work” to the donor (as the RBF *principal*), although the importance of a sufficient degree of alignment between principal and agent with regard to RBF “success”, just noted, somewhat nuances this theoretical perspective. In practice, asking for results that cannot be delivered with MoE capacity means donors, inside of or outside of the program boundary, step in to provide support. None of the three countries pursued the agreed goals in isolation. Technical assistance was valuable and necessary in all three countries. This was particularly the case for the PFM4R, in which TA was embedded in a way that made it look more like a standard project than RBF, or rather it was a “hybrid” approach. In Tanzania, government and donors alike agreed that TA was indispensable, and it played a wide range of roles.

Table 21 Costs of RBF across the three countries

Dimension	Mozambique	Nepal (SSDP)	Tanzania
Implementation costs	PFM4R included costs for a capacity development window and performance-based allocations (PBAs) to subnational government as well as schools. The latter was underspent while tracking costs has been difficult.	Implementation structure took on the Joint Financing Agreement approach used in the SSRP period, in that sense RBF within SSDP did not add cost.	Costs to GoT of using its structures for implementation, compensated by “easily won” funding (financial DLI, retention DLIs in early years). Potentially costliest DLI (constrained by national PTR) largely ignored.
Costs of reporting	No high direct administrative costs to donors/government. Little additional cost of RBF dialogue, embedded in sector dialogue already focused on indicators.	Very high in terms of time of reporting, with most of BRM / JRM reporting and work of technical working groups focused on DLIs. Extensive time costs in redesign and negotiation of new DLIs.	Costs to GoT of using its systems, largely offset by their strengthening & ensuing efficiency gains (e.g. “better” school-level allocation of teachers with better stats). Extensive – reporting, negotiations on DLIs, discussions on results, redesign proposals.
Costs of TA and verification to donors	TA & verification costs approx. 17% for PFM4R, not known for GPE VT and KfW.	Low cost of IVA relative to size of SSDP (approx. 0.1% of total RBF). TA facility provided via British Council, cost unknown.	Costs to donors: small direct administrative costs, TA & IV accounted for 2% of total RBF.

249. A very high-level summary of costs is provided in Table 21. There was a cost in terms of the “time spent” of government officials to the use of RBF. This particularly came up in the case of Nepal, in which the majority

of sector reporting is now focused on DLI reporting, to some extent “crowding out” discussion of other areas from the sector results framework and KPIs which are also reported. Many sector stakeholders set this position out to us; they noted that this includes sector and technical working groups, the annual budget review meeting (BRM) and joint review meeting (JRM) processes. The DLI reporting, planning of the sector dialogue, and organization of the verification process, as well as ongoing renegotiation of DLIs, is a very time-consuming process, and according to senior individuals who had experience of the previous SSRP system, was the major additional cost of RBF in the country.

250. Whilst budget support in all three countries entailed significant reporting and dialogue costs to satisfy donors about government performance and secure approval of successive tranches of budget support, even for so-called “fixed tranches”, one clear additional cost of RBF which is not present for other approaches is the addition of a “compartmentalized” verification process. The cost of verification can be viewed narrowly as the cost of the contract with the IVA itself, or also looked at in terms of its cost-effectiveness; that is, a low-cost verification option might save resources but then not adequately carry out the required function. A higher cost option might also add useful information and evaluative content beyond the specific verification function. In practice, across the three countries the IVA costs appear to have been small, with the greatest effectiveness in Tanzania. As set out in Section 3E, the approaches taken to verification, including the type of organization contracted, varied across the three countries, and experience varied significantly (use of the National Audit Institution in Mozambique, two universities in Nepal, and a consultancy contracted via DFID in Tanzania).

Overall cost-effectiveness

251. A pre-condition for cost-effectiveness, by definition, is effectiveness. As discussed above, RBF gives development partners a means to apply conditionality to incentivize governments to focus on higher-value policy areas – that should, however, also be seen as important by government, at least to some extent.

252. In **Tanzania**, many successes were clear, this includes EMIS improvement, more equitable teacher deployment especially across schools, some improvement in retention and survival (though might have been temporary), and some attention to learning and education quality which was sustained, although shallow towards the end. The introduction of SQA is a mixed story; it has potential but not if the approach is rushed and insufficiently supported, which has been the case thus far. Plausible claims can be made about significant contribution by RBF to some important and desirable outcomes for the sector. In Tanzania overall, there can be said to have been reasonable benefits for relatively low costs, compared to other contemporary and recent aid instruments (projects, budget support). However, as discussed elsewhere in this chapter, it is plausible that, for some DLIs, targeting other priorities might have been more cost-effective than those they targeted. Also, achievements resulted from a hybrid program and aid portfolio; it is not plausible that the same results would have been achieved with “pure” RBF alone. Moreover, even success stories (for example, EMIS) were either “work-in-progress”, and most if not all the gains achieved are unlikely to be maintained if the trend of underfunding of basic education is not reverted (for example, equitable teacher deployment, retention and transition).

253. In **Nepal**, a greater focus on process and output indicators meant that development partners have pursued the “means” to achieve sector outcome goals, and only a minority of the RBF allocated was to the goals themselves – that is, Year 5 (final year) targets on learning, and enrolment. An exception was earlier targeting of reductions to OOSC in targeted districts. While the DLI targets for this indicator have been achieved, questionable data was used, and quantitative analysis for this assessment could not find the enrolment effects when analyzing district-level EMIS data.¹⁰⁰ In terms of the effectiveness of the various outputs, as the program period is ongoing it is still early to say, however delays to various targets show that many may have been limited.

¹⁰⁰ See Holden and Chapagain (2021)

Thus, even where important and evidence-based strategies have been pursued, for example multilingual education, and monitoring of time spent teaching, implementation challenges have meant progress has been slow, while qualitative assessment of progress has been lacking.¹⁰¹ Other areas of DLI focus are much less clear in terms of being evidence-based approaches to lead to education outcomes, for example Model Schools, which have taken a very large share of non-salary sector expenditure, with a lack of clarity over definitions and no evaluation of what has been achieved. Further, it is unclear whether teacher deployment and equity in the system was effectively addressed by the DLI; and PFM reforms targeted in the original SSDP DLI framework have had to be fundamentally adjusted to fit with the new reality under federalism. Overall it is very hard to say at this stage that the DLI focus of RBF has been cost-effective, given the major time burden set out above, and challenging progress, with a major opportunity cost in terms of reducing the focus on adapting to federalism and the new context.

254. In **Mozambique**, the use of RBF was both more limited and more targeted than the other two countries. As above, the PFM4R had some important successes, and while some stakeholders characterized these as “low hanging fruit”, others pointed out that this was only the case in retrospect, and the reforms at the time were important and difficult to achieve, particularly in getting the MoE and MoF to work together. The PFM4R was an unusual program in this assessment as it ultimately originated outside the education sector, out of broader work the World Bank had done on PFM, with the program itself also covering health, and payments made outside of FASE, the pooled fund for the education sector. This design meant that it was very clearly focused on some specific bottlenecks, but also meant that it took a hands-on approach through the hiring of facilitators and coaches. The use of RBF by the GPE did not have this approach, but was well-aligned with some reforms the MoE was seeking to make. It was a short program with only two main years of incentivized results, and as such perhaps it not surprising that the changes occurred may not have been qualitatively robust, however they may have been important “stepping stones” in terms of IFPs having a clearer role in in-service (as opposed to just pre-service) training of teachers and school directors. The use of RBF in construction by Germany (KfW) was on the surface less successful, however we based our review on much less available documentation or interviews for this use of RBF, and it is also a complicated area; an important caveat. Overall, the limited use of RBF had some useful benefits, and the major challenge may be in the sustainability of the PFM focus, with the DLIs reached not clearly built on to move to the next stage. This may limit the ability of these reforms to lead to broader system changes or sector outcomes.

255. Finally, taking **equity** as a component of effectiveness, in subsectors characterized by high funding and outcome inequality as in all three countries, RBF did not, or only very marginally, engage with this issue (as discussed in Section 3G). In Mozambique, there was no evidence that RBF has begun to address long-term structural inequalities in education finance and outcomes particularly between provinces and the South and North. In Nepal there was one ambitious DLI aimed at equity via bringing OOSC to school, which led to a focus initially on the most marginalized districts as measured by an “Equity Index”. In practice, this did not lead to any significant change in resource allocation, most likely reflective of it being more of a donor priority than for government. As set out in Section 3G there were also various risks around equity in which the DLI frameworks may have led to prioritization of resources for certain subgroups that may already have been better off.

¹⁰¹ These are clear examples of high-benefit areas according to the rigorous literature – see for example the recent Evans and Acosta (2020) study.

4. Conclusions

256. In this chapter, we draw on the study findings to provide some conclusions on how RBF in practice aligned to the various theories for how it can work. This takes us back to the central question of this research: "To what extent, in what ways and under which conditions has RBF contributed to strengthen education systems to deliver results?" To refresh, the theories for RBF can be classified as:

1. *steroids* (RBF induces more effort and/or resources to achieve specified results);
2. *signposting* (RBF focuses effort and/or resources to achieve specified results and not others);
3. *autonomy/innovation* (RBF provides greater discretion for how to achieve specified results);
4. *labeling success* (RBF induces better results-based planning);
5. *sharpening minds* (RBF improves evidence and evidence-based policy);
6. *aligning all actors* (RBF improves coordination between key institutions); and
7. *sustaining attention* (RBF maintains contractual definitions of results during periods of change).

257. A point to note is that the seven RBF theories are interlinked. The *steroids* and *signposting* theories are two sides of the same coin – that is, if RBF is effective in prioritizing certain results, other results will be de-prioritized in the agents' allocation of attention, efforts and resources, which are always finite. Another cluster of closely linked theories relates to RBF effectiveness in improving result-based planning (*labeling success*) and in improving evidence and evidence-based policy (*sharpening minds*). To plan for results, one first needs to pay attention to results; for both to be done meaningfully, requires good evidence; and ideally, there should be a continuous iterative loop between results-based planning and monitoring, and evidence-based policy. There can be a tension between effectiveness in improving result-based planning (*labeling success*) and *innovation* through greater discretion, if results "labeled" to be achieved are detailed operational targets. There can also be a tension between *aligning all actors sustaining attention*, and *innovation* by some of the actors, particularly when priorities and the context changes with the need to adapt. It is important to keep these links and possible trade-offs in mind. These emerge throughout the sections below focusing on the seven individual RBF theories.

258. RBF experiences assessed for this report have included in Tanzania through the implementation of the BRNEd/EP4R since 2014; in Nepal through the late SSRP period (2013–15), and extensively in the SSDP period since 2016; and in Mozambique through the PFM4R (2014-18), the GPE additional financing (2015-17) and the German (KfW) (2013–16) use of RBF in construction. In all of this RBF programming, the financial incentive sat together with a wide range of other support, including various forms of technical assistance. As such, it is difficult to untangle the effect of the RBF mechanism from the broader support provided, and in practice, the two can be viewed as complementary modalities as this section sets out.

259. This study set out to find how RBF had been applied, whether by donors as an aid instrument, or by government within its own systems. The main use of RBF has been as a donor financing modality. Tanzania had the most sophisticated design with respect to subnational payments, representing a third of the RBF funding available and actually paid (see Box 10 in Section 3C); in Mozambique and Nepal, financial and other cascades were also envisaged but were not designed or implemented to the same degree of rigor as in Tanzania (see Section 3C and 3D). Overall, whilst subnational RBF was used in Tanzania, it cannot be considered as "government RBF" as it was clearly funded and designed as part of donor aid, through fund flows separate from the government's. In this assessment, therefore, RBF as a means of providing external finance has been the main focus of analysis.

260. The rest of this section looks at each of the theories, reflecting on the role of RBF and how it has worked, drawing on the qualitative and quantitative evidence and findings set out in this report.

4.1 Steroids

“Steroids” Theory claim:

RBF induces more effort (and/or resource) to achieve specified results. A goal is more likely to be achieved if a payment is made on its achievement, because more effort will be exerted to achieve the goal by agents.

Potential risks:

*Gaming risks relating to the RBF metric, such as fudging where success is in appearance only.
Demotivation by crowding out other accountability mechanisms.*

Conclusions:

- There was evidence of increased effort and resource to results linked to RBF incentives in each of the three countries. The degree of performance on DLIs was not found to be linked to the size of the associated RBF payment.
- There were instances in which DLIs aimed to incentivize additional financing for education from the MoF. It was less clear that RBF was effective in this respect.
- Tanzania’s EP4R included subnational RBF payments in its design. Local government (LGAs) recognized this incentive, including to improve survival and transition rates. Mozambique’s PFM4R also included subnational payments in design but these did not function as intended or align to existing subnational PFM systems. In Nepal, subnational government was reorganized during the program period, and the PFM system fundamentally reshaped. This complicated the ability of the MoE to implement changes represented by DLIs.
- Tanzania’s program aimed furthest up the education results chain and could be described as more transformational in its ambitions. In Nepal, and in particular Mozambique, the RBF concentrated on lower-level results. While Tanzania was comparatively “hands-off” in relation to how results would be reached, in Nepal and Mozambique, DLI results had more linear sequencing, similar to project results chains (e.g. process to output to outcome).
- In Tanzania, the steroids effect of RBF funding was present but the effectiveness of the learning focus was partly offset by the Presidential drive to increase access through the FFBEP (without commensurate resources), and a focus on investing in school infrastructure.
- Under the Nepal SSDP a steroids effect was clearly present for the MoE and CEHRD. The ability to achieve DLIs seemed to be constrained by their large number and complexity, and by the changes entailed by the federal transition, particularly from 2017. From that point, DLIs involving local-level delivery faced significant challenges.
- In Mozambique, foundational DLIs were well understood and aligned with MoE objectives, though it has been hard to disentangle a steroids effect from accompanying elements of the PFM4R design (technical support from coaches and facilitators). The GPE RBF did not have this support but appeared to have a steroid effect. The German (KfW) RBF was less successful in speeding up classroom construction.

Cases of risks materializing:

- The selection of simple indicators to measure complex change at times allowed reporting of apparent success when it looks unlikely that the underlying goals had been achieved. There were limited attempts to use a range of evidence to triangulate DLI reporting and offset this risk.
- There was often an absence of strong qualitative understanding of results and the link to impact. For example, the “last mile” of school grants being spent well at schools, or changes to skills and motivation of teachers and school directors as a result of their training.

261. The central claim of RBF through the principal-agent framework, is that a goal is more likely to be achieved if a payment is made on its achievement. The payment *aligns* the incentives faced by the agent with the interests of the principal making the payment. The theory is that more effort or resource may then be put in to achieve the goal. As the agent is the “government”, with a complex set of institutions and civil servants working within it, the incentive is not generated by the idea of personal gain, but by organizations’ collective goals and the potential reward to the sector. As set out in Chapter 1, the principal-agent paradigm itself is likely to be a simplification. In particular, as implemented, donor-funded RBF programs involve mutual accountability, both donors and government want to achieve results, and donors also fund technical assistance to help in achievement alongside any RBF.

262. As Section 3A set out, the designs of the RBF programs studied were very different, including with regard to their degree of ambition in the levels of targeted DLI results within the education results chain, and in terms of coverage of the chain. In all cases, the results targeted would require a wide range of system actors (or agents) to be reached. Our analysis identified the need to consider how specific DLIs involve different levels of

social, political and technical complexity, which broadly speaking (though not universally) increases when targeting behavior change among a wider range of actors (central agencies, local and school-level actors, society/parents). This requires an understanding of the institutional set-up within the education system and broader government structures, and how incentives work or do not work to align with existing accountability relationships. “Steroids” effects may work at the individual DLI level, or more broadly. A DLI may have higher or lower financial values, and it may have targets that are or are not well-communicated and understood. More broadly, agents may see value and take pride in achieving results, if there are reputational effects from being seen as an “achiever”, or alternatively see risks in not achieving the targets. For the “steroids” effect to work, it is important to identify who should be incentivized, and how to align multiple agents, and if it is not feasible or desirable to directly financially incentivize the most important agent in achieving a given result, how will agents that get the financial reward “transmit the pressure” (delegate) to those responsible.

263. In Tanzania, good alignment between donors and government goals and collaborative agreement on targets was important, alongside the financial incentive. While goals became more divergent with the FFBE decision made at the highest level in government just one year after the program start, the EP4R helped sector senior officials to maintain focus on learning goals through the DLI-associated incentive effect and funding, and accompanying donor coordination. The “steroids” effect of RBF funding acted at program level rather than in relation to individual DLIs; and was accompanied by an incentive for agents to be seen as achievers, including for local governments. However, given the focus on learning, the cascade and nature of incentives targeting national and local government did not account for the range, number and roles of other program agents (for example, households “left out”, teachers only indirectly incentivized). In at least one case, attempts to induce a “steroids” effect were, in the later years of the program, no longer correctly targeted (LGAs instead of the MoF/PO-PSM in relation to teacher deployment). The post-2015 presidential “drive for results”, with sanctions taken on adverse results, confounded the EP4R “steroids” effect in two ways: a) it meant that EP4R was not the only (and arguably not the largest) factor incentivizing sector actors to “deliver”; and b) it “curbed” efforts so that, whilst attention to EP4R goals continued, sector actors also paid attention to deliver what the President wanted (school infrastructure).

264. Under the Nepal SSDP the “steroid” effect was most clearly present at the central level, in that the MoE and CEHRD have put effort and resources into the delivery and achievement of DLIs; and as long as education provision was ensured through deconcentrated management, the steroids effect was possible to transmit through the prevailing central command and control system. However, the effort to achieve DLIs has been constrained by both the large number and complexity of DLIs, as well as in large part the changes entailed by federalism, particularly from 2017, which saw LGs as new actors over which the former command and control system was not applicable. In this context, “socially complex” tasks involving a large number of actors, especially any task involving local-level delivery, faced major challenges (for instance, the fiduciary and governance reforms under DLI 9 had to be redesigned; the new setup also raised complex and thorny issues of contested mandates with regard to teacher deployment under DLI 3). The mechanism relied upon to obtain implementation by the LGs (tasks with earmarked funding through government Conditional Grant, and a Program Implementation Manual specifying how it should be spent) has been hampered by unclear reporting systems, staffing challenges at LG level, and “political sensitivities” and a contested space in terms of the degree of autonomy for LGs post-federalism, for example including curriculum reform (DLI 2).

265. In Mozambique, particularly under the PFM4R, relatively simple, foundational DLIs were well understood by and well aligned with central and to some extent provincial and district officials and their priorities. Target results were not particularly socially or technically complex, but they did bring together central actors to address long-standing bottlenecks, and incentivized provincial and district effort for DLIs such as school grant timeliness and school supervision. The DLIs and reward to sector were well known, although there was some confusion on how PBAs would be used (cascaded financial incentives built in to the PFM4R funding model) and they were not used as intended at the outset of the program. It is hard to disentangle the “steroid” effect of RBF incentives from the accompanying elements of the PFM4R program (support from coaches/facilitators), however the GPE

DLIs without the same degree of support also appeared to have an effect. The top down mechanism from the MoE to provinces and districts provided a direct means for the central level to deliver RBF programming, however, the approach may have made it more difficult to build a strong qualitative understanding of results. This includes the quality rather than quantity of primary school supervision, the “last mile” of school grants being spent well at schools, and changes to skills and motivation of teachers and school directors as a result of their training. The PTR DLI while not formally achieved, saw statistically significant progress according to analysis carried out for our research, suggesting clear messages were passed to province level. The German (KfW) use of RBF was less successful in speeding up classroom construction.

4.2 Signposting

“Signposting” Theory claim:

RBF focuses effort (and/or resource) to achieve specified results. Within finite limits to efforts/resources, allocations of effort/resource to specified results reduce those to less important areas.

Potential risks:

Diversion risks may be present if there is a suboptimal resource allocation, i.e. with important areas deprioritized because of the RBF incentive.

Cherry-picking risks may also be created, in which less costly or politically favored subgroups or regions see greater resources, in part because they are more likely to reach a target.

Conclusions:

- DLIs were effective in providing signals to MoE civil servants responsible for their delivery, and in most cases the level or value of the DLI was not an important factor in this. DLIs became a central focus of annual sector reporting and dialogue for the program period.
- Tanzania had at the outset a “design lab”, with analysis and intensive discussions between government and donors, which ensured strong alignment. The high-level ownership of the BRN initiative varied over time with new political leadership. Learning goals had to compete with the legitimate but under-resourced priority given to expanding access with the FFBEF. During the 2017 redesign, there was less in-depth analysis and discussion of trade-offs than in the initial design process. Issues not addressed included important quality-focused measures such as in-service teacher training and regional equity.
- In Mozambique, the process of program development and negotiation was very focused. The MoE was effective at ensuring strong alignment with its own priorities. The PFM4R evolved out of government and World Bank focus on PFM at sectoral and subnational level and DLIs provided a strong signposting effect on specific priorities, including for district and provincial education staff. In the case of the GPE DLIs there was a signposting effect for reducing PTRs (though not sustained). The MoE’s original goal with PTRs was to secure additional financing for new teachers from the MoF, but this was not achieved. Other GPE DLIs were likely already government priorities and may have happened without RBF.
- In Nepal, the large-scale of the SSDP RBF included DLIs in multiple thematic areas, with an array of new initiatives and priorities. Alignment or agreement with the specified goals of the DLIs by the MoE and its agencies at times was variable. Despite this, a strong signpost effect was present but with strain on the ability to achieve many DLIs in the context of the rapid and profound changes faced by government. Much of the focus was on reporting and for many process DLIs it is unclear that targeted reforms took place even when DLIs were achieved (for example targeted improvements to geographical equity in teacher allocation).
- Signposting was supplemented by technical assistance funded by development partners and this was often essential to achieve results.

Cases of risks materializing:

- Results within sector frameworks that were not financially incentivized via DLIs in many cases had less focus. However, MoEs often pursued their own priorities even if they were not DLIs.
- The time dedicated to DLI reporting crowded out attention to other results.
- Equity in resource allocation and outcomes was an issue that DLIs did not systematically address. On occasions, DLIs may have had an unintended consequence of incentivizing more inequitable resource allocation.

266. Closely related to the *steroids* theory is the theory that RBF creates an incentive to *focus* effort to the specified results via a *signposting* effect, and that this may prioritize resource allocation over other areas, a reflection that any allocation of resource or time has an opportunity cost. The essential point of an RBF mechanism is to incentivize some results and not others, but there are risks to doing so. Most notably other important goals or results could become deprioritized in a suboptimal manner.¹⁰² In addition, depending on how results are specified there are equity risks in how results for different subgroups are incentivized. In assessing why certain results were chosen it is important to assess whether discussions on DLI selection include considerations of trade-offs – for example, why targeting textbook availability and not teacher in-service training, and how much “localized” evidence supporting these considerations was used. Depending on program design, other results can also be incentivized by non-monetized means, for example being part of a results framework –

¹⁰² Although given the complexity there is unlikely ever to be an objective answer to what is optimal or suboptimal.

how different results are prioritized may link to the program's planning, and support if needed (under/outside the program).

267. In Tanzania, under the BRNEd, discussion about trade-offs were well resourced at the initial design stage, with an intensive six-week broad-based stakeholders' "design lab", with in-depth, readily available evidence and analyses on the main issues affecting the sector. As noted above (steroids), this gave a strong focus on education quality. When access became the politically determined first order priority in the FFBEp era, we could say that quality then became a signposted goal of RBF. With goals not fully aligned, a number of important issues were not addressed in the redesigned DLI framework, and later on were also *de facto* ignored or paid scant attention to. However, our analysis suggests that this was also due to the fact that the redesign process, whilst it could draw on the better EMIS data available by then, missed the kind of in-depth, explicit trade-off analysis and discussion that the initial design process had enabled. Issues not addressed included important quality-focused measures (for example, in-service teacher training, where not much happened after the initial burst of early grade learning strengthening activities in 2015); and equity, with which the program minimally engaged, in spite of evidence of serious imbalances in most education results. Financial DLIs signposting a few lines in the government's education budget were insufficient, without DLIs targeting more of the government's basic education budget.

268. In Mozambique, the programs had different operational models. In the case of the PFM4R, DLIs signposted some key bottlenecks for a limited period, which had evolved and were designed as part of a broader government and World Bank focus on PFM at sectoral and subnational level. DLIs provided a very strong signposting effect on some specific priorities, including for district and provincial education staff. The PFM4R program's use of coaches and facilitators in large part put in place a team to deliver the changes, together with a facilitation role, particularly in bringing together the MoF and MoE, and coordination down to subnational levels. In the case of the GPE DLIs there was a signposting effect but generally for relatively simple quantity indicators rather than broader or deeper system changes addressing capacity gaps. However, using teacher training institutes (IFPs) for in-service teacher training and school director training (covering three out of four of the GPE DLIs) was part of the 2015 MoE operational plan, and associated achievements may have happened without RBF. For the most part, the GPE DLIs were based on "numbers trained", with targets passed directly from the MoE to IFPs. But they did not provide any strong qualitative measurement of school-level change, for example, teacher peer training, or school director practice. For the GPE DLI on PTRs aimed at provinces with districts with very high primary school PTRs, there appears to have been a signposting effect, although it was not sustained. The MoE's original goal with PTRs was to secure additional financing for new teachers from the MoF, but this was not achieved.

269. In Nepal's SSDP period, the signpost effect has been strong across the large number of DLIs in terms of attention and focus, and in a context of rapid and profound changes faced by central officials (mandate, civil service, PFM changes following the "big bang" federalism; and more recently the effects of COVID). It was very clear that the DLI framework has provided the main definition and therefore signposting of results for the years of the SSDP in terms of reporting and dialogue, and while there is a broader results framework and key performance indicators, both are said by most stakeholders to be secondary to the DLI framework. This in part is logical as the DLI framework was designed to be holistic; however, there are many areas of significant expenditure by government not covered by the DLI framework which include in ECED, mid-day meals, a wide range of low-value scholarships, WASH facilities, and textbooks. All are important policy areas, yet with the exception of ECED, have had significantly less focus than DLIs in the biannual BRM/JRM processes and reporting during the SSDP years, and there is therefore a risk that the "signposting" effect has been too strong in Nepal to date.

270. Guidance on RBF has suggested that *"the more important the DLI, the more money"*,¹⁰³ with the idea that more money will prompt more efforts. Across the three countries, it is not clear that this was consistently applied, nor that it mattered. First, with the exception of the LGA DLIs in Tanzania, when incentives do not cascade (like in Nepal and Mozambique, and in Tanzania for the learning DLIs), it is unlikely that the financial incentive gets passed down clearly enough for it to work. More fundamentally, in Tanzania it did not seem that the financial value of the DLI mattered in determining the level of efforts made. What seemed to matter more, in terms of attention – a prerequisite but not the same as effort – was the gap between results and targets. The most financially substantial DLIs (financial DLIs 2) possibly by design and experience did not have an incentive effect on the actor responsible to achieve them (MoF); and there is a sense that the MoE and PO-RALG rarely exerted much effort to try and get the MoF to "do more". Moreover, there was no convergence, even among donors, on what the most important DLIs were – and this, with successive additions of funding by donors, meant that the relative financial importance of DLIs changed every year. This assessment can also differ between government and donors – and there were clear examples of this in Nepal.

271. In Nepal, with six joint financing partners (JFPs) using DLIs, there were a range of values, with the few outcome indicators having some of the lowest allocations by value. The values of DLIs varied, depending on which donors focused on which DLIs and how much funding they had, as they tended to mostly split their respective contribution relatively evenly across the DLIs of interest to them (ranging from USD 1 million per result for Finland and up to USD 7 million per result for the World Bank). As a result, DLIs focused on by the World Bank and ADB, including governance and fiduciary management (DLI 9) and a range of targets regarding teachers (DLI 3), were the largest. They undeniably focused on important areas, but it is not clear that the framework as a whole followed the logic of allocating higher rewards to "more important" DLIs. Either way there was not full convergence on the "importance" of DLIs. Some DLIs were actively not viewed as valid priorities by government stakeholders, for example curriculum reform seen as "political" and outside of bureaucratic control. There were also relatively small value DLIs (such as Model Schools) that were heavily funded by government whereas other similar value DLIs (for instance on OOSC enrolment) received less government attention and/or resource.

272. While recognizing that RBF programs cannot address "everything" and should be seen in the context of a broader portfolio of support to the sector plan and goals, it was notable from our research that equity was not systematically addressed in any of the countries and programs. This seems to be a case of "diversion risk" associated to strong signposting effects that did not consider equity, or considered it insufficiently or in contradicting ways – an "omission" all the more significant considering that in the contexts studied (and in most low and lower middle-income countries), inequitable resource allocation and education outcomes are a major issue; and that, especially in Nepal and Tanzania, the RBF programs were a proportionally large injection of aid financing in the sector. Whilst equity is also often a highly sensitive issue, with political economy dimensions related to identities, geographical specificities, and political alignment, there could at the very least have been more focus on the equity implications of incentivized changes, as well as of the way rewards at subnational levels were calculated, and the implications of often widely varying levels of capacity among agents (for example, schools or local governments) to implement reforms to achieve results.

¹⁰³ Adil and Sabarwal (2017), p.13.

4.3 Autonomy and innovation

“Autonomy and innovation” Theory claim:

RBF provides greater discretion regarding how to achieve specified results. Paying for results on a goal can allow the agent to work out the best way to achieve that goal and to use their local knowledge to innovate, if the principal leaves space for this.

Potential risks:

Loss aversion, with innovation curtailed, owing to perceived riskiness of actions in relation to RBF (non-) disbursement. The number of indicators, their “position” in the education result chain, and their detailed design may all be factors in reducing the space for policy discretion.

Conclusions:

- RBF programs mostly gave autonomy to central government in how finance was used. Though some looser conditionality targeted additional finance for the sector, it was only indirectly incentivized via RBF, in spite of this likely being critical to sector results and performance.
- Nepal’s SSDP had the largest number of DLI results covering the widest range of process, output and outcome targets. These were often sequenced – for example, design a scholarship scheme (process), implement the scheme (output), and improve net enrolment (outcome) – staggered over the five years of the program. This linear structure combined with detailed DLI definitions often limited autonomy in terms of which interventions government should pursue to achieve outcomes.
- The EP4R in Tanzania had a greater focus on outcomes than Nepal or Mozambique. The focus on learning provided a wide range of potential policy options to pursue. However, the relative autonomy implicit in this approach did not stimulate the policy learning/action cycle needed, and the program’s “missing middle” of outputs may have led to some inertia.
- The PFM4R in Mozambique was focused on four specific process DLIs, alongside targeted technical support. This limited its goals to some important reforms, for example improving school supervision by districts. In practice, such changes required cooperation and autonomy to achieve, particularly between the center and local levels of the education system.
- RBF incentives were combined with technical assistance, capacity development, joint sector dialogue, and other projects and technical support. This mix was likely critical to the delivery of many results. Technical assistance played other important roles, including in the complexities of RBF reporting and offsetting some of the costs to government in dealing with the reporting bureaucracy required for RBF.
- In each country, there was more autonomy and discretion at central level than for local government or schools. At local level, space for autonomy and innovation was constrained by restrictions in the use of RBF rewards and by insufficient technical support. For example, improving equity in teacher deployment would require local governments’ discretion, but often this was constrained in contexts with high centralized control, particularly with regard to decision-making in hiring and firing teachers.

Cases of risks materializing:

- DLI frameworks could limit scope for innovation. When DLIs were achieved, successful reforms should have been built on to target subsequent reforms; but such a process was often absent.
- With complex and wide-ranging DLI frameworks, development partners often promoted their own priorities, rather than processes of learning and innovation. Complex and detailed DLI definitions could mean “micro-managing” reforms rather than allowing home-grown learning and innovation to take place.

273. The third theory of how RBF can improve results is the idea that by paying on results, and allowing the agent to work out the best way to achieve them, better use can be made of their local knowledge, as well as unlocking the potential for their creativity and innovation. This implies that there are a number of ways in which a goal can be achieved and the DLI rewards the work, incentivizing the agent to seek ways to achieve goals. The nature of this theory will vary depending on the level of the results chain incentivized – for example for a DLI at process level such as the timeliness of school grants, it may be about the MoF, MoE and agencies working out the best way to reform the PFM system; for a DLI at outcome level such as learning, it encompasses the more complex formulation of policies and strategies that the MoE choose to pursue (and they may be many) to target improved learning.

274. RBF incentives do not act in isolation. For all the programs looked at for this assessment, RBF financial incentives were combined with a mix (varying across programs) of technical assistance and capacity development

associated with the RBF program, joint sector dialogue, and the presence of other projects and technical support. In other words, there was a hybrid element either internal to the program (for instance, technical assistance built into the RBF package in Tanzania and Mozambique's PFM4R) or external (that is, through the presence of other support interventions with related objectives). Even in the case of Nepal where the SSDP linked over 80 percent of external financing to RBF for the 2016–21 period, the involvement of technical assistance as well as dialogue meant that there has always been a significant degree of mutual accountability between development partners and government. Our assessment indicates that it is this mix that has been critical to the delivery of many of the results. In other words, RBF financial incentives in and by themselves would not have delivered the same results. This does not necessarily mean there is less autonomy or discretion, but shows that it may rarely be as simple as direct and independent reaction to the incentive, without support.

275. In all three countries, RBF funding was not earmarked beyond "education" as a sector – hence giving full financial autonomy to the "main agent" with regard to the use of these funds. The PFM4R program in Mozambique was a special case as the funds did not flow through the sector support fund (FASE) but to the MoF, with limited tracking. For the other RBF programs, in both Tanzania and Nepal, large amounts of sector finance have gone into infrastructure development, which in both countries, was only weakly incentivized by the DLI framework – in Tanzania through a DLI incentivizing a strategy; in Nepal retrofitting of classrooms and Model Schools were the only areas in which annual expenditure was significantly higher than the "value" of the DLI disbursement. In Mozambique, the GPE and KfW use of RBF went through FASE, which funds infrastructure as well as the majority of the sector's non-salary expenditure lines, many of which are essential recurrent expenditure (school grants, teacher training funds, textbooks), and funded by donors. The RBF period saw no essential changes to this model, beyond a small portion of the funds being "at risk" and linked to DLIs. At the same time, in none of the countries did RBF engage meaningfully with the broader government basic education financing issue, even though this was problematic in all three cases. In summary, across the study cases, the large financial autonomy given in the designs allowed sector decision-makers to spend on what they saw as most important, whilst the broader program designs also "de facto" allowed governments not to increase funding for basic education.

276. Capacity building can be seen as a way to support agents to wisely use the discretion they have. This was the least well articulated and least well used area of the RBF programs we studied. This is possibly due to lack of clarity as to what the goal should be. There clearly may be a case for structured capacity building when new systems are being developed and rolled out. Governments and funders may also want to reduce reliance on technical assistance in terms of "running the RBF program" (reporting etc.). That may, however, be less realistic in a low capacity context, and possibly less desirable, in the case of complex RBF programs, except if the government was envisaging adopting RBF as its own management paradigm more broadly. Support to "policy learning" – that is, to learn from program implementation and feed this back into policy development – may also be important in many contexts, to strengthen the effectiveness of the choices made by agents, though often more sensitive. In practice, in all three countries, technical assistance facilities played important roles, including in the complexities of RBF reporting and offsetting some of the government costs in dealing with the additional bureaucracy (in multilateral sector review fora and bilateral reporting), and in system strengthening. Their "policy learning support" role was limited – which relates to a point arising from our analysis for most of the RBF theories; that is, that RBF programs, by focusing quite narrowly on the DLIs, arguably paid insufficient attention to policy learning. In the Tanzania case, it was clear that other technical assistance roles were taking precedence – an example of short-term priorities trumping long-term ones. In addition, extensively programmed technical assistance may have limited the innovation and/or autonomy aims of the RBF programming.

277. In terms of autonomy, innovation and DLIs, each program presented differences in how DLIs were designed and framed. As above, the Mozambique PFM4R provided some limited autonomy to achieve the specific DLIs, although the very specificity of the DLIs kept the program scope very narrow. For example when the program achieved the DLI on timeliness of the first tranche of school grants (due at the start of the school year in February), it could in theory have moved on to the remaining delays on the second tranche of school grants (due in the second half of the school year and generally used to prepare for examinations). However, flexibility

and adaptation did not happen in this or other similar cases, and in this sense the DLI framework limited the program scope and degree of innovation. The program also targeted the specific bottlenecks, but had an absence of intermediate outcome indicators. For example, while teacher attendance had been central to the program design and theory of change, it was not measured meaningfully or regularly during implementation. It was therefore not particularly clear that the program moved beyond “target fixation” of the DLIs, to the broader goals of improved PFM or accountability; this was particularly the case for school councils where the targets and data collection did not provide a deep or qualitative engagement in whether functioning had improved, but rather a select number of “tick box” criteria for what “functioning” would mean.

278. In Nepal’s SSDP, the DLIs as designed presented a highly linear results chain approach to RBF, with a focus on process and output DLIs for the majority of its five-year period. In this, development partners highlighted various of their own priorities, particularly in the very specific DLIs that were developed – for example very specific scholarship schemes (which made up the target results for the DLI, rather than targets on another specified goal to “consolidate scholarships”); and for a range of DLI areas specific targets on the number or percentage of geographical areas adopting some new initiative, which would be designed through a DLI result on technical guidance or a particular strategy to be developed in Year 1 of the SSDP. Thus, targets for Year 2 to 5 of the SSDP were in some degree, micro-managing what should be delivered via the detailed definitions of DLIs (for example, the conditions for rollout of the EGRP for DLI 1, exact numbers of books to schools etc.). This approach, particularly when carried out for so many DLI areas, may have limited the degree of autonomy or innovation in programming. As set out in Section 3A and 3B, Nepal’s SSDP provided a clear contrast in relation to teacher deployment to the cases using PTR as the DLI in Mozambique and Tanzania. The convoluted and complex teacher deployment indicator (DLI 3) in Nepal has not (yet) motivated substantial action. While this is linked to the challenges of federalism, it points to the need for clear intermediate outcome indicators if autonomy and innovation are seen as important drivers of success.

279. In Tanzania, the EP4R effect on autonomy and discretion has been mixed, both regarding the extent of discretion provided and its effectiveness. National-level agents had full discretion on the use of RBF rewards. However, arguably, EP4R could have been made more cost-effective by more support to research, synthesis of analyses, evidence-based policy formulation and result-based planning. At local level, the space for autonomy and innovation was constrained both by restrictions in the use of RBF rewards (in spite of emerging international evidence linking increased local autonomy to service-delivery gains)¹⁰⁴ and by similarly insufficient support to communicating a “mindset change”, and to local education planning processes. In some cases (for example, equitable teacher deployment), the discretion required to achieve the results specified was out of the reach of the program agents (LGAs, MoE, PO-RALG) according to the government institutional set-up, while the government actors that had greater discretion (MoF, PO-PSM) were out of the program’s reach. A risk exists of a similar situation for the financial DLIs.

¹⁰⁴ Rasul and Rogger (2018); Rasul et al. (2018); Williams (2017); Pradhan et al. (2014) in the education sector.

4.4 Sharpening minds

“Sharpening minds” Theory claim:

RBF improves evidence and evidence-based policy. As more accurate measurement is required to document and verify results, RBF may increase the quantity and quality of evidence available and, also help to improve policy and planning.

Potential risks:

Compromised evidence systems or verification process, owing to the need for the RBF to be perceived as successful and funds to flow.

Conclusions:

- The RBF programs studied included DLIs to incentivize the strengthening of evidence systems, while also drawing on evidence from these systems. DLIs also created new evidence outside government systems. Some programs sought to introduce elements of subnational RBF using government evidence systems, but potentially before they were of sufficient quality, introducing risks in the process.
- Policy learning was not coherently incentivized by RBF; even in Tanzania, where a specific DLI was in place for this purpose. The evidence in DLI reporting directly from government was not supplemented with sufficient evidence from non-incentivized sources, or research and evaluation outside of the incentivized DLI framework, for example studies on “why” outcome results were what they were.
- A variety of approaches were taken with independent verification, including who hired the verification agent (IVA), how their scope of work and methods were agreed, the resources allocated, and the degree of donor involvement in the process. At least two donors did not require any formal verification for payments to be made for all DLIs. Each approach had limitations, and often the IVA was commissioned to provide a “tick-box” approach rather than a rigorous evaluation of results.
- Mozambique’s PFM4R used the national audit institution for the IVA role. This had the advantage of using government systems and strengthening national capacity. A challenge was the degree to which the role expected was a simple crosscheck of MoE data, as opposed to a more evaluative performance audit.
- In Tanzania, the contracting mechanism for the verification agent gave a relatively higher degree of independence from government. Combined with specified protocols this led to a detailed assessment of achievement and triangulation with qualitative evidence. All sides found this to be effective and it was a respected process.
- In Tanzania and Nepal, DLIs included a focus on learning assessments. This produced important evidence on learning performance; for example, NASA reports in Nepal are now regularly produced and include a high quality of subgroup analysis. It is less clear that this has led to specific improvements in policy.

Cases of risks materializing:

- With the use of outcome DLIs, there were risks when ‘bad’ performance signaled failure on sensitive and complex results. Risks were to both transparency and rigor in measurement.
- There were cases where methods for assessing outcomes were not rigorous. This could create the risk of the impression that a complex outcome was achieved when the DLI result was achieved, when the reality remained unknown.
- There was tension between strengthening the DLI evidence, and developing a wider evidence base, including on whether results led to higher-level goals. For example there were gaps in analysis of drivers of out of school incidence in Nepal; constraints to accountability of school directors in Mozambique; and an absence of analysis on what caused retention to improve in Tanzania, which left stakeholders surprised when it subsequently deteriorated.

280. The fourth theory for RBF is that it can increase the availability and quality of evidence on important indicators of progress and change within the education system, which in turn can be part of feedback loops to improve policy over time. The claim is that more accurate measurement can be generated through the evidence on DLI results, with added rigor from a verification mechanism. There is a flipside risk in which RBF decreases the quality of evidence due to “high stakes” of RBF results (in terms of disbursement or reputation) with inaccuracies in either DLI reporting, verification, or associated evidence systems in cases where results are not achieved.

281. Across all three countries there has been some tension between strengthening the DLI evidence, and developing and strengthening a wider evidence base, going beyond just reporting DLI results. Our analysis suggests the importance of a wider evidence base if RBF is to improve evidence-based policy, which includes the need to monitor important trends and change beyond the program; the need to assess whether DLI results lead to the intended higher-level goals; the importance of triangulating DLI result evidence with non-incentivized

evidence; and the difficulty of capturing more complex and qualitative information, for example on the level of institutional and/or behavioral change, which quantitative indicators often fail to do. Examples of the latter challenge were present in all programs – for instance, the OOSC DLI in Nepal, in which results reported significant reductions in OOSC but with little evaluation on supply and demand constraints to school access, or on how and why improvements may have been made; the school council functionality DLI in Mozambique, successfully achieved target results but with no information on the extent of change in real parents' "empowerment"; and the retention DLIs in Tanzania, where there was no analysis of what caused retention to improve in the first year, which then left stakeholders surprised when it subsequently deteriorated.

282. We found for all three countries, reporting during the period has focused on the DLIs, their performance, and the process of verification, with a significant cost in time of all involved, which stakeholders noted was very substantial. Far greater attention was paid to this than to a wider evidence strengthening strategy. The Tanzanian RBF design which as will be seen below, perhaps most strongly engaged with government evidence systems, paid insufficient attention to institutional issues of data ownership or use as these systems multiplied. Moreover, and in spite of initial intentions, there was little research undertaken directly under the program, no visible use of third-party evidence and analyses, and no feedback loop from beneficiaries – plausibly due to "core program processes" being more onerous than expected.

283. One "direct" way in which RBF may improve evidence is through DLIs interacting with government evidence systems. The programs had different starting points with respect to education statistics systems (including, but not exclusively: EMIS). The EP4R in Tanzania was more structured to use the EMIS data for its quantitative indicators, with the DLIs selected lending themselves to this approach; and through high-value DLIs directly incentivizing the reestablishment and strengthening of the system. In Nepal, EMIS was a focus via specific DLIs carrying out "verification of EMIS data", but the data from EMIS was used only rarely for DLI reporting. Whilst the overall narrative from the "EMIS verification" DLI was of an accurate web-based EMIS being successfully implemented, our review found the data to be more problematic, particularly since the shift to federalism. In Mozambique, EMIS was not a focus for the RBF, although its data was used in the GPE PTR DLI; and there is suggestive evidence there are potentially large numbers of "ghost" students and teachers in the EMIS data, with possible implications in terms of actual achievement levels for the DLI.

284. In Mozambique and Tanzania, there were DLIs linked to school supervision systems. In Mozambique, district supervision of primary schools was a DLI of the PFM4R, and one of its main successes. The DLI focused on the quantity of supervision via the proportion of schools that would be visited in the first and second half of the school year. The work may have kick-started a supervision system, with work from other development partners also strengthening this system including the use of electronic forms and work on the specific data to be collected. Despite this, supervision still does not constitute a new data and evidence system and is not regularly reported at annual reviews, and it is not clear the degree to which the MoE use the information, though this may be a growing focus in coming years. In Tanzania, the pace for the rollout of the new school supervision system incentivized by the DLI represented a risk to the objectives of moving away from top-down inspection towards a "whole-school-visit" comprehensive and participatory performance assessment system.

285. In Tanzania and Nepal there was also a focus on learning assessments. In both this was linked to an early grade learning outcome, with much higher stakes in Tanzania due to the value of the DLI. In the SSRP and SSDP in Nepal, use of NASA was also incentivized, with a view to policy learning. NASA reports have been regularly produced and are of a high quality including subgroup analysis in terms of marginalization, differences by caste, differences between private and public schools etc. It is less clear that this has led to specific improvements in policy however.

286. In Tanzania, the learning DLIs not being achieved in 2017, directly associated to the societally-important goal of improved learning and in the prevailing political economy context, caused political embarrassment which may have borne costs in terms of transparency and longer-term policy ownership and learning. In practice, the

DLI measurement was credible, although narrow, and contemporary assessments by reputable third-party researchers measuring learning more broadly provided more nuanced results that were not considered. Generally, the RBF program did not engage with transparency or government communication, including of potentially adverse program-associated results. When this then happened for the DLIs 6, the reaction (with lengthy discussions about the validity of the measurement approach etc., and government finally not agreeing to results being publicly available) suggests that this may have been counter-productive in terms of willingness for openness and transparency, and detrimental to policy learning too. The experience thus shows significant risks of linking central goals of learning to high stakes incentives. Moreover, it also points to a trade-off between high rigor and cost (associated with the DLI measurement approach) and “good enough” evidence. The case of OOSC in Nepal shows a flipside risk in which a relative lack of rigor in methodology can lead to an impression that a complex issue has been solved (and fairly easily, at low cost etc.) when the reality remains unknown.

287. The three countries provided very different models and experiences with verification. In Mozambique, the PFM4R used the national audit institution (the *Tribunal Administrativo*), the GPE DLIs were verified by a consultancy / audit firm (Ernst & Young), while the German use of RBF in construction had no verification. In Nepal, two universities were used for the formal independent verification agent (IVA) role in the SSDP; and both were also used for the World Bank’s use of ILIs in the SSRP period. For the SSDP, the IVA was required for all World Bank and EU DLIs; though not all of the ADB DLIs were subject to verification, but based on the “endorsement” of the other JFPs. In Tanzania’s EP4R, whilst the procurement was done by one of the funders for a consultancy firm (OPM), the scope of work and verification protocols were jointly agreed by government and all RBF funders. The Tanzanian case appears to feature the strongest verification role, the most detailed and comprehensive in its appraisal of achievement and including some triangulation with qualitative evidence. It was also accepted as authoritative by the MoE, which meant it had a broad-based credibility.

288. In Nepal the Year 1 IVA had one instance of questioning the government’s reporting (for DLI 6 on OOSC), and its contract was not renewed, with a replacement found for Year 2. While the connection is circumstantial, this provides a questionmark on the message this would send for subsequent verification and the degree of independence that would be allowed. In Mozambique, there was also one case in which the IVA, the *Tribunal Administrativo* for the PFM4R, found data conflicting with the government reporting (for the school supervision DLI), and in this case the principal (the World Bank) chose to use the government’s data instead, with methodological fault found with the IVA. Such experiences show that while the IVA holds a critical contractual function, the degree to which its findings are held to be valid is ultimately a decision for the principal and its internal policies.

289. The use of the national audit institution in Mozambique most closely aligned with the RBF ideal of using government systems and building national capacity through the verification exercise. It was clear that this was valuable, although because the terms of reference focused so much on verification, often with “tick-box” methods, opportunities were missed to make the most of a performance audit approach, which the *Tribunal* perhaps wanted to pursue and which is ultimately its remit within the constitution. The choice to switch to a private firm for the GPE variable tranche DLIs provided clear and professional reports, but again these were limited in focus, with insufficient analysis to generate much learning. For example, no methods were used to look at the quality of the in-service teacher training conducted. In Tanzania, whilst the verification was “external”, it provided feedback on the EMIS and school supervision systems and in the case of the EMIS, this feedback was found useful and mostly acted upon, hence further strengthening both the capacity of the team in charge, and the system itself.

4.5 Labeling success

“Labeling success” Theory claim:

RBF induces better results-based planning. As goals are labeled as the results of interest, planning may focus more systematically on results – identifying bottlenecks to these results, and actions to address these, with associated indicators.

Potential risks:

Lack of ambition if results are simple to achieve, or achieved within the time-scale, with years still to go to the end of the plan period. Financial flow risks because payment is typically in arrears or via non-achievement, meaning fewer resources for sector or unpredictability of fund flow.

Conclusions:

- All three countries saw RBF reporting align to annual reporting cycles, following the structure of joint sector reviews which predated RBF and underpinned government to donor dialogue. This structured planning and dialogue approach was helpful to focus on results, though for bigger programs there was a risk that RBF crowded out dialogue on non-incentivized results.
- Collaboration in the RBF design process varied. In some cases, it was done in tandem and very closely aligned to the sector plan with DLIs selected from a wider sector results framework. In other cases, donor prioritization in planning and DLI selection was stronger.
- Where DLI frameworks were extensive in terms of the number and range of DLIs, the burden on planning systems was significant, reporting could take large amounts of time and resources, and this was not always sufficiently addressed. With more limited and targeted frameworks such as that of the PFM4R in Mozambique, the results-based planning approach was clearer, although potentially also less ambitious.
- In all three countries, the use of technical assistance alongside DLIs supported the planning and management of reforms as well as relationship management with the development partners. This support was often needed, but was also more often focused on DLI reporting than strategic sector planning.
- Given that basic education management is decentralized or deconcentrated, an important part of the “labeling” effect and of strengthening result-based planning depends on how this is being transmitted to subnational levels. RBF programs engaged with these levels through existing systems but did not usually give enough support or attention to subnational planning processes, and in some instances created ad hoc systems for implementation or reporting, with a very specific focus on DLIs.

Cases of risks materializing:

- Financial risks were found and even though payments were often deferred or ‘rolled-over’, due to resource constraints even delayed payment could lead to less resources for the education sector in a given financial year.
- Detailed DLI definitions for processes and outputs meant, particularly in Nepal and Mozambique, that programming was at times “micro-managed”, leaving less room for learning and adaptation or planning towards outcomes.
- RBF “labeled success” but this could become an “indicator fixation”, in which the strategic underpinning of interventions set out in DLIs was at times perceived to be lost. This focus could take away time for problem-solving on key bottlenecks in the system and this was linked to what some government stakeholders called a “target obsession” of development partners linked to the need for DLIs to be achieved and funds to flow.

290. The fifth theory for the benefits of RBF is that the presence of DLIs improves the overall approach to results-based planning and management. There are two important ideas here – one is that goals are labeled as a starting point; the second is that (joint) planning is undertaken with a focus on these main goals via what can be called a “results chain” approach; that is, to identify the specific bottlenecks that need addressing to achieve goals, be they educational equity, learning or access.¹⁰⁵ This process could be integrated into the sector plan itself, improving the quality of the overall plan, or accompanying operational or annual implementation plans. Or it could be part of how delivery is cascaded through key institutions and actors within the education system, and review processes feed back into planning and policy.

291. The RBF programs assessed showed some of the ways this can work, but equally that RBF can undermine results-based planning in some instances. The programs all aimed to support the implementation of the country’s education sector plan, to achieve results and strengthen the education system. In making this contribution, the

¹⁰⁵ World Bank (2017a)

RBF programs' results chains took a broad approach in Nepal and Tanzania, covering a range of results areas, and were narrower and more selective in Mozambique. In Mozambique, the programs focused on a select number of process and output DLIs and while this was appropriate to the context, it is important that higher-level outcomes and the longer-term goals remain in focus given that they were not themselves specified as DLIs.

292. Even where results chains were longer, more oriented on outcome indicators, and with a mix of DLI levels, RBF programming would always still only relate to a subset of sector bottlenecks. The design of an RBF result chain and selection of DLIs requires careful consideration, as a trade-off between specificity and leaving room for local choice and autonomy (see *Autonomy* theory above). There is not a linear path from one DLI to more complex outcomes, and to strengthen results-based planning, an RBF program needs room for learning and adaptation (see *Sharpening minds* theory above), and strengthening the results chain. For this, greater attention needs to be paid to monitoring that looks beyond the DLI results, to questions around why the results are what they are; contributing to the next steps where initial (process/output) results are being achieved, and to addressing obstacles where not. The monitoring and associated dialogue has at times been seen to focus too much on specific DLI results, with some stakeholders highlighting the risk of "target fixation", rather than on the broader results chain and higher-level goals. Monitoring of this kind should ensure it does not neglect non-DLI results that also constitute or contribute to the goals of the sector and sector plan.

293. To improve results-based planning, RBF would need to align with and reinforce or strengthen existing planning processes. From this research, this seemed to work more with national-level processes; while there was less support evident for local-level planning and its interface with national planning processes, or the balance between vertical and horizontal planning and accountability. This varied across the country contexts studied. In Tanzania, overall, the results-focused dialogue of RBF worked well with the government's BRN and with the post-2015 presidential drive for results, but in practice this was limited to key individuals both at central and local levels. Whilst through its support to strengthening the EMIS and the sector performance review report, the program was instrumental in enhancing the content of the broader sector review dialogue, the government-donor dialogue was almost exclusively focused on the DLIs. Moreover, the broader sector annual review process lacked operational planning content, there was not an effective process in place for bottom-up planning, and the RBF stakeholders did not address this through the programming. Even under the EP4R, results were identified, but there was little systematic planning of the actions needed to achieve them, and a lot was left to local initiative. The DLIs alone did not track a coherent theory of change in terms of process and outputs to learning outcomes (for example PTRs and SQA are not likely enough to directly impact early learning to the magnitude required), this means there is significant autonomy (see above) in the program but this may have meant that strategies to target learning were, ultimately, "missing".

294. In Mozambique there was good alignment with and use of the existing process of operational planning and annual activity planning; however, this was seen as a top-down process which may have worked well for the results and DLIs in focus, but would be harder for more complex behavioral change requiring more context-specific planning and responses (that is, more bottom-up and/or local). The PFM4R was a well-planned program, identifying DLIs which focused on key bottlenecks with coaches and facilitators to ensure joint planning and implementation of changes to achieve them. This shows to some extent that a discrete RBF program can add value. Across each of the RBF programs looked at, there was more focus on the specific DLI and metrics than the higher goal (for example, focus on criteria for school councils rather than on the more complex local reality; the timely arrival of school grants rather than whether they were spent well). The DLIs were justified in their selection, but with a limited time span, once achieved they were not built upon with follow-up indicators. This was the case for example for the second tranche of school grants following the success with the first tranche in the PFM4R.

295. In Nepal, the sector is undergoing huge upheaval, and will require significantly more bottom-up planning in the context of federalism. Under the SSDP, the DLIs became the focus and in this sense the RBF "labeled success", with significant attention given to reporting and seeking to deliver the DLI results. However, an

“indicator fixation” has been described to us, in which the strategic underpinning of the interventions set out in DLIs was at times lost. Some stakeholders questioned whether the RBF / DLI approach might reduce incentives for problem-solving on key bottlenecks in the system and this was linked to what government stakeholders called the “target obsession” of development partners; as one put it, “we are missing visionary planning”. The nature of contractual DLI targets constrained the degree of flexibility needed to respond to the new reality of federalism. Flexibility in changing DLI results and in the interpretation of detailed DLI definitions may have allowed the SSDP to avoid non-achievement of many DLIs from Year 3. However, the process itself has been time-consuming, and has perhaps been a distraction from the broader effort needed to adapt to federalism and to think through, in the new set-up, an efficient yet “acceptable” approach for the center to cascade its objectives and for these to be both successfully implemented and reported upon by the local level.

4.6 Aligning all actors

“Aligning all actors” Theory claim:

RBF improves co-ordination between key institutions. As results are clearly identified, the policy dialogue is more focused, with ensuing collective action made more likely by the incentivized nature of the framework.

Potential risks:

Dispersion of actors and coordination; for example, donors focusing on “their DLIs” as opposed to the overall sector plan.

Conclusions:

- The alignment of sector actors did not automatically ensue from RBF but required attention to ownership of strategies to achieve results. At its best, consultation was broad-based, included all relevant actors, and time was taken to build consensus. Where development partners promoted reforms without sufficient buy-in on the government side, they were either not achieved, or achieved but only in a narrow sense (for example DLIs “achieved” but without associated deeper reforms).
- In Nepal, the RBF period was concurrent with the “big bang” of federalism in which the delegation, financing, motivation and reporting functions of different tiers of government underwent radical change. The new tier of local governments became responsible for most basic education management tasks. Federalism heralded greater autonomy but also led to gaps in civil service positions and technical capacity. The DLI framework was adapted, but took time to be made relevant to the new reality.
- In Tanzania, there was a sophisticated program structure, with RBF-financed subnational payments to LGAs for DLIs they were directly responsible for. This may have secured greater buy-in from local governments, as well as greater awareness of the DLIs at local level than in either of the other two countries. However, this awareness did not appear to extend to DLIs attributed to the national level, even though some, such as the learning DLIs, also required local and school-level attention and action.
- In Tanzania, strong program governance from development partners and government, supported the EP4R’s coherence in spite of the priority shift to access with the FFBEP. Within government the program contributed to greater MoE and PO-RALG collaboration; and to improved coordination between national and local levels, although from a low base, and with a narrow focus on the DLI framework.
- Mozambique had a more stable, top-down structure of relations from the MoE to deconcentrated provinces and district services. The PFM4R was well-designed to work within this structure, and also played a role in ‘aligning actors’ by bringing together the MoF and MoE, with a focus to improve timeliness of fund flow for school grants in particular.

Cases of risks materializing:

- There was at times a power imbalance between the donors and the government in negotiating changes to DLIs. Government was at the behest of development partners’ priorities and internal systems to ensure that DLIs could be changed, even when significant contextual changes made this essential.
- The membership of third party voices including civil society, NGOs, and advocacy organizations within joint sector fora, provides potential for enriching dialogue on sector results. There was at times a risk that simplistic narratives of success linked to DLI achievement could undermine these voices, framing their nuanced views as “dissident” and legitimate to dismiss.
- A combination of PFM, core financing, and political economy constraints were behind many of the issues addressed by DLIs, but were rarely discussed. This can create a degree of confusion and lack of transparency in sectoral discussions.

296. The sixth theory of change for RBF considered is its potential ability to bring stakeholders within the education sector together and improve their coordination. This relates to RBF as an accountability mechanism and how it fits within the existing system of relationships between institutions. RBF could provide a focus for improving dialogue around results between development partners and government. RBF could also play a role in improving alignment and coordination among national government agencies (including central agencies such as the MoF), between national agencies and subnational deconcentrated structures or local government institutions, as well as down to school level. These relationships are formalized through definitions of planning and management mandates and through financial, reporting, audit and other accountability functions; and RBF introduces a new factor which interacts with these mandates and functions, and may enhance them in the process.

297. The experience of all three countries has shown the centrality of both local-level accountability and donor-government accountability in sector performance. In terms of the former, the three countries provide different

contexts and also different models in engaging with the local levels. In Nepal, as extensively outlined, the SSDP period has been concurrent with the “big bang” federalism in which the delegation, financing, motivation and reporting functions of the different tiers of government underwent radical change, transforming basic education management. Adapting to this change was a central challenge to the sector and the RBF program. Mozambique in contrast had a more stable, top-down structure of relations from the MoE to deconcentrated provinces and district services; the PFM4R was well-designed to target the structure as it existed in terms of supervision and transfers, though the intended school council changes, relating to horizontal accountability to parents and communities, were more difficult to achieve. The PFM4R also brought together the MoF and MoE, with a focus to improve timeliness of fund flow in particular. In Tanzania, there was a more sophisticated program structure with RBF-financed subnational payments to LGAs for DLIs they were directly responsible for, and this may have secured greater buy-in from local governments, as well as greater awareness of the DLIs at local level than in either of the other two countries. However, this awareness did not appear to have extended to DLIs attributed to the national level, even though some of these, such as the learning DLIs, also required local and school-level attention and action.

298. Part of the alignment that RBF can bring arises from the “introduction” and design process. The “results lab” model that underpinned the RBF design in Tanzania, building on shared recognition of the central challenge (the “learning crisis”) including at the highest level in government, was a successful example of collaborative process which clearly contributed to developing a coherent structure in the DLI framework with good alignment of the development partners involved. This was also helped by the fact that donors and government explicitly discussed the use of RBF as a new modality and how the risks entailed by this new way of working could be mitigated. Mozambique’s RBF was also based on good alignment with donors and government, with the PFM4R growing out of a long process of diagnosis on sectoral PFM constraints through work of the World Bank and government, having grown out of the government’s long-term “PFM vision”. The GPE variable tranche RBF programming was directly tied to the 2015 Operational Plan, with very good MoE ownership, although the GPE at the time may have wanted more stretch in the DLIs than ultimately achieved. The MoE had clear ideas of the reforms targeted. In the case of Nepal’s SSDP, the use and development of RBF and of the DLI framework was simultaneous and grew from the development of the sector strategic plan, costed framework, and results framework. With so many DLIs, the RBF program had the ambition to be comprehensive, and its development involved close working of joint sector technical working groups. Ultimately, however, RBF funders may have had a strong sway in how DLIs were designed and in some cases, they clearly represented their own priorities as much as government’s. As a result, some of the DLIs were questioned in terms of relevance, with some donors skeptical about Model Schools (DLI 5), and on the government side some frustration at DLIs perceived as “political”, such as curriculum reform (DLI 2).

299. In Tanzania, the EP4R program did not, and did not attempt to, strengthen broader coordination in terms of the links with other basic education projects and relevant projects in other sectors. Central agencies such as the MoF, and PO-PSM, remained out of reach for education sector dialogue. Given that the most critical issues affecting the program, and more broadly basic education results, were political, development partners failed to engage meaningfully at that level to support senior sector technocrats, who felt unable to influence politicians on their own. More broadly, the program lacked a continuously refined institutional map spelling out the institutional and behavioral changes required to achieve the results specified, and this undermined its ability to align all actors. Despite this, there was strong within-program goal coherence in spite of the FFBEF priority shift at highest government level. Strong program governance from development partners, the MoE, and PO-RALG, supported this. The EP4R contributed to develop greater MoE and PO-RALG collaboration; and to improve coordination between national and local levels, although from a low base, and with a narrow focus on the DLI framework.

300. During Nepal’s SSDP period, there have been major challenges to align sector working with the new shape of government, with LGs now responsible for most basic education management tasks having greater autonomy but facing gaps in the civil service (with many LGs not having an education officer in place) at the

same time as lacking the institutional capacity of the former deconcentrated education management structures and of the network of experienced staff in these structures. In spite of the “joint approach” instigated via technical working groups, in practice, the detailed renegotiation of changes to the DLI framework have often been bilateral. The flexibility required to deal with the changes of federalism has been critical to ensure that funds will still flow during the SSDP period, and changes make the program more rational in response to the new government system. However, processes have been slow, and it is not at all clear RBF has helped the process of transition – indeed, reporting on and renegotiating DLIs required time whilst arguably, attention was needed and urgently, to develop ways of building alignment and strengthening institutional capacity among LGs.

301. In Mozambique, sector coordination mechanisms are well established, with the Local Education Group and the “Troika” mediating donor to government dialogue. RBF has not substantially changed this, though has worked well within this. Despite the joint approach, most of the RBF programming involved strong bilateral approaches. This included construction, in which Germany (KfW) has provided technical assistance, as well as the use of RBF to try to improve its speed and efficiency. While we were unable to analyze this in significant detail, it is notable that reporting on the construction bottlenecks and the use of KfW RBF was limited in joint sector annual review reporting, despite it being one of the largest and most problematic areas of FASE expenditure.

302. As was clear in all three countries, political economy constraints are behind many of the issues addressed by DLIs, but are rarely discussed directly. This can create a degree of confusion and a lack of transparency in sectoral discussions. Across the countries and likely in many contexts, challenges in education core government financing and PFM are also manifold, and the RBF programs we studied did not clearly engage with these, leaving full discretion to government in how RBF funds were being spent in the sector whilst not addressing the broader financing and PFM issues, with limited exceptions. While this may be in line with the “hands off” approach supposed to underpin the use of RBF, it may mean central sectoral bottlenecks are not being adequately addressed. Yet, as seen above, RBF programs were arguably not that “hands-off” in other respects.

303. The membership of third party voices including civil society, NGOs, unions and advocacy organizations within joint sector fora, provides a potential for enriching the RBF dialogue as the basis for sector actors’ alignment. However, in many contexts this raises tricky political economy issues as well. There is a danger that the simplistic narratives of success that some DLIs bring may undermine these voices, framing their more nuanced views as “dissident” hence legitimate to dismiss. Whilst there was no strong evidence of this, there were cases in which research was produced by third parties including NGOs which implicitly questioned the DLI results and underlying goals as reported, but which was not engaged with by RBF programming. There was also some evidence in the Tanzanian case of the opposite risk – that these voices might be politicized, and in using adverse results for politically-motivated aims in an unwelcoming context, actually undermine the potential for genuine dialogue on results.

4.7 Sustaining attention

“Sustaining attention” Theory claim:

RBF maintains the focus on specified results during a period of change through its contractual nature, more than would be possible with a traditional (non-contractual) results framework.

Potential risks:

Inflexibility: inability to deal with or adapt to changing context over time, because of the fixed nature of the RBF “contract”.

Conclusions:

- All three countries experienced major changes to context during the programming period. Most pertinently, there were changes to the political leadership and associated priorities in Tanzania in 2015; the federalism changes in Nepal in 2017-18 and ongoing; and in Mozambique, the suspension of GBS in 2015-16, which reduced the fiscal space for the education sector.
- In practice, RBF programs have room for flexibility. The Tanzanian RBF was in two main phases with a mid-term review (MTR) leading to new areas of prioritization from 2017. The SSDP in Nepal also underwent a MTR in 2019 which was a key point in addressing the realism of program targets in light of federalism. In both countries, DLI targets were changed once formally agreed between development partner and government, but this could take time.
- In all three countries, there were cases in which non-payment occurred. While this demonstrated the “credibility” of the RBF mechanism, there were many cases where a combination of rollover provisions, program restructure, flexibility in interpretation, and redesign, were applied. This meant the value of disbursements was largely maintained, and indicates that the need and/or desire for disbursement could override considerations of credibility.
- In Mozambique, RBF was used over a relatively shorter period (mostly during 2015-17), reducing the risk that the focus might shift over time. Despite this, the suspension of budget support constrained the national budget, with likely implications for DLI targets such as PTRs in terms of hiring more teachers. While there may have been a ‘sustaining attention’ effect for the program period, there is less evidence this lasted once the RBF programs were complete.
- In Nepal, despite the SSDP moving into a ‘transition phase’ from 2018, it took two more years to change many of the DLIs that were either no longer relevant or substantially more difficult to achieve due to federalism. It is possible that the focus on DLIs may have distracted attention from deeper realignment to the new institutional reality.
- In Tanzania, the EP4R RBF focus on quality and learning had to compete with the non-RBF priority of increasing access. The latter was a change not even two years into the program, but despite this the RBF did ‘sustain attention’ on learning and quality in the 2017 revisions to DLIs. The effect was mixed, as government prioritized access by using RBF rewards to fund school infrastructure, over a number of possible more quality-focused actions.

Cases of risks materializing:

- The contextual changes in Nepal and Tanzania did not lead to a coherent reassessment of the “stretch” or “risk level” of the program. MTRs could have provided a systematic reappraisal but were not always planned in this way.
- Flexibility in changing DLIs and in the interpretation of detailed DLI definitions may have avoided non-achievement of many DLIs. However, the process itself can be time-consuming, and may have been a distraction from the broader effort needed to adapt to major changes to the contextual environment.

304. The final potential theory of change considered for RBF is the idea that the firm nature of an RBF contract could maintain a “focus on results” over time. In particular, there could be a benefit for RBF to be useful in maintaining a (contractual) definition of achievement and goals (results) during a period of change, which has been framed as keeping “*eyes on the prize despite the political economy*”.¹⁰⁶ That said, there are some major risks that “sustaining attention” may have unintended negative consequences – arising from preventing adaptation in the face of major contextual changes.

305. The experience of RBF assessed across the three countries have had different timelines, and various alignment with sector planning timeframes. In Mozambique, the RBF programs, with target years between 2015 and 2017, all fell within the 2012–19 Strategic Plan period, with the 2015–18 Operational Plan linking to the GPE additional financing, and there was therefore strong alignment of timelines. The RBF period was short, and this reduced the risk that the focus might shift over time, but this, and the fairly “simple” design of the programs, also reduced the risk that significant flexibility would be needed. The most important shock within the RBF period

¹⁰⁶ World Bank (2017a)

in Mozambique was the “hidden loans” scandal and associated suspension of budget support. This constrained donor inflows to FASE and to the sector budget more broadly, the former raising challenges for construction and the achievement of the KfW targets, and the latter linking to reduced numbers of teachers entering the system each year, making PTR targets harder to achieve. Despite this, the PFM4R may have provided a strong model for focusing on a set of very specific results. With only four DLIs it was also less challenging than RBF programming in Tanzania and Nepal.

306. In Nepal, the first experience with RBF involved only one development partner, the World Bank, in the 2013–15 period towards the end of the SSRP (2009–16), with only four indicators. A notable point is how little documentation there was on performance and possible lessons in using RBF, including in the project completion report. This may have limited lesson-learning leading into the major use of RBF of the SSDP period in which six JFPs were involved. The SSDP period since 2016 has involved a large DLI framework with 84 separate results across ten DLI areas, and defining the focus hence “sustaining attention” up to the final year of the program. Federalism, however, has required adaptation and flexibility; and mechanisms for this have been slow. Moreover, the flexibility required has been at the discretion of the JFPs – except for the possibility for some of them to rollover funding for delayed result achievement, it was not “written in” to the contractual set-up of the SSDP. Even the TORs of the SSDP mid-term review did not include a review of DLIs, although it was undertaken by the consultants as a side-deliverable. As such, despite the SSDP moving into a “transition phase” from 2018, it took two more years to change many of the DLIs that were either no longer relevant or substantially more difficult to achieve due to federalism. Moreover, it is possible that the focus on DLIs may have distracted attention from deeper realignment to the new political reality, and led to ignoring developments which had potential to drive better education service delivery. For instance, with federalism the network of support to schools was taken away without an adequately staffed or capacitated alternative, but the RBF framework, before being adapted, acted as though nothing had changed. Anecdotally, the most proactive LGs may have instigated improvements to service delivery that have nothing to do with the DLI framework or central-level guidance, but this was not likely to be noticed.

307. In Tanzania, the EP4R continuity in focus on quality and learning, alongside the FFBE high-level priority of increasing access that arose as a major and sudden change not even two years into the program implementation, was agreed with the MoE and PO-RALG, who felt “sustaining attention” was important. The mid-term review and redesign process explicitly provided space for assessing the changes entailed by the FFBE, and led to the selection of new DLIs with additional funding to address some of their potential undesirable effects. However, the “sustaining attention” effect was mixed, as the MoE and PO-RALG prioritized access in the use of RBF rewards (mostly on school infrastructure) over a number of possible more quality-focused actions. There were other cases in which sustaining attention arguably prevented relevant adaptiveness: development partners’ refusal to amend the equitable teacher deployment DLI in spite of deteriorating results, as a way of putting pressure on the government to recruit more teachers, was misguided, as the pressure targeted powerless agent(s). In not making the financial DLIs more demanding at redesign, the EP4R partners continued their lack of attention to what is now the single most pressing challenge to the program results: government underfunding of basic education and insufficient teacher numbers. In addition, the complexity of the DLI framework, measurement approaches, and the reward allocation framework, may have weakened the sustaining attention effect, that is, given the number and range of DLIs; demanding measurement standards distracting focus from results; and a lack of or weak understanding of reward flows.

308. The above highlights that adaptation may be required in case of a major contextual change – sudden as with federalism in Nepal, or gradual as with the growing “mismatch” between government policies (encouraging rapid enrolment increase without commensurate resource allocation) in Tanzania. The experiences of these two countries suggests the usefulness of trying to anticipate such change as much as possible, and that this may often be feasible, at least partly. However, this requires a systematic process of anticipation and projecting the program in the future, to reassess, in light of the possible or anticipated change(s), the relevance of “sustaining attention” to the same set of results. This is in terms of feasibility and reassessing the “stretch” parameters and

“risk level” of the program; and even in terms of desirability of maintaining the focus on the same areas. Usually such “large contextual change” goes beyond affecting the education sector only, and even if it is education-specific (FFBEP in Tanzania), does not emanate from the sector’s usual decision-making processes. Whilst this may make them harder to anticipate, it suggests the value for education stakeholders to engage with important beyond-sector dynamics (preparations of the legislative framework for decentralization in Nepal, budget process in Tanzania). Then, there are also *force majeure* events like the COVID-19 pandemic, requiring the same kind of “re-think” process. Mid-term reviews may be a forum for this kind of “re-think”. However, in Nepal this was not planned in that way, and more broadly contextual changes requiring flexibility may not “fit” with programming time schedules. The Nepalese experience, especially, also highlights the risk that the reporting system and processes of RBF, with the associated time and dialogue, may reduce the ability of sector actors to respond to new realities.

309. DLIs only form a subset of sector indicators, for example those in the results framework of the sector plan which in principle should also be monitored on an ongoing basis. As above (*Signposting* theory), there are therefore risks of “diversion”, in which important but “non-RBF” programming areas are deprioritized. There are also risks to sustainability in which once the RBF period finishes, any focus disappears. In Mozambique the PFM4R provided a short-term focus on some key bottlenecks, but was lacking in ongoing and sustainable planning around longer-term sector PFM improvement. In Tanzania, as discussed elsewhere, it is far from clear that government attention to early grade learning, which has already become more narrowly focused on the DLI measurement process and results with less attention to measures that would ensure better teaching/learning at these levels than in the first years of the program, would be maintained if it was no longer incentivized in RBF programming.

5. Lessons and recommendations

310. This research is informed by the theory of RBF as set out above, with the assessment based on three countries and six RBF programs, two of which it was not possible to study in as much depth.¹⁰⁷ The programs were very different in terms of design, and have been programmed and implemented in very different contexts in terms of the political economy and the nature and histories of each education system. Despite the limited size of the sample, from the in-depth nature of the study it is possible to draw a number of insights that relate to one or several of the program features we examined (design, results, flexibility/adaptation, etc.) and/or one or several of the theories underpinning the use of RBF.

311. The research began with Preliminary Assessments with RBF mapping undertaken on the three countries.¹⁰⁸ The majority of RBF found in each case was from aid-financed RBF programs, and these programs have become the main focus of the assessment including the Final Assessment Reports for each country,¹⁰⁹ the findings of which this report has synthesized. Whilst some of the insights proposed here may be useful to RBF programs that would be run by governments and with government financing, this is not our aim here. In a bid to be useful to aid-financed RBF practitioners on both donors and governments' sides, the suggestions we make are based on the fact that such RBF programs always entail a donor-government relationship. We also aim to be practical, at the same time recognizing the importance of assessing the relevance of any suggestion in the particular context one is dealing with, hence trying to identify, based on the cases studied here, what were the element(s) that made or might have made what we suggest effective, in one or several of the contexts we studied.

312. The Table 22 below sets out the recommendations and connected lessons that we hope will help those involved in either the scoping, designing, implementing or evaluating of RBF programming, to think through some of the major challenges and try to maximize the opportunities of RBF. It should be clear from this assessment that RBF is not a magic bullet, and should work within realistic confines of change, embedded in education systems with established actors and relationships, specific strategies that government sector planning seeks to pursue, and taking account of the broader, government-wide dynamics and prioritization processes. RBF is also likely to be part of a mix of modalities including other sector support, projects, and the many range of many third sector actors in the system including CSOs and NGOs, as well as private sector education providers. Our recommendations then seek to provide some important and realistic ways to improve RBF in such settings.

¹⁰⁷ The main programs of focus were EP4R in Tanzania (2014-); the SSDP in Nepal (2016-21); the PFM4R (2014-18) and GPE (2015-17) use of RBF in Mozambique. Included though with less focus in terms of interviews and documentation were the RBF used by the World Bank in the SSRP period in Nepal (2013-15) and Germa's (KfW) use of RBF for construction in Mozambique.

¹⁰⁸ Preliminary assessment reports for Tanzania, Dom et al. (2020); Mozambique, Patch et al. (2020); and Nepal, Holden et al. (2020).

¹⁰⁹ Final reports for Tanzania, Dom et al. (2021); Mozambique, Patch et al. (2021); and Nepal, Holden and Chapagain (2021).

Table 22 Recommendations and lessons from the RBF assessment

Recommendation	Connected lessons	Explanation
1. Be clear why you are using RBF while anticipating and mitigating risks	1A. There are multiple objectives of RBF, and some of them may compete, so priorities should be made clear from the outset.	The selection, design and use of DLIs should vary depending on the intention of RBF to incentivize effort and use of resources by different actors, the balance of support in planning how to achieve results and the autonomy/discretion of agents. Ensure that the design process is collaborative both with government and among donors, and it is clear there is strong alignment on the direction and specifics of policies and strategies pursued.
	1B. Risks that are specific to RBF – information risks, motivation risks, equity risks – should be set out in the design phase, with clear strategies for mitigating each risk.	Program design processes can pay little attention to specific risks associated with the use of RBF. Design teams should assess these risks and wherever possible build safeguards, e.g: - Cheating/fudging: Strengthen government evidence systems and triangulation; support government communication including when results are adverse. - Cherry-picking: Calibrate DLI rewards to prioritize the most difficult situations (e.g. remote schools). - Diversion: Explicitly embed the RBF program in the sector results framework, results chain and broader monitoring. - Demotivation: Ensure systematic identification and consideration of change in stretch and discussion of processes for flexibility and adaptation.
	1C. Financial and non-disbursement risks come from delays and under-performance. Risks can be offset by rollover and flexibility, but with implications on the credibility of RBF as a contractual mechanism.	The main financial risks are for government. The timeline of disbursement of RBF should be realistic in light of need for reporting, verification, and application of rollover provisions and flexibility. Risks can be mitigated through formalizing a mechanism for their ongoing assessment and building-in frequent and explicit stocktake/break stages to reconsider stretch/financial risk balance. This, in turn, should be based on broad evidence building, policy learning processes and anticipation exercises.
2. Identify, discuss, and agree how RBF can be part of a mixed portfolio of approaches	2A. Targeted technical assistance may be needed; and this should link to clear planning on its contracting, its role and how it is structured.	RBF incentives do not act in isolation. Technical support can play a key role in achieving targets. This support can focus on process/systems development, specific technical areas, policy learning, broader capacity development. It can be provided under the RBF program or through parallel technical assistance teams and projects. So, clear design and strong operational and technical coordination is needed.
	2B. RBF is only one part in a portfolio of sector support. Donors should be clear on the mix used, and plan for complementarities.	There can be instances of explicit coordination and of unplanned but useful complementarity with other projects. RBF programs should assess overall coherence of the basic education aid portfolio, to strengthen complementarity between policy-focused RBF and more operational instruments. This should also contribute to ensure links in the government’s sector strategy, including how non-incentivized results are addressed (identified, planned for, supported).
	2C. RBF alone will not address broader governance and political economy bottlenecks: a broader set of approaches are likely to be required.	There may be areas which RBF / DLIs are too “crude” an instrument to address, and this includes “deeper reforms” to governance, PFM and addressing political economy risks and issues. Evaluation and analysis, dialogue, and possibly the provision of technical and capacity support, are as important as DLIs or possibly more in some instances, to address these issues (e.g. education financing). It may also be important for RBF to seek synergies or build links with other work outside of the sector – in broader governance and PFM reform and political dialogue, for example work with the MoF and other central agencies as

Recommendation	Connected lessons	Explanation
		well as the MoE. Where possible, learning from RBF in other sectors on how they address governance issues (e.g. with regard to LGs' role in service delivery) is also advisable.
3. Have clear results chains identifying the changes and goals targeted, including a balance between feasibility, desirability and flexibility	3A. While DLIs target specific results, these are also likely to link to higher-level sector goals to which the RBF design is contributing, alongside other initiatives. Ideally these goals should be made explicit at design, monitored as part of the program monitoring framework, and considered in the light of results from the other initiatives.	Where RBF programs' results chains are short, it is important that the higher-level goals that are intended are clear. Assumptions for how DLIs lead to better student outcomes should be clear, with an assessment of whether these are likely to hold and what will be required (e.g. finance, donor support, human resources, government buy-in, government policy change etc.), and indications of how thorough evaluation will be undertaken on these changes.
	3B. There can be a trade-off between using the DLIs to map each level in the chain towards outcomes (specific processes, outputs etc.) vs. fostering ownership, autonomy and adaptive management.	RBF programming designs can span a spectrum of very little specification and few DLIs targeted at outcome level; few DLIs but targeting a set of well-designed process results; or a larger number of DLIs at all levels in the results chain, therefore giving the impression of being "complete". Each approach has advantages and disadvantages, with trade-offs in autonomy and flexibility vs. internal program completeness. However, the strength of linkages and coverage of the DLIs (or lack thereof) should be clear to all parties, with realistic expectations and a clear evidence base underpinning them. Substantial external support may be required to strengthen planning in the absence of clear wayposting from the DLIs.
	3C. There may be additional costs, and increased complexity, of "larger" DLI frameworks, and when very large, there is a risk of crowding out other important results from sector dialogue.	While there is not an "optimum" number of DLIs, each DLI requires detailed definitions, disbursement rules, verification protocols, and therefore will create reporting cost and complexity during implementation. In addition to the cost and complexity of achieving and measuring each DLI, cost and complexity also arise from the structure of the DLI framework as a whole (number of DLIs, points in results chain etc.). Even when RBF programs aim high, and/or cover the education results chain more broadly, one single program cannot be expected to unblock all key bottlenecks. Collaborative and robust dialogue to identify key bottlenecks is critical. As said above, there will be some bottlenecks for which RBF might not be effective and would need to be complemented by other support modalities.
4. Assess, discuss and explicitly agree on the degree of stretch including a focus on complexity, and the potential trade-off with disbursement	4A. Results vary in the degree of stretch for each DLI, although actors often have no consistent way of assessing it. Ideally it should be explicitly assessed and methods considered could include assessing the degree of social and technical complexity to achieve the result.	There are different aspects of complexity that may drive the degree of stretch with a DLI – this includes social complexity in terms of the number of agents involved and the political economy of their coordination, and the technical complexity of the result itself. While it may not be easy to objectively assess these aspects, it may still be useful to systematically think about each DLI in this way, to assess the likely stretch that DLI targets represent.
	4B. The likelihood of achievement of result targets in light of their stretch, will have implications in terms of incentives including risks of demotivation. The	It is likely that all targets will not be achieved if there is sufficient stretch. This may be "acceptable" (to government) to an extent. However, if the stretch is too high or unrealistic – e.g. because agents are not sufficiently "in control" of what is needed to achieve the targets, RBF/DLIs are likely to no longer be incentives, and demotivation with the

Recommendation	Connected lessons	Explanation
	<p>perception of how much stretch there is, is also likely to change over time.</p>	<p>RBF mechanism may result. If the stretch is such that DLIs become punishment of poor performance, it can reduce the buy-in to the program as a whole.</p> <p>The stretch may change over time, for example with large changes in government’s set-up, or other external changes, and therefore should be continually reassessed.</p>
	<p>4C. Disbursement is a legitimate objective of any financing arrangement; discussions between donor and government should be clear on the trade-off in likely rates of disbursement when there is higher stretch.</p>	<p>It is important to ensure that the financial implications of non-payment are clearly understood prior to starting any RBF project, in particular as external finance is likely to be helping to fund essential expenditure. The higher the degree of aid dependence, the more severe the implications may be in terms of potential costs of non-disbursement. Having a clear view of the likely extent of achievement of DLI results given their stretch, is essential to understanding the specific financial/disbursement risks of the RBF modality, in addition to the implications in terms of incentives just noted. The financial risks should be explicitly discussed at the outset, and there is a similar need to regularly reassess them.</p>
<p>5. Analyze and discuss how RBF will work with and reach out to all actors in the sector (including schools and subnational government)</p>	<p>5A. Significant results require a large number of agents to act, and cooperate; program documents should spell out plausible reasons why the incentive should ensure agents work together. The accountability arrangements in the sector should be continually assessed, as they are not static.</p>	<p>The RBF program and its financial rewards operate in a complex landscape of preexisting (and evolving) incentives and relationships. The RBF reform focus and associated DLIs may require action at different system levels. Actors at different levels may be reached by the financial incentive (a reward to the system or specific level) and/or may need to pass on the incentive or “transmit the pressure”.</p> <p>Not all actors whose action is required can or should be reached through financial incentives, but they still need to be reached. Unless it is acceptable to government and well-designed, cascaded RBF incentives can cause confusion and even demotivation. Even where there are cascaded financial incentives, these may not be enough; the understanding, PFM, political economy and capacity may all also need to be addressed. For example, strong communication may be a critical strategy to ensure successful policy implementation at school level, with or without other incentives.</p>
	<p>5B. Local-level stakeholders’ engagement and action is essential to achieve results, but mechanisms may vary in how much voice and autonomy they have to act and contribute. Where they have little of both voice and autonomy, RBF may not enhance this, but should seek to strengthen the work of local stakeholders and combine with non-RBF approaches to local accountability.</p>	<p>Horizontal accountability – at subnational level and especially school/community level – is often weak in education systems. Whilst RBF may not lend itself to incentivizing improvements at this level, given the need for measurable indicators, RBF program designs can support and strengthen the work of local stakeholders and non-RBF approaches to strengthen local accountability. For instance, program monitoring can include feedback loops from local-level DLI agents and those that are being impacted by the measures that are being incentivized. This will also enrich the evidence basis enabling RBF stakeholders to assess program performance.</p>
	<p>5C. Consideration should be given to how the RBF design, including DLI selection, supports and/or interacts with <i>de facto</i> and <i>de jure</i> levels of decentralization.</p>	<p>Each context is different but an important dimension is the degree of decentralization to local levels, and the specific responsibilities at different levels both <i>de jure</i> and <i>de facto</i> (e.g. for teacher management / for where to prioritize infrastructure development / for classroom pedagogy etc.)</p> <p>There is a potential trade off in RBF design, to support genuine decentralization or to support centrally determined policy and strategy. Even in cases where decentralization is not a genuine government priority, an RBF design should consider how to foster local government alignment and buy in, and/or more local autonomy in the use of</p>

Recommendation	Connected lessons	Explanation
		resources. This may involve collaborative design in the target-setting / measuring process; incentive cascading; and support to decentralized actors. It also may make certain types and levels of DLIs more desirable, and influence the design of any subnational RBF.
6. Build adaptive management in from the start	6A. The majority of DLIs have a “sell by date”, i.e. may become redundant, irrelevant, or counter-productive, and this is most true for process and output DLIs. Actors should be proactive when there is a need to remove DLIs, revise targets, or add new DLIs.	<p>Planning for five or more years of target results is difficult. Some DLIs will lose relevance, others may be easily achieved and the “bar could be raised” to retain stretch and relevance. For other DLIs the targeted results chain may no longer hold (i.e. assumptions made at design may turn out to be incorrect), due to external changes or new evidence or evaluation, with implications in terms of realism / stretch.</p> <p>A good process should provide regular review of these points and anticipate change, while maintaining consensus. When new DLIs are agreed to be useful during the programming period, a thorough design process should also be undertaken, including an assessment of the trade-off associated with selecting specific areas of focus rather than others.</p>
	6B. A formalized, multilateral, transparent and regular process helps to enable adaptive management; where this is not present, the processes for making changes to DLIs can be ad hoc and take time.	<p>The RBF dialogue and management process should involve regular review of the DLIs to be timely and reflected in program design. Such a review should be an essential part of the terms of reference of any overarching program review period (for example mid-term reviews). Ideally, annual reviews should provide at least a “sense check” on DLI relevance, beyond merely focusing on DLI result reporting. Processes should cover all donors operating RBF where possible. Donors also need to think through how changes once agreed can be made quickly in their programs, avoiding overly bureaucratic delays.</p>
	6C. The method to react to <i>force majeure</i> situations could be more clearly stipulated – i.e. external or large-scale changes to context that have effects across all DLIs – and with processes to manage change across multiple donors. Any of the parties, including government, should have the means to trigger a reassessment in light of any new <i>force majeure</i> situation.	<p>With any programming over several years, shocks will occur, that must be addressed promptly hence outside of the regular, time-fixed review process. Some shocks will have an impact for specific DLIs, while others may impact the feasibility / complexity of all DLIs. These latter <i>force majeure</i> situations require a rapid and structured response to adapt RBF programming. There should be a clear process of consensus in identifying when such situations have occurred (for example natural disasters, major droughts, or major constitutional change etc.), with clear processes for program revisions if necessary. This may be more complex to deal with when there are multiple donors using RBF, and mechanisms should ideally cover this coordination.</p>
7. Seek to use RBF in line with principles of good donor coordination	7A. Multiple donors may be involved in RBF for a particular sector plan, and should work together in design to reduce contradictions and/or duplication for government, for example with program or indicator reporting.	<p>Donors using RBF at the same time, as well as non-RBF donors, bring their own priorities, their own individual program designs, internal processes and systems. This has the potential to create duplication and cost for all sides. Working together is therefore critical to minimise cost, ensure that efforts are complementary and not competing, that demands on government are realistic and proportionate, and to avoid the risk of diluting attention through the sheer number of different incentives.</p>
	7B. Where multiple donors and sectors use RBF at the same time, the risks to government in terms of finance flow can be multiplied. Donors should ideally work together to understand the	<p>There are financial / disbursement risks created by any individual donor when using RBF. Where there are multiple donors doing so or even where there are few donors, RBF can represent a large volume of aid funding, and this can create very big risks for government in terms of volatility in the flow of aid finance. Depending on the context of domestic financing and the degree of aid dependency, donors should consider the collective risk they bring when</p>

Recommendation	Connected lessons	Explanation
	collective risks that RBF poses including whether the scale of use is proportionate.	considering how much funding to put at risk for RBF, which may greatly outweigh the risk brought by the single contribution of any individual bilateral donor.
	7C. RBF links to, aligns with and potentially can strengthen the role of sector coordination and dialogue, including with non-RBF donors and projects. However, there is a risk that RBF/DLIs can come to dominate dialogue when used by multiple donors in a way that crowds out other important results.	Ideally, sector forums such as the local education group for dialogue and review will cover DLI and non-DLI results, and provide a basis to more systematically plan complementarity with other initiatives. RBF programs should not dominate the dialogue, but can leverage and strengthen broader (beyond program) coordination mechanisms. Where the use of DLIs is significant by many partners there is a real risk that it can dominate sector dialogue, which at times this may be appropriate, it may “crowd-out” vital dialogue on non-incentivized results. While maintaining reporting and transparency on joint DLI frameworks, it may be important to ensure transparency across all partners for bilateral agreements and frameworks, to ensure alignment of RBF with non-RBF support.
8. Design the interaction of RBF with evidence systems so as to ensure monitoring of program performance beyond DLI reporting	8A. Plans and strategies to improve data systems are likely to be in place but with objectives beyond just RBF. The use of RBF should be clearly incorporated in these plans including data collection, the use of data in planning and policy, and the trade-offs and risks in creating high-stakes data.	Rigorous, accurate and timely information is vital for good policy. EMIS systems are foundational for evidence-based policy and strategic planning. Other information systems are likely to exist, for example learning assessment / exam results and school supervision systems, separate systems for subsectors such as infrastructure programming or scholarship provision, and financial data systems. Data and information can and should be improved, even when the measurement of DLI results does not draw on any of these systems. Improving evidence more broadly also requires paying attention to interfacing systems and data. The channels by which data and evidence feeds into policy-making processes should also be understood and continually assessed. This also needs to be resourced, as well as when required, support to syntheses of evidence and formulation of policy options/ strategies/ programs, responding to what the evidence indicates.
	8B. The environment for communication and transparency in evidence systems and processes is a key component for how much learning and improvement to policy is likely to flow from RBF.	There is a need to pay attention to and support government in strategies to communicate results, including when results are adverse. As this is tricky for all governments, especially on sensitive matters (e.g. learning), and may be trickier still in highly polarized political contexts, it is important to engage with political economy issues around transparency. National systems of performance audit may be an important component of this.
	8C. The use of a broad base of evidence in assessing program performance can be vital in understanding sector trends, including non-incentivized information, evaluation, and learning processes.	Ensuring and encouraging separate strands of research and evidence that do not directly feed DLI results, i.e. non-incentivized evidence, may be essential to provide a “reality check” on sector changes. Close attention needs to be paid to continuously monitoring not just the DLI results, but also “why” the results are what they are and whether the assumptions underpinning the link between them and higher-level goals are still holding, as well as the ongoing monitoring of important non-DLI results. There may be a case for RBF programs to incentivize not just results but also policy learning, through DLIs that would reward research, analyses, and policy / strategy development on issues arising in relation to the program “results DLIs” and broader sector trends relevant to the program focus. Data can be generated in a number of complementary ways, including commissioning research focused on issues arising from implementation and/or on weak / less well established links in the program or sector’s theories of change, identifying and drawing from relevant third-party research and evidence, integrating data from existing

Recommendation	Connected lessons	Explanation
		<p>separate systems (including “beyond-education” data e.g. on poverty), and evaluations of specific programs and projects which are particularly important for assessing whether strategies (e.g. for infrastructure) are efficient, effective and offer value for money.</p>
<p>9. Carefully design and fund the verification system</p>	<p>9A. It can be challenging to ensure the independence of verification, including in the designs for the scope of work, processes of reporting, and joint management and buy-in from government and donors. A variety of approaches have been used with different trade-offs.</p>	<p>Verification needs to be done in a way that it is clear in terms of who will define the scope of work and how, who the independent verification will report to, and the transparency of reporting as well as processes of feedback (including on specific sampling methodologies, how many schools to sample, which tools to use, etc.).</p> <p>If the verification reports to government “first” for “sign-off” then in practice it will not be independent. If scope of work and reporting are joint then this may mitigate risks to the independence of the verification agent, while ensuring maximum buy-in to the credibility of the process and the potential uses of any research undertaken. Multiple donors may complicate the process, but in all cases the design phase should include consideration of all of these trade-offs.</p>
	<p>9B. The amount of funds allocated, and the source of funds, to the verification process should ideally be clear and transparent at all stages. However, this is often not the case. Sufficient funds should be considered or face trade-offs with quality, sample size, rigor, and the timeline.</p>	<p>As with any process of research, there will be trade-offs between the amount of resources allocated to verification, and the rigor and quality of the process. The capacity and skills of the verification agent are of course equally important, but when combined with the trade-offs, parties to the RBF contract need to be clear where to set their expectations, including in how “credible” the process will ultimately be. It should also ideally be clear how much is allocated, where it is coming from, and how competitive tendering will be done (assuming a single source government agency is not preferred).</p> <p>The definition of verification protocols can be useful, but will depend on the nature of the DLI. For simpler process indicators, cheaper, “good enough” evidence, may work. For more complex outcome indicators, sample size, representativeness, and survey tool design, as well as enumerator skills and training, will all be important drivers of quality. Trade-offs need to also consider the risk that the more rigorous the protocol, the more time will be spent in formulating it, reaching agreement between government and funders, and in cases in which the verification process challenges government reporting, time spent in “arguing” about it.</p>
	<p>9C. The verification role may or may not include a role for qualitative research and evaluation, i.e. going beyond strictly “verification”. For this to be effective clear expectations are required as well as incentives for all parties to learn from the experience rather than treating it as a bureaucratic exercise.</p>	<p>The RBF program independent verification process is one among several potential sources of evidence. Any expectations for this process to “go beyond verification”, for example conducting performance evaluation, original primary research, or promoting the learning content of this process, should be very carefully designed.</p> <p>In practice, the learning role and evaluative content of the verification role will link to government willingness and the degree of flexibility in disbursement in cases of problematic lessons (i.e. if lessons question the quality of performance of a particular DLI). Donors should be aware that the stricter they are in “not paying”, the less they will incentivize learning from the verification process itself, especially if government has a significant role in selecting and/or liaising with the verification agent, particularly if government does not “buy into” the way in which the independent verification scope of work is defined.</p>
<p>10. Use a combination of process, output and intermediate outcome indicators in a</p>	<p>10A. Process reforms that are socially complex and require extensive change at local level, take time and require alignment and buy-in from a wide range of country stakeholders. They may not</p>	<p>Meaningful process reform is only possible with buy-in and ownership from government stakeholders. This may include a mixture of central agencies, as well as politicians, and may be outside the control of the MoE or other sector agents to complete. Depending on the reform, they may also involve a wide array of subnational government and agencies, which require buy-in, capacity building and/or technical support.</p> <p>When DLIs with such greater complexity are desirable, time and resources are likely to be needed to adequately communicate the complexity, and ensure attention to results themselves and to planning.</p>

Recommendation	Connected lessons	Explanation
<p>way that helps reformers</p>	<p>easily lend themselves to time-bound DLIs.</p>	<p>Process indicators may require a simple metric (e.g. “Policy/strategy developed”), but relate to a more complex change involving multiple actors. In this case, the progress of reform, including whether all actors effectively align behind the change intended, should be closely monitored. Lessons from the use of RBF in other sectors should be carefully monitored.</p>
	<p>10B. Output indicators require clear processes of design and clear project management responsibilities, as well as understanding and measurement of output <i>quality</i> in addition to more common measures of <i>quantity</i>.</p>	<p>Output indicators may incentivize outputs that are already occurring (for example if teachers are being trained, outlining a <i>number</i> of teachers to be trained as a target), or new outputs (for example teachers to be trained in a <i>new</i> methodology). In the latter cases there will again be a need for buy-in and ownership into the idea that the new direction is the right one, as well as clear project management (in the example above, <i>who</i> will design the <i>new</i> methodology, <i>who</i> will manage its rollout), and the degree of donor support with this via technical assistance or other. Specific targets for new outputs set before the output has been designed (e.g. how many teachers to train in the new methodology) should be done with caution, as the costs and complexity of rollout at the outset may be unknown. Particular attention should go into whether outputs have a quality measure in addition to any quantity measure, otherwise DLI achievement can become superficial representations of results or system changes.</p>
	<p>10C. For results that constitute intermediate outcomes, it may be better to have a clear and simple metric (e.g. PTRs) than complex process, output or proxy indicators.</p>	<p>Intermediate outcome indicators may be more likely to spur autonomy and innovation over time, but are closer to outcomes in being challenging to achieve (socially and technically) and to measure. As such, clear and simple metrics should be used, ideally drawing on established data sources, though while assessing the risk of potentially distorting the data source through the RBF incentive; and as suggested earlier, thoroughly triangulating the DLI result measures with related information from other sources.</p>
<p>11. Be careful, and frugal in the use of outcome indicators.</p>	<p>11A. For outcome indicators, their complexity and often low degree of control and knowledge on how to achieve them bring particular risks (stretch too high, risks to disbursement, and demotivation and/or to transparency). While outcomes are desirable to achieve, careful attention should be paid to these risks.</p>	<p>Student outcomes, particularly in terms of learning, are perhaps the best expression of education system goals, and are therefore likely to be considered for selection as DLIs. However, they are also the most complex to achieve and may not be in the likely discretion or control of the agent over the three- to five-year life of an RBF instrument. This makes them risky, including in the degree of actual and perceived control, uncertainty in achievement and potential demotivation and political risks if they are not achieved, as well as disbursement risk. Where used and if targets for results are above trend (i.e. have a degree of ambition) it could be considered to present the outcome DLI from the outset as a “challenge indicator” – a nice-to-have, but higher risk, portion of the overall financing instrument.</p>
	<p>11B. Outcome indicators can be difficult to define and measure, particularly where only one metric is used. Incentivizing the measurement of outcomes brings lower risks than incentivizing the results of outcome measures; and may therefore provide a better basis for policy learning.</p>	<p>There are risks in measuring a DLI outcome indicator with one metric or evidence source, particularly if the disbursement stakes are then high when linked to the results of the outcome as measured. Gaming risks, or the use of an incomplete, too narrow or otherwise imperfect metric then have a double risk – disbursement on results that are not what they seem, and providing a false sense of success, which may take away attention or bring complacency on extremely important areas. Given their complexity, one metric or source of evidence will never be “enough” to understand outcome progress.</p> <p>Measurement and the outcome definition can be complex, and therefore there is a need to ensure sufficient time for consensus and understanding on measures, including at subnational levels if they are also incentivized. If only the measurement is incentivized and not the results, the associated risks are reduced, i.e. by making the outcomes less “high-stakes”.</p>

Recommendation	Connected lessons	Explanation
	<p>11C. Given their importance, non-incentivized measures of outcome indicators are vital, as is transparency in discussion and learning on these results.</p>	<p>Complexity for outcome indicators brings particular challenges, as above. With complex indicators, complex measurement and verification methodologies and protocols for the DLIs, the risk increases that the RBF dialogue will be dominated by discussions about the methodology and measurement and verification standards, at the detriment of attention to results and “why”. In these instances, there is a risk of crowding out time for considering evidence on non-directly RBF-related results and trends.</p> <p>To respond to this risk, and in line with the recommendations above on evidence systems, it is extremely important to build in non-incentivized measures for comparison and deeper understanding of outcomes. This may create conflicts of evidence, for example measures conflicting with “official” DLI metrics, and this risk shows why outcome indicators should be handled very carefully, in order to avoid false narratives and to create unintended consequences to disincentivize attention to learning.</p>
<p>12. Pay continual attention to equity</p>	<p>12A. Equity should be a consideration in RBF design and implementation including in terms of the different capacity of agents, particularly at subnational level. There can be structural inequalities in resource allocation by geographical or other criteria and there is a risk that RBF can exacerbate them.</p>	<p>The differential ability (due to capacity, or context) of agents (e.g. of local governments, or of schools) to implement actions required to achieve targets, creates a tension between rewarding performance and ensuring equity. Even when financial incentives do not cascade to these agents, their overall ability to reach targets can be undermined when differential ability is not considered. It matters even more when incentives are cascaded as it may compound prior imbalances in resources.</p> <p>As a result, targets may need to be calibrated for the reward system to be “fair” / reflect the diversity of challenges faced by different agents at the same level. Time and resources are required to make the DLI reward system as well-understood as possible.</p>
	<p>12B. Program-level implications for equity during design should be treated as a “crosscutting” programmatic issue, and continue to be so during implementation.</p>	<p>There are likely to be significant political economy considerations at design, concerning the focus of DLIs and the balance across the program as a whole. In any context, the overall implications of the RBF program for equity should be considered; including a crosscutting appraisal of equity implications of all incentivized changes. This should include an analysis of “cherry-picking” risks for DLIs for which this is relevant (risks of prioritizing more advantaged, less costly to work with, easier to access, or better off groups to achieve targets).</p> <p>Equity risks should be regularly reviewed, in the same way as other RBF risks. Wherever possible, DLI targets or reward calculations should be designed so as to address/ mitigate these risks.</p>
	<p>12C. While it may be beneficial to have DLIs specifically addressing equity, it is also important to look out continually for unintended equity consequences of all DLIs.</p>	<p>Beyond assessing and explicitly addressing the equity implications of any DLI, there may be room for specific DLIs that bring a clear focus on equity, including for example addressing marginalized groups or regions if there are significant such equity challenges. Careful consideration should be given to whether and how the RBF DLIs can or should address these. In some contexts, RBF may not be the best mechanism.</p> <p>Caution should be applied that by incentivizing one particular aspect of equity (e.g. within-province distribution of teachers), broader and more important dimensions of equity may be lost (e.g. between-province distribution of teachers), and therefore even explicitly equity-focused DLIs need careful consideration.</p>

References

Short reference	Full reference	Location
Acharya & León-González, 2018	Acharya, Chakra Pani & Roberto León-González. 2018. The Quest for Quality Education: International Remittances and Rural-Urban Migration in Nepal. National Graduate Institute for Policy Studies, Japan. February 2018.	link
Adil and Sabarwal (2017)	Adil, M., and Sabarwal, S. (2017). <i>Twelve Ideas for Effective Education PforRs</i> . World Bank.	link
Angrist et al. (2020)	Angrist, Noam, David K. Evans, Deon Filmer, Rachel Glennerster, F. Halsey Rogers, and Shwetlena Sabarwal (2020). <i>How to Improve Education Outcomes Most Efficiently? A Comparison of 150 Interventions Using the New Learning-Adjusted Years of Schooling Metric</i> . CGD Working Paper 558. Washington, DC: Center for Global Development.	link
Bandstein (2007)	Bandstein (2007). <i>What Determines the Choice of Aid Modalities? – A framework for assessing incentive structures</i> . Swedish Agency for Development Evaluation.	link
Bassi et al. (2019)	Bassi, Marina, Octavio Medina Pedreira, and Lucia Jose Nhampossa (2019). <i>Education Service Delivery in Mozambique: A Second Round of the Service Delivery Indicators Survey (English)</i> . Washington, DC: World Bank Group.	link
Cilliers et al. (2018)	Cilliers, Jacobus, Isaac Mbiti, and Andrew Zeitlin (2018). <i>Can public rankings improve school performance? Evidence from a nationwide reform in Tanzania</i> . RISE Working Paper. December 2018.	link
Clist (2016)	Clist, Paul (2016). "Payment by Results in Development Aid: All That Glitters Is Not Gold." <i>The World Bank Research Observer</i> , 31(2), 290–313.	link
Clist and Verschoor (2014)	Clist, Paul & Arjan Verschoor (2014) <i>The Conceptual Basis of Payment by Results</i> . University of East Anglia, Norwich, UK. Research for Development, DFID.	link
Cuevas-Rodríguez et al. (2012)	Gloria Cuevas-Rodríguez, Luis R. Gomez-Mejia, and Robert M. Wiseman (2012). <i>Has Agency Theory Run its Course?: Making the Theory more Flexible to Inform the Management of Reward Systems</i> . <i>Corporate Governance: An International Review</i> , 2012, 20(6): 526–546.	link
DANIDA (2014)	Ministry of Foreign Affairs of Denmark, International Development Cooperation (DANIDA) (2014). <i>Review of Budget Support Evaluations</i> . Evaluation Study. 2014/1.	link
Development Cooperation Ireland and Mokoro (2005)	DCI (Development Cooperation Ireland) (2005). <i>Aid Modalities in Ethiopia</i> . Dublin: Development Cooperation Ireland and Mokoro.	link
Dom et al. (2020)	Catherine Dom, Alasdair Fraser, Edward Mhina (2020) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Tanzania</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
Dom et al. (2021)	Catherine Dom, Alasdair Fraser, Edward Mhina (2021) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Tanzania</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
EC et al. (2013a)	EC, Danida, Irish Aid, Ministry of Foreign Affairs of the Netherlands, Ministry of Finance Tanzania (2013). <i>Joint Evaluation of Budget Support to Tanzania: lessons learned and recommendations for the future – Final Report: Volume 1</i> .	link
EC et al. (2013b)	EC, Danida, Irish Aid, Ministry of Foreign Affairs of the Netherlands, Ministry of Finance Tanzania (2013). <i>Joint Evaluation of Budget Support to Tanzania: lessons learned and recommendations for the future – Final Report: Volume 2</i> .	link
Evans and Acosta (2020)	David Evans and Amina Mendez Acosta (2020). Education in Africa: <i>What are we learning?</i> Center for Global Development, Working Paper 542. August 2020.	link

EY (2018)	Ernst & Young (EY) (2018). Ministry of Education: Education Sector Support Fund (FASE) Financial Reports for the year ended 31 December 2017. Audit Report.	
GAN Integrity (2020)	GAN Integrity (2020). Mozambique Corruption Report, regularly updated, accessed October 2020.	link
GPE (2019)	Global Partnership for Education (GPE) (2019). "Guidance Note on GPE Variable Part Financing." March 2019.	link
Holden et al. (2020)	Joseph Holden, Frances Hitchens, Yadab Chapagain (2020) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Nepal</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
Holden & Chapagain (2021)	Joseph Holden and Yadab Chapagain (2021) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Nepal</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
IMF (2016)	IMF (2016). Republic of Mozambique, Selected Issues. IMF Country Report No. 16/10.	link
IMF (2019)	IMF (2019). Republic of Mozambique : 2019 Article IV Consultation-Press Release; Staff Report; and Statement by the Executive Director for the Republic of Mozambique.	link
KfW (2017)	KfW (2017). <i>Ex post evaluation – Mozambique</i> . Project: Education SWAp ESSP FASE III-VI DC programme.	
KPMG (2019)	KPMG (2019). Ministry of Education: Education Sector Support Fund (FASE) Financial Reports for the year ended December 31, 2018. Audit Report.	
KUSOED (2017)	KUSOED (2017). <i>SSDP DLI Year 1 Verification Report Part II. Independent Verification Report of SSDP Year 1 Disbursement Linked Indicators (DLIs) (Part II)</i> . Submitted to Department of Education (DOE), Sanothimi, Bhaktapur. Submitted By Kathmandu University School of Education (KUSOED), December 29, 2017.	
Lee and Medina Pedreira (2019)	Lee, Ling Jessica Diana and Medina Pedreira, Octavio (2019). <i>Results-based Financing in Education : Learning from What Works (English)</i> . Washington, DC: World Bank Group.	link
Macamo (2006)	Macamo, Elísio (2006). <i>Political Governance In Mozambique Final Report</i> By (For DFID – Mozambique). June 2006.	link
MINEDH (2016)	MINEDH (2016). <i>Desempenho do sector da educação, 2015</i> . Ministry of Education and Human Development. Annual Review Meeting (RAR). 2016.	
MINEDH (2019)	MINEDH (2019). <i>Análise do Sector de Educação (ESA). 2019. Relatório Final</i> . Moçambique, Julho de 2019. UNESCO.	
MINEDH, (2020)	MINEDH (2020). <i>Plano estratégico da educação Moçambique (Mozambique Education Sector Strategic Plan)</i> . 2020-2029.	link
MoEST Nepal (2017)	Ministry of Education (2017). <i>Teacher rationalization and deployment plan</i> . Annex 2 to SSDP DLI Achievement report Year 1.	
MoEST Nepal (2019)	Ministry of Education, Government of Nepal (2019) Draft Consolidated Flash Report 2018-2019	
MoEST & PO-RALG (2019b)	United Republic of Tanzania, Ministry of Education, Science and Technology and President's Office – Regional Administration and Local Government. Education Program for Results (EPforR) Annual Report 2018-2019. Final, November 2019.	
MoEST & PO-RALG, 2020a	Ministry of Education, Science and Technology and President's Office – Regional Administration and Local Government. Tanzania Education Program for Results (EPforR) Programme Operations Manual. Revised June 2020.	
MoEST Nepal (2020)	Ministry of Education / Local Education Group / Joint Financing Partners. BRM May 2020 Aide Memoire, and supporting documents.	
Norwegian Institute of International Affairs (2017)	Norwegian institute of international affairs (2017). <i>Mozambique: A political economy analysis</i> . Aslak Orre & Helge Rønning, October 2017.	link
Nuvunga and Orre (2019)	Nuvunga, Adriano and Orre, Aslak (2019). "The 'secret loans affair' and political corruption in Mozambique". Chapter 6 in <i>Political Corruption in Africa Extraction and</i>	link

	<i>Power Preservation</i> . Edited by Inge Amundsen. Published in print: November 4, 2019. ISBN: 9781788972512.	
OPM (2018a)	Oxford Policy Management (OPM) (2018). <i>Education in a Decentralized Context: The case of Nepal</i> . Shrochis Karki, Pramod Bhatta. July 2018.	
OPM (2018b)	Oxford Policy Management (OPM) (2018). Operational Research Assignment on the World Bank's PFM for Results Program in Mozambique, 2018. Rogerio de Sa.	
OPM (2019a)	Rawle, G., Binci, M., Gelander, G.P., Harb, J., Jasper, P., Khan, S., Medardi, D., Romarri, A., Rorich, M. and Ruddle, N. for Oxford Policy Management (2019). EQUIP-Tanzania Impact Evaluation Endline Quantitative Technical Report, Volume I Results and Discussion. January 25, 2019.	
OPM (2019b)	Rawle, G., Binci, M., Gelander, G.P., Harb, J., Jasper, P., Khan, S., Medardi, D., Romarri, A., Rorich, M. and Ruddle, N. for Oxford Policy Management (2019). EQUIP-Tanzania Impact Evaluation Endline Quantitative Technical Report, Volume II Methods and Supplementary Evidence. January 25, 2019.	
Orlowski (2016)	Orlowski, D (2016). FASE Review Part 2.	
Patch et al. (2020)	John Patch, Joseph Holden, Francis Ruhumbika (2020) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Preliminary Report, Mozambique</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
Patch et al. (2021)	John Patch, Joseph Holden, Francis Ruhumbika (2021) <i>Results-Based Financing in the Education Sector: Country-Level Analysis. Final Assessment Report, Mozambique</i> . Submitted to the REACH Program at the World Bank by Mokoro Ltd.	
Paul (2015)	Paul, Elisabeth (2015). "Performance-Based Aid: Why It Will Probably Not Meet Its Promises." <i>Development Policy Review</i> , 2015, 33 (3): 313–323.	link
Pérez Niño and Le Billon (2016)	Pérez Niño, Helena and Le Billon, Philippe (2016). "The effects of foreign aid on political settlements: the case of Mozambique and the contrast of Angola." In Hagmann, T. and F. Reyntjens (Eds). <i>Development without democracy: foreign aid and authoritarian regimes in Africa</i> . Zed Books.	
Poyck et al. (2016)	Poyck, Maria Catharina, Dr. Bidya Nath Koirala, Dr. Prem Narayan Aryal, Ca. Nanda Kishor Sharma (2016). Joint Evaluation of Nepal's School Sector Reform Plan Programme 2009-2016. Contract n° 2014/357774 funded by European Union. GFA Consulting Group GmbH.	link
Pradhan et al. (2014)	Pradhan, M., Suryadarma, D., Beatty, A., Wong, M., Gaduh, A., Alisjahbana, A., & Artha, R. P. (2014). Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia. <i>American Economic Journal: Applied Economics</i> , 6(2), 105–126.	link
Pritchett (2015)	Pritchett, Lant (2015). Creating Education Systems Coherent for Learning Outcomes: Making the Transition from Schooling to Learning. RISE-WP-15/005 December 2015.	link
Quinn Patton (2008)	Michael Quinn Patton (2008). <i>Evaluating the Complex: Getting to Maybe</i> . Powerpoint. 29 May, 2008	link
Rasul and Rogger (2018)	Rasul, I. and Rogger, D. (2018). "Management of bureaucrats and public service delivery: Evidence from the Nigerian civil service." <i>The Economic Journal</i> 128 (608): 413-446.	link
Rasul et al. (2018)	Rasul, I., Rogger, D., and Williams, M. (2018). <i>Management and Bureaucratic Effectiveness: Evidence from the Ghanaian Civil Service</i> . World Bank Policy Research Working Paper. No. 8595	link
Sabarwal et al. (2019)	Sabarwal, Shwetlena; Mohan Aryal; Maya Sherpa; Uttam Sharma; Unika Shrestha; and Laxman Timilsina (2019). "The Promises and Perils of Education Decentralization in Nepal." Draft Policy Report. World Bank. Additional JRM Documents #6.	
Schmidt (2006)	Petra Schmidt (2006). "Budget Support in the EC's Development Cooperation." <i>Studies Deutsches Institut für Entwicklungspolitik</i> . Bonn 2006.	link
SOFRECO and FBC, 2019	SOFRECO and FBC (2019). School Sector Development Plan (SSDP) Mid-term Review (MTR).	

Tilley and Tavakoli (2012)	Helen Tilley and Heidi Tavakoli (2012). <i>Better aid modalities: are we risking real results?</i> Literature review. Overseas Development Institute (ODI).	link
Todd and Attfield (2017)	Todd, Robin and Ian Attfield (2017). <i>Big Results Now! in Tanzanian education: Has the Delivery Approach Enabled Teachers & Delivered Learning?</i>	link
TU-CERID (2019)	DLIs Verification Report Submitted to Ministry of Education, Science and Technology (MOEST), Center for Education and Human Resource Development (CEHRD) Sanothimi, Bhaktapur. Tribhuvan University, Research Center for Educational Innovation and Development (TU-CERID) Balkhu, Kathmandu. January 2019.	
Turner et al. (2019)	Turner, Stephen, Christine Fenning, and Yadab Chapagain (2019). GPE 2020 Country-level Prospective Evaluations. First Annual Report: Nepal. April 2019. Global Partnership for Education. Mokoro Ltd.	link
UNESCO (2017)	UNESCO (2017). Global Education Monitoring Report 2017/8: Accountability in education: meeting our commitments.	link
UNESCO (2018)	UNESCO (2018). Mozambique Education Policy Review. January 2018.	
UNESCO (2019)	UNESCO (2019). Global Education Monitoring Report 2019. April 2019.	
UNICEF (2017)	UNICEF (2017). Mozambique Budget Brief: Education 2017.	link
UNICEF (2018)	UNICEF (2018). Mozambique Budget Brief: Education 2018.	link
Universalialia (2019)	Universalialia (2019) Summative GPE country program evaluation, Mozambique. Final evaluation report. July 2019. With Results for Development and ITAD.	
Williams (2017)	Williams, M. (2017). "The Political Economy of Unfinished Development Projects: Corruption, Clientelism, or Collective Choice?" <i>American Political Science Review</i> 111(4), p. 705-723.	link
World Bank (2013)	World Bank (2013). "Nepal, Small Area Estimation of Poverty, 2011." Vol: I. The World Bank and Government of Nepal National Planning Commission Secretariat Central Bureau of Statistics.	link
World Bank (2014)	World Bank (2014). International Development Association Program Appraisal Document on a Proposed Grant in the Amount SDR 33 Million (USD 50 Million Equivalent) to the Republic Of Mozambique for the Public Financial Management For Results Program. May 30, 2014. Report Number 72624-MZ.	link
World Bank (2015a)	World Bank (2015). Mozambique Service Delivery Indicators (SDI): Education. March 2015.	link
World Bank (2015b)	World Bank (2015). Project Paper On A Proposed Additional Grant In The Amount Of SDR 36 Million (USD 50 Million Equivalent) And A Global Partnership For Education (GPE) Grant In The Amount Of USD 57.9 Million To The Republic Of Mozambique For An Education Sector Support Project. July 2, 2015.	link
World Bank (2016)	World Bank (2016). "What Works? Program-for-Results - An Early-Stage Assessment of the Process and Effects of a New Lending Instrument." An Independent Evaluation. IEG, World Bank.	link
World Bank (2017a)	World Bank (2017). "Results-based financing in education: Financing results to strengthen systems." Peter A. Holland and Jessica D. Lee.	link
World Bank (2017b)	World Bank (2017). "Bank Policy. Program-for-Results Financing." Catalogue Number OPS5.04-POL.107 Issued November 10, 2017.	link
World Bank (2017c)	World Bank (2017) Implementation Completion And Results Report. Nepal. SSRP. March 15 2017 Report No: ICR00003994	link
World Bank (2019a)	World Bank (2019). Tanzania Mainland Poverty Assessment. World Bank.	link
World Bank (2019b)	World Bank (2019). Implementation Completion and Results Report for the Mozambique Public Financial Management for Results Program. Program-For-Results. International Development Association. November 7, 2019. Governance Global Practice Africa Region, the World Bank.	

World Bank (2019c)	World Bank (2019). International Development Association Restructuring Paper on a Proposed Restructuring Of Tanzania Education Program For Results (Epffor) Approved on July 2014 to the United Republic Of Tanzania. September 12, 2019.	
World Bank (2020)	World Bank (2020). "Improving Learning and Empowering Girls in Mozambique". International Bank for Reconstruction and Development + International Development Association. Project Appraisal Document. Project P172657.	