

52nd GEF Council Meeting
May 23 – 25, 2017
Washington, D.C.

EVALUATION OF PROGRAMMATIC APPROACHES IN THE GEF

VOLUME I – MAIN REPORT

(Prepared by the Independent Evaluation Office of the GEF)

VOLUME I: MAIN REPORT

TABLE OF CONTENTS

Executive Summary.....	v
Abbreviations and Acronyms.....	viii
I. Introduction	1
A. Background and Objectives	1
B. Approach and Methodology.....	2
C. Limitations	5
II. Context	6
A. Programs Evolution, Typologies and Definitions	6
B. Portfolio.....	6
C. GEF versus Comparable Donor-based Programs: Similarities and Differences.....	9
III. Findings	12
A. Programmatic projects compared with stand-alone projects.....	12
B. Broader and Longer-term Programmatic Results.....	18
C. Addressing Drivers of Environmental Degradation through Programs	22
D. Ownership	23
E. Coherence	27
F. Cost-effectiveness, Efficiency and Coordination issues.....	29
G. Monitoring and Evaluation	35
H. Governance	38
IV. Conclusions and Recommendations.....	40
A. Conclusions.....	40
B. Recommendations.....	41
Annexes	
<i>(Annexes are provided separately)</i>	
1. Approach Paper	
3. Methods and Tools	
4. Portfolio	
5. List of Interviewed Stakeholders	
6. Countries and Sites Visited	
7. References	

Figures

Figure 1: Program and Child Project Geographic Scope by Share of GEF Grant.....	7
Figure 2: Program/Project Focal Area (#)	8
Figure 3: Program/Project Focal Area (% GEF Grant)	8
Figure 4: Program/Project Agency by GEF Grant.....	8
Figure 5: APR rating comparisons highlights	12
Figure 6: Comparing APR ratings by program complexity	13
Figure 7: Incentives and disincentives to join a program	14
Figure 8: Heterogeneity in remote sensing findings along relevant dimensions (I)	15
Figure 9: Heterogeneity in remote sensing findings along relevant dimensions (II)	15
Figure 10: Heterogeneity in remote sensing findings along relevant dimensions (III)	16
Figure 11: Heterogeneity in remote sensing findings along relevant dimensions (IV)	16
Figure 12: Spatial distribution of NDVI around the Coringa Wildlife Sanctuary, India (2015).....	17
Figure 13: Comparing the average May-Aug NDVI Al Huseinieh Reserve, Jordan (2013 – 2015)	17
Figure 14: Broader adoption of project outcomes	19
Figure 15: Country stakeholder perceptions on broader adoption.....	19
Figure 16: Broader adoption of project outcomes by program complexity	20
Figure 17: Factors contributing to broader adoption of project outcomes	21
Figure 18: Typologies of drivers addressed by pre-2008 and post-2008 programs	22
Figure 19: Share of total program costs funded from STAR by geographic scope	24
Figure 20: GEF-4 - Program Use vs. Country allocation (\$ million).....	24
Figure 21: GEF-5 - Program Use vs. Country allocation (\$ million).....	24
Figure 22: GEF-5 - percent of STAR given by countries for programs out of their total allocation	25
Figure 23: Program Components addressed through Child Projects	27
Figure 24: Program Components addressed through Child Projects by complexity	27
Figure 25: Stand-alone and child projects GEF grants and co-financing	31
Figure 26: Co-financing for child projects by source	31
Figure 27: Types of program coordination mechanisms	32
Figure 28: Total coordination budget by program complexity	35
Figure 29: M&E Design Ratings.....	35
Figure 30: M&E Implementation Ratings	35
Figure 31: Coherence of Project and Program M&E and RBM	36

Tables

Table 1: Post-2008 Programs by Geographic Scope and Focal Area	7
Table 2: Comparing GEF programs with major program typologies	10
Table 3: Program and Child Projects document requirement vs. Stand-alone projects	30

Boxes

Box 1: Definitions used in this report	3
Box 2: Program Case Studies	4
Box 3: Mechanisms for Broader Adoption of Project Outcomes.....	18
Box 4: Evidence of Mainstreaming in Two Country Programs	20
Box 5: Absence of scaled-up results of the MENA-DELP	21
Box 6: Tackling Drivers of Environmental Degradation in the India and MENA-DELP programs	23
Box 7: Country Ownership in National and Regional Programs	26
Box 8: Coherence between GEF Programs and their Child Projects.....	28
Box 9: Coordination between Child Projects and Programs.....	33
Box 10: Monitoring and Evaluation at Program and Child Project level.....	37

VOLUME II: TECHNICAL DOCUMENTS

(Volume II with the evaluation Technical Documents is provided separately)

TD1 – GEF Programs and Beyond: A Comparative Analysis

TD2 – Geospatial Impact Analysis of Programmatic Project Implementations in the GEF

TD3 – Global Online Survey

TD4 – Program Case Studies:

1. PRC-GEF Partnership on Land Degradation in Dryland Ecosystems, China
2. India GEF Coastal Marine Program
3. MENA-Desert Ecosystems and Livelihoods Program
4. Reducing Industry's Carbon Footprint in South East Asia

EXECUTIVE SUMMARY

1. Programmatic approaches, formalized in 2008, are particularly relevant to the GEF, given the long-term nature of the environmental problems the GEF addresses. This evaluation assessed the mechanisms and conditions by which GEF programs have delivered broader- scale and longer- term results by comparing them to stand-alone projects. It focused on the extent to which GEF programs addressed drivers of environmental change; performance issues such as coherence, ownership, efficiency, M&E have also been evaluated. The evaluation is based on evidence from a wide array of sources, analyzed with a mixed-methods approach. In this evaluation, complexity is a function of the degree of homogeneity of a program's child projects and whether they belong to one or multiple countries, agencies and/or focal areas. The evaluation covers a total of 38 programs and their related 301 child projects, 63 of which are completed. The ratings evidence is based on 42 project terminal evaluations of which 29 are categorized as belonging to simple and 13 to relatively more complex programs.

2. Following are the key findings of the evaluation.

- (a) **Child projects under programmatic approaches performed better than stand-alone projects that are not part of programs.** Child projects, implemented as part of programs, performed better than stand-alone projects on all dimensions. In addition, in terms of vegetation density and forest cover, child projects have improved local environmental conditions as compared with no GEF interventions, and single focal biodiversity projects provided more benefits than their stand-alone comparators.
- (b) **Complexity matters for outcomes.** A simple regression analysis on 42 completed projects suggests that complexity as measured by multi-country, multifocal, multi-agency dimensions and project heterogeneity, is negatively correlated with outcomes. Based on the sample of closed child projects in complex programs (n=13), these projects underperformed relative to those in simpler programs (n=29) or standalone projects on 5 dimensions including outcomes, M&E implementation, execution quality, effectiveness and efficiency, and outperformed these comparators on implementation, sustainability and M&E design.
- (c) **Program design for broader adoption has improved substantially over time across focal areas, but actions were limited.** International water programs are the only exception, and have shown well-designed programmatic thinking from the early GEF phases. Data as to whether improved design for broader adoption has translated into better performance is not yet available. The available data from terminal evaluations indicate that while child projects rated higher than stand-alone projects on the design for broader adoption, they demonstrated less concrete action for broader adoption during implementation.
- (d) **Programs represent a shift toward a more integrated systemic approach to address drivers.** GEF programs have evolved from a narrow approach focused on mitigating the negative effects of both food and energy production on biodiversity loss, land degradation, and climate change to applying a systemic integrated approach encompassing a wider set of drivers such as food and energy production and consumption, buildings and infrastructure construction, and transportation.
- (e) **Program ownership at the country level is highly linked to the degree of alignment with national priorities.** With the notable exception of programs addressing transboundary issues (i.e. international waters) GEF programs progressively shifted over time from a country to a multi-country focus. STAR funds are a substantial share of total program resources regardless of the

program geographic scope. Central and country level stakeholders stated that country programs have stronger ownership than regional/global ones, as they tend to be closely aligned with national priorities. Country focused programs typically employ more of their STAR allocations and tend to receive higher cofinancing from national budgets. Regional/global programs rely heavily on set-asides.

- (f) **Program/child project coherence has improved in recent programs.** Program objectives are better defined; child projects have improved in design and are better linked to the overall program. This improved coherence of programs and the associated child projects is notable in the design of increasingly complex programs, under which projects more specifically address the outcomes of their programs.
 - (g) **Cost-effectiveness and efficiency decline as programs become more multidimensional.** Overall, based on the terminal evaluations, child projects scored higher on efficiency and leveraged higher cofinancing, but efficiency ratings decline with increased complexity. Child projects don't differ much from standalone in terms of project cycles. GEF Agencies consider simple programs, particularly those composed of homogeneous child projects, as having lower transaction costs and being easier to manage. Most programs involve more than one GEF Agency, but child projects tend to be implemented by a single agency. Due to their diversity in mandates and operational approaches, GEF Agencies often find it challenging to work together. The increased costs in coordinating large complex programs in terms of resource and coordination requirements, are increasingly being addressed through better design and are being resourced to improve knowledge management and coordination.
 - (h) **M&E has improved in the design of recent programs, but still faces challenges.** Child projects achieved higher M&E design ratings compared to stand-alone projects, indicating that child projects tend to be more cognizant in designing their M&E frameworks. However, these projects show weaker implementation of M&E than their stand-alone counterparts. More complex programs have similar M&E ratings to simpler ones, but again their ratings drop from design to implementation. M&E is mainly undertaken at project level. Little evidence of program level M&E has been found. When present, it is most likely because of individual GEF Agency requirements. Early evidence from the Integrated Approach Pilots (IAPs) suggests more attention to M&E design, but the systems have yet to be implemented.
 - (i) **The roles for partners in program design have evolved with changing focus on global programs and IAPs.** While this was not a concrete objective within the scope of this evaluation, partners expressed a need for greater clarity on roles in program formulation.
3. Based on the above findings, the evaluation has reached five main conclusions:
- **Conclusion 1:** GEF programmatic approaches have promoted projects that are better designed to produce broader and more sustainable results than stand-alone ones;
 - **Conclusion 2:** The multidimensional nature of programs has generated greater need for coordination and management, with implications for efficiency, results and performance;
 - **Conclusion 3:** Alignment of program support with country priorities has generated strong program ownership;
 - **Conclusion 4:** Program design has improved, but M&E systems have not adapted to measure and demonstrate program level results and additionality;
 - **Conclusion 5:** Decision making on program design needs to reflect greater transparency and clear roles for all players in the partnership.

4. Following are the three main evaluation's recommendations:

- **Recommendation 1:** The GEF should continue with appropriate programmatic interventions, addressing issues that are likely to impede outcomes and performance, efficiency, and management, as they become multidimensional;
- **Recommendation 2:** The GEF should continue ensuring that programs are relevant to the national environmental priorities of the participating countries while meeting the requirements of the Conventions;
- **Recommendation 3:** M&E should be implemented at the program levels, with a clear demonstration of the additionality of the program over projects.

ABBREVIATIONS AND ACRONYMS

ADA	Agence pour le Développement Agricole
ADB	Asian Development Bank
AfDB	African Development Bank
APR	Annual Performance Report
BD	Biodiversity
BELP	Badia Ecosystems and Livelihoods Project
BPAs	Program-based Approaches
CC	Climate Change
CIDA	Canadian International Development Agency
CPF	Country Programming Framework
CPP	Country Program Partnership
DELP	Desert Ecosystems and Livelihoods Program
EA	Enabling Activity
EBRD	European bank for Reconstruction and Development
EnMS	Energy Management Systems
FAO	UN Food and Agriculture Organization
FIP	Forest Investment Program
FSP	Full Size Project
GEF	Global Environment Facility
HFDJB	Hashemite Fund for Development of Jordan Badia
IADB	Inter-American Development Bank
IAP	Integrated Approaches Pilot
IEM	Integrated Ecosystem Management
IEO	Independent Evaluation Office
IFAD	International Fund for Agricultural Development
IFIs	International Financial Institutions
IUCN	International Union for Conservation of Nature
IW	International Waters
KM	Knowledge Management
LD	Land Degradation
M&E	Monitoring and Evaluation
MDBs	Multilateral Development Banks
MENA	Middle East and North Africa
MFA	Multi-focal Area
MPPs	Multi Project Programs
MSP	Medium Size Project
MTR	Midterm Review
NCARE	National Center for Agricultural Research and Extension
NDVI	Vegetation Density
OPS	Overall Performance Study
OSS	Observatoire du Sahara et du Sahel

P2I	Progress to Impact
PA	Programmatic Approach
PBA	Program-Based Approach
PFD	Program Framework Document
PIF	Project Identification Forms
PIR	Project Implementation Report
POPs	Persistent Organic Pollutants
PPCR	Pilot Program for Climate Resilience
PPG	Project Preparation Grants
PRC	Peoples Republic of China
PRSP	Poverty Reduction Strategy Paper
PRT	Program Review Template
RAF	Resource Allocation Framework
RIE	Rapid Impact Evaluation
RSCN	Royal Society for the Conservation of Nature
SIF	Strategic Investment Fund
SIP	Strategic investment Program
STAP	Scientific and Technical Advisory Panel of the GEF
STAR	System of Transparent Allocation of Resources
SWAp	Sector-Wide Approach
TE	Terminal Evaluation
UN	United Nations
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNIDO	United Nations Industrial Development Organization
WB	World Bank
WWF	World Wildlife Fund

I. INTRODUCTION

A. Background and Objectives

1. Programmatic approaches (herein after referred to as programs) are particularly relevant to the Global Environment Facility (GEF), given the long-term nature of the environmental problems the GEF was tasked to address. Although the most common form of GEF support to recipient countries has historically and still is being provided through projects, programs have been part of the GEF since its establishment. As early as in 1999, the GEF Council supported the evolution of GEF support to recipient countries through a programmatic approach.¹ Shortly after, in 2001 the Council clarified that programs should “*secure larger and sustained impact on the global environment through integrating and mainstreaming global environmental objectives into a country’s national strategies and plans through partnership with the country*”.²

2. The shift to a more strategic partnership between the countries and the GEF was also discussed during the third GEF replenishment, where parties proposed a performance-based resource allocation system. This led to the introduction of the Resource Allocation Framework (RAF) in 2006, replaced by the System of Transparent Allocation of Resources (STAR) in 2009. These reforms influenced the way programs—particularly regional and global ones—were to be financed, i.e., either from the RAF/STAR or from ad hoc “set-aside”³ funds.

3. In May 2008, the Council endorsed the objectives and basic principles for programmatic approaches. For the first time, detailed procedures for designing programs were approved, including the requirement of preparing a Program Framework Document (PFD) when submitting a financing proposal to the GEF Council for approval. Through this major reform the program support modality was formally introduced in the GEF. This reform resulted in an increase in the submission of programs to the Council, and a change in their nature from phased activities to clustered ones.

4. Before the formal introduction of the program support modality in May 2008, the GEF allocated \$868.29 million to 34 phased/tranched and other country, regional and/or global programs without PFD, corresponding to 5 percent to the total GEF grants to that date. Post May 2008 program funding amounts to \$1,486 million for 38 programs, corresponding to 8.7 percent of the total GEF funding as of this evaluation’s cut-off date (April 2016).

5. This evaluation assesses the mechanisms and conditions by which GEF programs have attempted to deliver broader scale and longer term results by comparing them to stand-alone projects. Annex 1 presents the evaluation’s approach paper. The evaluation aims at contributing to the further development of GEF programs in the context of the GEF’s strategic move towards multifocal and integrated solutions to environmental challenges proposed in the GEF 2020 Strategy.⁴

¹ Global Environment Facility. *Corporate Business Plan (GEF/C.14/9)*. November 5, 1999, p.19.

https://www.thegef.org/gef/sites/thegef.org/files/documents/gef_c14_9.pdf

² Global Environment Facility. *The GEF Programmatic Approach: Current Understandings (GEF/C.17/Inf.11)*. May 9, 2001, p.3.

http://www.thegef.org/sites/default/files/council-meeting-documents/C.17.Inf_11_5.pdf

³ In the GEF, “set-aside” funds are targeted toward reinforcing the focal area mandate through investments that complement country activities under the STAR.

⁴ Global Environment Facility. *GEF 2020 Strategy for the GEF*. May 2014, p. 21.

https://www.thegef.org/gef/sites/thegef.org/files/publication/GEF-2020Strategies-March2015_CRA_WEB.pdf

B. Approach and Methodology

6. The evaluation was conducted applying a mixed methods approach that encompassed both quantitative and qualitative data gathering and analyses. A limited number of key questions guided the evaluation. These included two main effectiveness questions, focusing on: (i) the extent to which the different typologies of GEF programs delivered broader scale and longer term environmental outcomes and impacts compared to stand-alone projects, and (ii) the extent to which GEF programs addressed the main drivers of environmental degradation. The latter question was approached retrospectively, taking into consideration that many GEF programs covered under the scope of this evaluation were not explicitly designed to address the drivers of environmental change currently recognized by the GEF. The performance with regard to drivers was therefore analyzed with a formative approach. Other key evaluation questions focused on: (iii) the factors having influenced program ownership by participating countries and in turn the relevance of those programs to national environment and development needs and priorities; (iv) the program coherence, assessed in terms of the degree of integration between project and program level objectives; (v) GEF project cycle efficiency and program cost-effectiveness issues; (vi) governance, management arrangements and coordination issues; and (vii) monitoring and evaluation (M&E).

7. While the evaluation looked back at the overall historical evolution of program support in the GEF since 1999 to date, it covers the period from May 2008 to the present in particular detail, with a focus on the 38 programs designed after the introduction of the PFD requirement. These 38 programs encompass a total of 301 child projects. Thirty-three programs comprising 175 projects implemented prior to May 2008 were the subject of an in-depth retrospective meta-analysis encompassing 88 Terminal Evaluations (TEs).⁵ The meta-analysis only focused on the first two questions, on program effectiveness and drivers.

8. Several other tools and methods were used to gather and analyze data. These included:

- (i) a portfolio analysis covering 34 out of the 38 post-2008 programs and their related child projects (n=237);⁶
- (ii) a broader adoption analysis based on the GEF Generic Theory of Change Framework⁷, conducted using the available 52 TEs of child projects. These projects are part of 15 out of the total of 38 programs covered by this evaluation;⁸
- (iii) a geospatial analysis conducted on 105 child projects belonging to 13 programs and encompassing observations on 653 project sites. The selection of these programs was based on their maturity, expressed in terms of implementation status of their child projects. The evaluation considers mature those programs having either more than 60 percent of their child projects under implementation for more than two years (i.e. having been under implementation before April 1st, 2014), or having been completed, or both. Mature programs are expected to have produced results that can be geospatially observed in terms of vegetation productivity, expressed through the Normalized Difference Vegetation Index (NDVI), and forest cover, expressed in km² of avoided forest loss.

⁵ Of the 34 pre-2008 programs, one had to be dropped from the meta-analysis as it had no TE.

⁶ Two programs composed by projects at the Project Identification Form (PIF) stage were excluded from this analysis as documentation for those was limited or non-existent. The analysis also excluded two umbrella programs (namely the Biosafety program and the Technology Transfer program), as these have been conceived more as administrative arrangements than explicit programmatic approaches.

⁷ GEF Independent Evaluation Office. *Fifth Overall Performance Study of the GEF, Final Report*, 2014, p. 47-50.

⁸ For a detailed explanation of the broader adoption analysis approach, see: <https://www.gefio.org/sites/default/files/ieo/ieo-documents/ops5-td12-progress-toward-impact.pdf>

- (iv) a global online survey administered to country and regional stakeholders (n=684) involved in GEF programs, which obtained a 27 percent response rate. The survey focused on assessing country stakeholder opinions on their experienced incentives and disincentives in being part of a program, broader adoption, knowledge sharing and M&E, potential for leveraging co-financing and coordination issues.
- (v) Central level interviews conducted with program stakeholders (n=26) in the GEF Secretariat, the Scientific and Technical Advisory Panel (STAP) and a broad spectrum of GEF Agencies involved in GEF programs (i.e. United Nations agencies, funds and programs; multilateral development banks; and international non-governmental organizations).
- (vi) Four program case studies, encompassing country visits in China, India, Jordan, Tunisia, Morocco, Vietnam and Indonesia. The four selected programs represent the following combinations: (i) two single- and two multi-country programs; (ii) two single- and two multi-agency programs; and (iii) two single- and two multifocal area programs. Two programs were homogeneous, i.e. composed of highly similar child projects, implemented in different regions in a country or in different countries, two were not. This allowed the evaluation to observe how different levels of program complexity affected performance and results. In this evaluation, complexity is a function of the degree of homogeneity of the programs' child projects and whether they belong to multiple or single country, agency and/or focal area programs. Box 1 here below presents the main definitions used in this report.

Box 1: Definitions used in this report

Program: programmatic approach approved under the post-2008 programmatic approach modality, composed of a parent program and a variable number of child projects, designed to contribute to the overall program objective. Programs conform to the requirement of having a PFD.

Child project: project belonging to and approved under a post- May 2008 Program.

Pre-2008 programs: large projects approved prior to formalizing programmatic approaches at the GEF by introducing the PFD requirement. These were often implemented as a set of consecutive phased projects.

Program scope: global, regional, or country coverage of the programmatic intervention.

Homogenous program: a program whose projects are similar in structure and outcomes, and are executed at multiple locations within a country (for country programs) or region (for multi-country programs).

Program Complexity: degree of homogeneity or difference among child projects, as well as whether they belong to multiple or single country, agency and/or focal area programs:

- **Low complexity programs**, with two or less of the following attributes: Homogeneity, Scope, Agency, and Focal Area.
- **High complexity programs**, with three or all of the following attributes: Homogeneity, Scope, Agency, and Focal Area.

9. Case study data were collected during the country visits through interviews and focus group meetings, documentation review, geolocation information data gathering and field observations in selected project sites. The programs assessed in the four case studies are described in Box 2 here below. Detailed program case study reports are available in a separate technical document in Volume II of this report. Volume II also includes a detailed study report of the geospatial analysis conducted on 13 mature programs.

Box 2: Program Case Studies

PRC-GEF Partnership on Land Degradation in Dryland Ecosystems, China – GEF ID: 3482

The Peoples Republic of China (PRC) and the GEF set up in 2003 the PRC-GEF Partnership on Land Degradation in Dryland Ecosystems Program (“the Partnership”), initially in the form of a Country Programming Framework (CPF) for land degradation. It was originally composed of only one project, the Capacity Building to Combat Land Degradation (GEF ID 956), under the GEF Operational Program 12 (OP12).

India-GEF Coastal and Marine Program (IGCMP) – GEF ID: 3661

The main objective of this 4-year program is to demonstrate multi-sectoral approaches to mainstreaming biodiversity conservation objectives into economic activities in two marine eco-regions of the country. Conservation and Sustainable Use of the Gulf of Mannar Biosphere Reserve’s Coastal Biodiversity (GEF ID: 634) was selected as a counterfactual for this program.

Middle East and North Africa Desert Ecosystems and Livelihoods Program (MENA-DELP) – GEF ID: 4620

The goal of MENA-DELP is to contribute to the enhancement of livelihoods in desert ecosystems by harnessing their value in an environmentally and socially sustainable manner so that the flow of desert goods and services can be optimized. According to the Project Appraisal Document (PAD), “GEF financing leveraged through the MENA-DELP will enable interested countries in the region to operationalize their existing or planned investments in desert ecosystems”.

Reducing Industry’s Carbon Footprint in South East Asia – GEF ID: 3756

The program has two main objectives: (a) controlling the growth of greenhouse gas emissions attributable to rapid industrialization in the countries of the South East Asia region; and (b) helping these industries reduce their costs of fuel and electricity which continue to rise due to the recent steep increases in oil price. The program has been implemented through child projects in five South East Asian countries: Indonesia, Thailand, Vietnam, Philippines and Malaysia. Vietnam and Indonesia were selected for case study missions. This case study was conducted using Rapid Impact Evaluation (RIE). *

* Rowe, A. *Introducing Rapid Impact Evaluation (RIE) and Associated Concepts. The Scenario-Based Counterfactual and Simplified Measurement of Effects - Expert Lecture*. 35th Evaluation Conference, Canadian Evaluation Society, June 2014.
http://evaluationcanada.ca/distribution/20130618_rowe_andy.pdf

10. Different cohorts of stand-alone projects have been used for programmatic/non-programmatic comparisons. Each cohort gives a different perspective in the analysis of programmatic versus stand-alone implementations. These stand-alone project cohorts included:

- (i) Database of TEs from the IEO’s Annual Performance Report (APR): this database includes the TEs of 1184 completed projects, 308 of which were approved in or after May 2008. These include 258 stand-alone project TEs, 50 child projects under programmatic approaches. Eight more projects belonging to the GEF Biosafety program were also excluded.⁹ The child projects sample used for this comparison (n=42) represents 14 percent of the total of child projects belonging to the 38 programs under analysis (n=301), similar to the 14 percent representing stand-alone project TEs (n=258) out of the total number of projects approved between May 2008 and April 2016 (n=1795);
- (ii) Database of TEs analyzed for the broader adoption analysis conducted for OPS5: this cohort includes 447 pre- May 2008 stand-alone projects. The use of this cohort in the comparative analysis allows an assessment on whether broader adoption increased as a result of the introduction of the program modality;
- (iii) Database of multifocal stand-alone projects: this cohort is currently being analyzed in the context of the Evaluation of Multiple Benefits of GEF Support.¹⁰ It includes a total of 68 completed projects, 18 of which are post- May 2008. This multifocal cohort allows to

⁹ See footnote 6.

¹⁰ See the evaluation approach paper, on <https://www.gefio.org/evaluations/evaluation-multiple-benefits-gef-support>

compare programmatic versus stand-alone project focusing on one of the four complexity dimensions used in this evaluation;

- (iv) Database of biodiversity projects: this cohort was analyzed in a recent IEO impact evaluation.¹¹ It includes 553 projects and 3096 project sites; and
- (v) Database of land degradation projects: this cohort was analyzed for the “*Value for Money Analysis for the Land Degradation Projects of the GEF*”, recently conducted by IEO.¹² It includes 202 projects and 1047 project sites.

11. The last three databases were used for comparisons with the programmatic projects cohort (n=105 projects, 653 project sites) in the context of the geospatial analysis mentioned in paragraph 6 under (iii) above.

C. Limitations

12. This evaluation had an ambitious design, justified by the multifaceted nature of the issues that had to be covered. The main constraint related to the fact that despite the long time span covered by the evaluation, only two of the 38 programs included in the scope of this evaluation had all their child projects fully completed and evaluated. One of these was composed by only one child project. Overall, availability of child projects TEs for each program varied. The 52 available TEs belong to half of the 38 programs covered by this evaluation. Of them, two had at least half of their projects fully closed and evaluated. Program evaluations were virtually inexistent. Only the PRC-GEF Drylands partnership program underwent a program review commissioned by ADB, the lead agency.

13. These limitations on the overall body of evidence were compensated for by adjusting the existing evaluation data gathering and analysis activities, deep diving in the case studies and in the geospatial impact analysis, and adding new cost- and time- effective analyses. One such addition was the global online survey mentioned in paragraph 6 under (iv) above, which could be conducted using to the exhaustive lists of country stakeholders provided in a timely fashion by the GEF Agencies involved in the GEF programs under study.

¹¹ GEF Independent Evaluation Office. *Impact Evaluation of the GEF Support to Protected Areas (PAs) and PA Systems*, 2015. <https://www.gefio.org/evaluations/biodiversity-impact-evaluation-support-protected-areas-and-protected-area-systems>

¹² See <https://www.gefio.org/sites/default/files/ieo/council-documents/files/c-51-me-inf-02.pdf>

II. CONTEXT

A. Programs Evolution, Typologies and Definitions¹³

14. Most of pre-2008 programs were phased/tranched ones,¹⁴ implemented during the GEF Pilot Phase, GEF-1 and GEF-2, with a few single focal area programs with a country focus. A notable exception to this trend is observed in the case of the International Waters focal area.¹⁵ Phased/tranched programs tended to be discontinued by the end of GEF-3, when a new generation of programs was introduced. These were composed of a parent program and a variable number of child projects, designed to contribute to the overall program objective. Single focal area programs were the norm up to GEF-3. From the end of GEF-3 onwards programs increasingly became multifocal and programs with a regional/global focus gained prominence in the portfolio.

15. In May 2008 Council formally approved the GEF program support modality. Council document GEF/C.33/6 clarifies the scope, approval procedures, and value added of programmatic approaches as compared to the predominant project-based approach. This reform marked the official start of programs at the GEF. Since then, programs and child project PIFs began constituting a substantial volume of Council work programs. The majority of programs covered by this evaluation were approved in the last two years of GEF-4. Early post-2008 programs tended to be designed and implemented through several child projects brought together under an objectives' framework that aimed at securing larger-scale and sustained impact on the global environment. During GEF-5 program design started to become increasingly complex: compared with earlier programs, the GEF-5 shows a greater range of non-homogenous, multifocal, multi-agency, and/or regional/global programs.

16. For much of GEF history, program definitions evolved as a function of their operational and financial features. The GEF had to accommodate for the diversity of: (i) programs' financial, administrative and operational categorizations; (ii) characteristics of GEF Agencies, with the main distinction between the International Financial Institutions (IFIs) and UN Agencies; and (iii) topics of interest. This changed in October 2014, when the GEF Council approved a revised programmatic approach modality defined in terms of the program scope: (i) *Thematic*: the program addresses an emerging issue (e.g. a driver of environmental degradation); and (ii) *Geographic*: the program addresses an established need to secure large-scale and sustained impact for the environment and development in a particular geography (landscape, ecosystem, district, provinces, country, region, among others).¹⁶ In GEF-6 the GEF introduced the Integrated Approach Pilots (IAPs). These programs align with the GEF 2020 Strategy, which emphasizes the need to support transformational change and achieve impacts on a broader scale, and calls for the GEF to focus on the drivers of environmental degradation by supporting broad coalitions of committed stakeholders and innovative and scalable activities.

B. Portfolio¹⁷

17. The post-2008 portfolio covered by this evaluation is diverse in type, scope, focal area, and implementation arrangements of programs and their respective child projects. Nine country programs account for \$269 million of GEF grant financing (18 percent of the total program financing). Twenty-one

¹³ A detailed description of the evolution, typologies and definitions of programs is provided in the approach paper (Annex 1).

¹⁴ Pre- 2008 long-term programmatic interventions could only be implemented as subsequent phases of large projects, funded through discrete financing tranches.

¹⁵ For example, the Black Sea and Danube Basin initiative evolved from being a phased project prior to May 2008 to a program with parent and child projects afterwards.

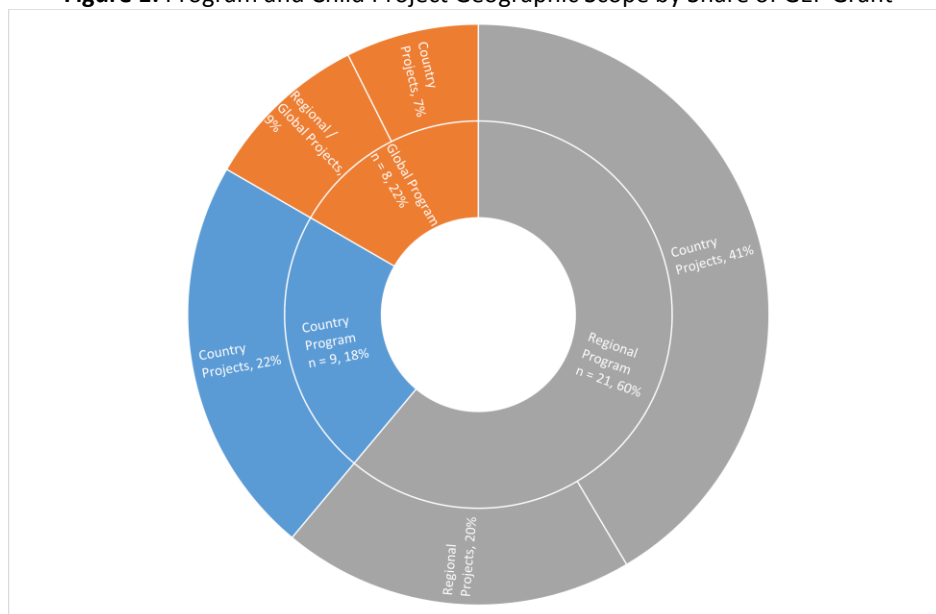
¹⁶ Global Environment Facility. *Improving the GEF Project Cycle (GEF/C.47/07)*, October 9, 2014, p. 9.

https://www.thegef.org/gef/sites/thegef.org/files/documents/19_EN_GEF.C.47.07_Improving_the_GEF_Project_Cycle.pdf

¹⁷ A detailed description of the portfolio is provided in Annex 3.

regional programs account for \$892 million (60 percent), and eight global programs for \$325 million (22 percent). Child project financing shows comparable shares (Figure 1).

Figure 1: Program and Child Project Geographic Scope by Share of GEF Grant



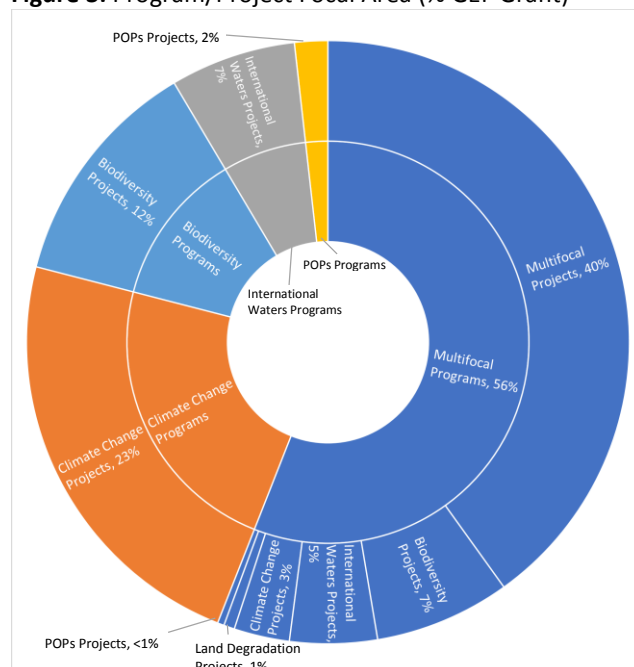
18. Most child projects are implemented in a single country: 230 projects, accounting for \$856 million (71 percent of the total child project financing). The remaining 71 (\$347 million, 29 percent) are regional or global projects. As for the regional distribution, Asia is predominant, with 127 projects (\$520 million, 35 percent), followed by Africa with 111 projects (\$310 million, 21 percent). A total of 171 child projects (57 percent) are currently under implementation, while 63 (21 percent) have been completed. GEF-4 country and regional programs are mostly multifocal, biodiversity and climate change. Multifocal programs became increasingly predominant in GEF-5 and 6 (Table 1).

Table 1: Post-2008 Programs by Geographic Scope and Focal Area

	GEF - 4			GEF - 5			GEF - 6		
	#	GEF Grant (\$Million)	Co-financing (\$Million)	#	GEF Grant (\$Million)	Co-financing (\$Million)	#	GEF Grant (\$Million)	Co-financing (\$Million)
Country	7	215	2,337	2	54	453	-	-	-
BD	2	53	775	1	26	143	-	-	-
CC	2	101	875	-	-	-	-	-	-
Multifocal	3	62	687	1	28	310	-	-	-
Global	4	125	554	1	51	223	3	149	770
BD	1	41	48	-	-	-	-	-	-
CC	2	79	501	-	-	-	1	12	56
Multifocal	-	-	-	1	51	223	2	138	715
POPs	1	4	5	-	-	-	-	-	-
Regional	9	366	1,760	11	402	5,009	1	124	683
BD	1	34	128	-	-	-	-	-	-
CC	2	55	544	3	38	1,103	-	-	-
IW	1	34	133	2	49	479	-	-	-
Multifocal	4	225	934	6	315	3,427	1	124	683
POPs	1	18	21	-	-	-	-	-	-
Total	20	706	4,651	14	507	5,685	4	273	1,453

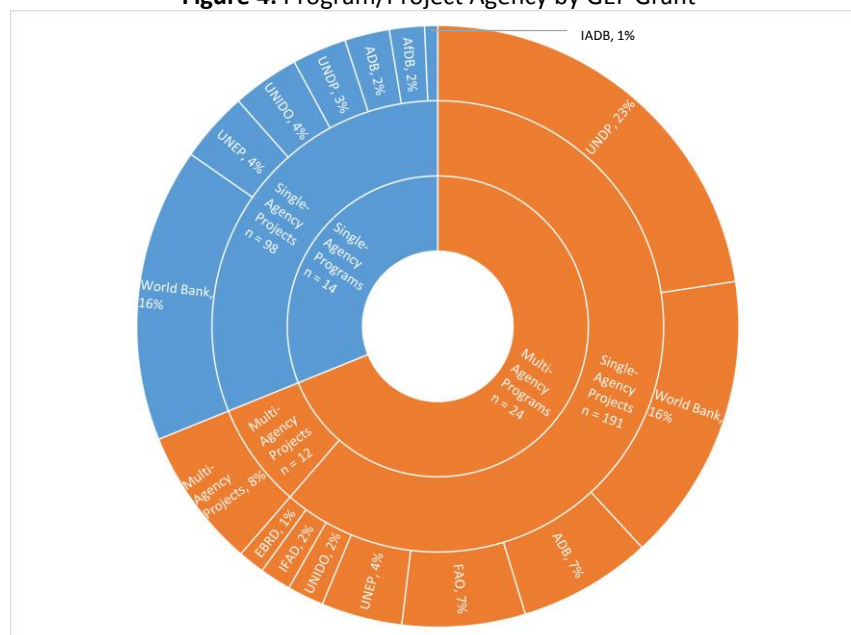
- Figure 2: Program/Project Focal Area (#)**

Figure 3: Program/Project Focal Area (% GEF Grant)



20. Twenty-four out of 38 are multi- agency programs, accounting for \$1,079 million (73 percent of the total program financing). However, the projects themselves tend to be implemented by a single agency (Figure 4): 191 projects under multi- agency programs (61 percent of the total program financing) are implemented by a single agency. Overall, the majority of child projects is implemented by the World Bank (\$384 million, 32 percent), followed by UNDP (\$325 million, 27 percent) and UNEP (\$119 million, 10 percent).¹⁸ Together, these three agencies comprise 69 percent of the total project financing.

Figure 4: Program/Project Agency by GEF Grant



¹⁸ These percentages include all single- and multi-agency child projects, be these under a single- or a multi-agency program.

C. GEF versus Comparable Donor-based Programs: Similarities and Differences

21. Over the years, the GEF financed programs that were collections of individual projects (country-based or otherwise); programs that represented long-term strategic sectoral engagement; multi-country programs; and sequenced interventions. Discussions held with central level stakeholders during the development of the approach paper for this evaluation suggested that the GEF's conceptual framework for programs is unique. In the introductory chapter, the approach paper of this evaluation uses a definition of program from the Organization of Economic Co-operation and Development (OECD), which is not fully applicable to the GEF.¹⁹ For this reason, a comparative study has been conducted as a component of this evaluation, with the objective of highlighting differences and similarities in conceptual frameworks and practices of GEF and other donor-based program approaches. The main findings of this study, presented as a separate technical document in Volume II of this evaluation, are summarized hereafter.

22. The idea of programs in the development cooperation context emerged in the late 1980s in response to slow progress in achieving tangible impacts in developing countries through the project support modality. Notably, there was concern among the donor community that the project-based approach was unsustainable and inefficient in creating economic growth and self-reliance. Among the factors considered as causing the absence of sustained results was the lack of ownership of the development process by recipient countries, and the dispersion of efforts into many discrete projects that were neither related to nor coordinated with one another or with national policies.

23. During the 1990s, various new aid delivery mechanisms emerged in line with the program approach. Among them the Sector Investment Programs (SIPs), designed as an investment mechanism to channel funding towards covering expenditures of a given sector in the country economy; the Sector-wide Approaches (SWAs), under which funds contribute directly to a country-defined sector policy; and the Program-based Approaches (PBAs), a generic program approach based on comprehensive and coordinated planning in a given sector, thematic area, or under the aegis of a national poverty reduction strategy. The pressures to increase coordination, maximize impact and reduce transaction costs while increasing national ownership continued to increase during the second half of 1990s and early 2000s. This led to a series of major international policy responses, including the 2000 Millennium Development Goals, the 2002 Monterrey Consensus, the 2003 Rome Declaration on Aid Harmonization and the 2005 Paris Declaration on Aid Effectiveness, the latter followed by the Accra Agenda for Action in 2008 and the Busan High Level forum on Aid Effectiveness in 2011, among others. These major policies aimed at delivering more coordinated development support, increasing national ownership, and streamlining the development cooperation efforts for increased impacts.

24. Programs in the GEF have evolved differently than those in the broader development context, with their own definitions and set of procedures. While the GEF policy documents do make reference to most of the same principles invoked in the various program-based approaches highlighted above (e.g. country ownership, coordination, among others), not all of these principles have applied to GEF programs.

25. Among the earliest GEF programs, Country Program Partnerships (CPPs) were designed to provide long term and large scale focus on a set of specific issues within a single country. Regional programs, regrouping countries intending to work together to achieve environmental impact in a given shared geographic unit, were constituted by highly inter-dependent child projects. Multi-country

¹⁹ OECD defines program-based approaches as “a way of engaging in development cooperation based on the principle of coordinated support for a locally owned program of development.” OECD. *Reporting Directives for the Creditor Reporting System – Corrigendum on Programme-Based Approaches (DCD/DAC (2007)39/FINAL/CORR2)*, October 15, 2008. P. 2. <http://www.oecd.org/dac/stats/44479916.pdf>

programs grouped countries, co-located or not, to work separately to achieve similar objectives under a common overarching goal, sometimes using similar approaches. Public-Private Partnerships (PPPs), added to the programmatic portfolio in GEF-5, involved the setting up of investment funds to be disbursed according to specific rules in one or more countries, towards a set objective. Table 2 below provides an overview of the main characteristics of the most common types of programs from the broader development context, highlighting the main differences and similarities to those that have been implemented with GEF support to date. These typologies do not intend to be exhaustive, nor are they mutually exclusive.

Table 2: Comparing GEF programs with major program typologies

Program typology	Key characteristics	Comparison with GEF Programs
Sector Investment Programs	- Channel large-scale, long-term investment into specific economic sectors by targeting themes and topics that go beyond traditional economic sectors; - Direct funding to cover all expenditures, including recurrent and investment ones of a given sector; - Have to be based in national strategy and policy framework; - Government or private sector has to manage the expenditure and policies; - Promote use of local capacity; - Use multi-donor and multi-stakeholder approach; - Are at least partially loan-financed.	- Similar to some of the earlier GEF sequenced programs, allowing for channeling long-term funding (e.g. international waters); - GEF Agencies could participate as co-financiers into a SIP, but would likely have to submit a project through GEF Council for operationalization, highlighting the incremental costs covered; - GEF grants cannot cover recurring or investment costs.
Sector-wide Approaches	- Funds are used for a sector-specific defined policy under the government leadership; - Usually a framework setting a direction of change; - Coordinate multiple sources and types of financing under the umbrella of a sector policy or plan; - Use multi-donor and multi-stakeholder approach - Use national systems for expenditures monitoring; - Target social sectors in highly dependent and low income countries; - contribute to facilitate the dialogue between donors and government and so strengthen the government leadership and coordination.	- Environment sector-based SWAPs exist but have experienced mitigated success. The GEF has been involved in a few of them as a funding partner; - The GEF cannot initiate or lead a SWAP, even in the environment and natural resources sector; certain types of costs have to be financed from other sources; - The lack of institutional weight that often characterizes environmental ministries makes environmental SWAPs difficult to operationalize, and the cross-sectoral nature of some environmental issues does not lend itself well to SWAP-like arrangements.
Program-Based Approaches	- Focus more on the national policy objectives (multi-sectoral, sectoral or sub-sectoral) and support locally owned development programs; - High degree of institutional flexibility; - More suited to environmental issues.	- By design, all GEF projects and programs are to be based on a national policy priority; - GEF planning horizons and time limits on fund availability exclude long-term recurring costs of PBAs from GEF processes.
Multi-Project Programs	- Most widely used type of programmatic approach; - Make no assumptions about the degree of donor coordination or country ownership; - Use of multi-donor and multi-stakeholder approach and multi type of financing; - Comprise projects that must be linked among them by some kind of unifying principle; - Offer flexibility, allow for higher-level impact monitoring and provide the possibility for donors to channel larger amount of financial assistance in a smaller number of transactions.	- Frequently found in GEF programming; - Lend themselves well to GEF co-financing as they encourage multi-donor approaches, with blended types of financing; - GEF MPPs have sought to create internal coherence and consistency through various means (integrative projects, shared methodologies and approaches); - Reflect the intention of more effective means of channeling funds, higher level impacts, smaller number of transactions.

26. In summary, this brief comparative overview indicates that although GEF programs are designed with most of the same internationally agreed principles of coordination, harmonization, country ownership, and higher efficiency and effectiveness, they differ mainly in terms of the degree of flexibility they allow in their operationalization. GEF programs mostly fall in the category of MPPs, essentially because the main operational tool for channeling GEF resources remains the project. Even the most recent group of GEF-6 programs, including the IAPs, is operationalized through individual projects, with clear time and resource limits and a strong attention placed on individual project results.

27. The fact that the GEF is replenished on a voluntary basis every four years does not favor its ability to engage in long-term partnerships. Unsurprisingly, GEF programs tend to be narrower in scope than those implemented by development cooperation agencies, be they bilateral or multilateral. MPPs have so far provided the most appropriate program approach through which the GEF can strive to channel more strategic and programmatic assistance to countries for environmental issues, which fit with the operational requirements of periodical replenishments and the principles of incremental financing.

III. FINDINGS

28. In this chapter, we first examine the evidence base assembled from the full range of quantitative and qualitative methods employed²⁰ on each of the key evaluation issues and then present the findings derived from this process. These findings assemble evidence on the key areas addressed by the evaluation questions on effectiveness, relevance and efficiency. While the over-arching comparison is between programmatic and project approaches, the analysis builds upon a variety of data sets including pre- and post-2008 stand-alone and child project cohorts, as well as single focal and multifocal area projects.

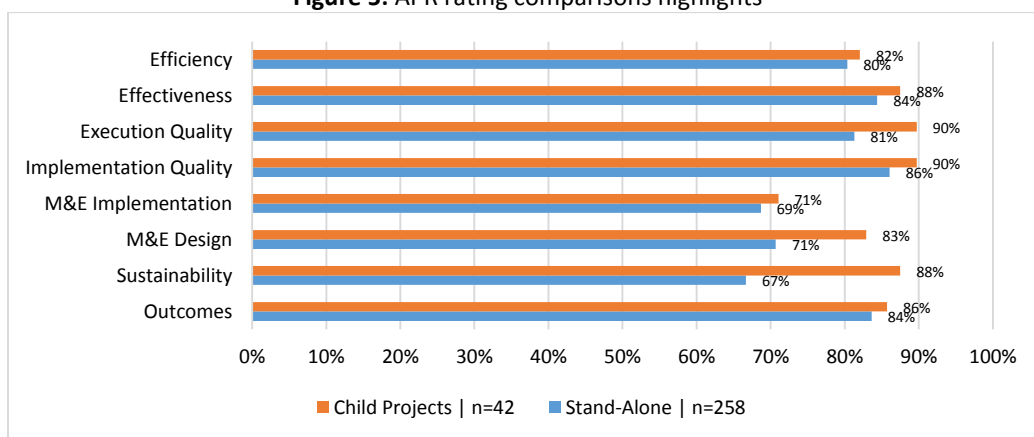
A. Programmatic projects compared with stand-alone projects

Finding 1: Child projects under programmatic approaches performed better than stand-alone projects that are not part of programs; complexity matters for outcomes.

Child projects performed better than stand-alone projects on all dimensions, especially on execution quality, sustainability and M&E design. Attention to synergies and longer-term results at the program design stage has been mentioned by stakeholders as a contributing factor. However, child projects in complex programs underperformed relative to those in simpler programs or stand-alone projects, except for implementation quality, sustainability and M&E design. In terms of vegetation density and forest cover child projects have improved local environmental conditions as compared with the null case, and single focal biodiversity projects provided more benefits than their stand-alone comparators.

29. An analysis of ratings from the IEO's 2016 Annual Performance Report (APR) database of post-2008 project TEs shows that post-2008 child projects were rated higher than stand-alone projects in the APR database on all counts (Figure 5)²¹. Eighty-six percent of child projects had moderately satisfactory and above outcome ratings compared to 84 percent of stand-alone projects. The most pronounced differences are on execution quality, sustainability and M&E design. Stakeholders interviewed²² mentioned attention to synergetic and longer term results at the program design stage as a contributing factor.

Figure 5: APR rating comparisons highlights



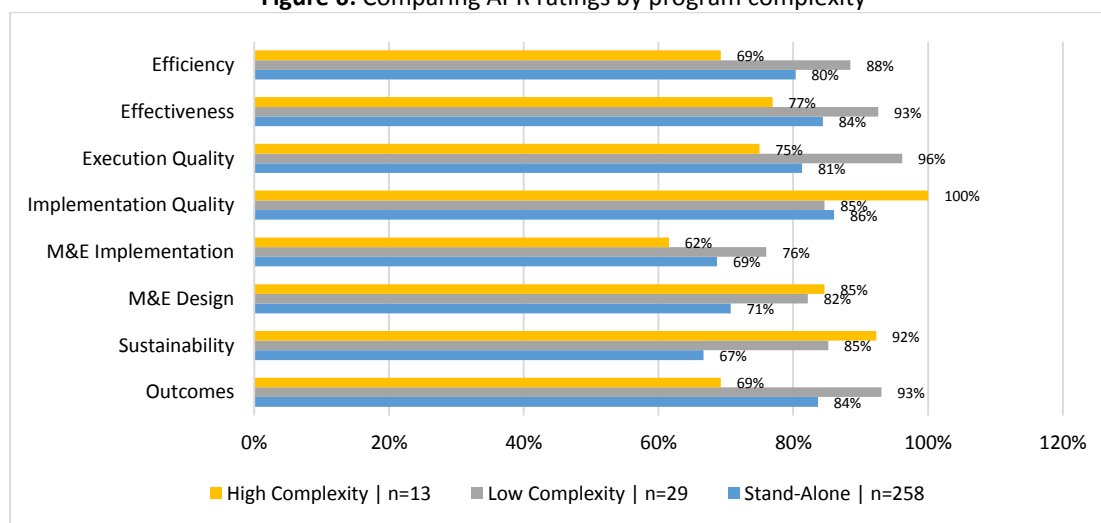
²⁰ See Section B, Chapter I.

²¹ For effectiveness and efficiency, GEF Agency ratings were used when available, otherwise supplemented by IEO ratings.

²² Stakeholders from headquarters of GEF Agencies and the GEF Secretariat.

30. Other trends emerge when programs are classified according to their complexity.²³ Splitting the limited number of available TEs of completed projects in two sub-cohorts of projects, namely those belonging to low and those to high complexity programs, further reduces the number of observations available for comparison. However, tested for statistical significance, the relation between APR outcome ratings from available TEs and the four complexity factors described in Box 1 shows that complexity is a good predictor of outcomes (both at program and at child project levels) at $p=0.05$, and is inversely related to the APR outcomes: the higher the complexity, the lower the outcomes. As shown in Figure 6 below, child projects under low complexity programs rate significantly higher than those under high complexity programs on five out of eight parameters. Exceptions are M&E design, sustainability and implementation quality, where projects under more complex programs perform better. All child projects in high complexity programs were rated in the satisfactory range as far as quality of implementation by GEF Agencies is concerned. Interviews reported that the design process for complex programs stimulates discussions within and among GEF Agencies, preparing the ground for better project implementation. A comparison between child projects of low complexity programs and stand-alone projects shows that the child projects score higher on all dimensions of performance except for implementation quality, where the two cohorts are very similar. Sustainability and M&E design of these child projects are substantially better than for stand-alone projects.

Figure 6: Comparing APR ratings by program complexity



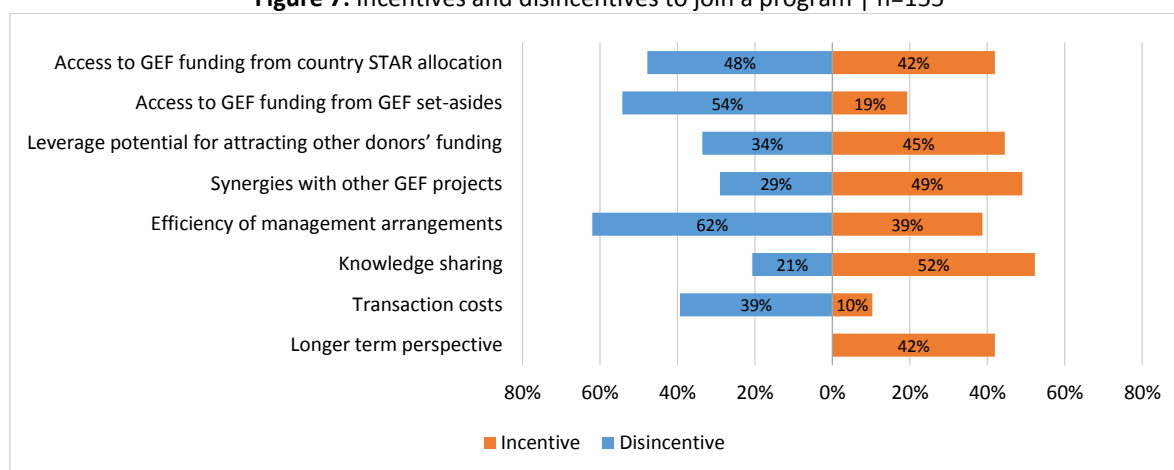
31. This evidence clearly shows that while complex programs may have longer term sustainability and better M&E design, they are substantially more difficult to deliver than are simple ones. While the programmatic approach has led to a broad range of improvements in performance and results, these have not yet been consistently attained by more complex programs. Child projects' substantially higher sustainability and M&E design ratings than stand-alone projects suggest that programs are designed with a long term perspective, which is reflected in their child projects. Analysis of program complexity along the lines followed here has been applied to the full cohort of child projects throughout this report.

32. Country stakeholders responding to an online survey conducted by this evaluation indicated that improved knowledge sharing and the potential for synergies with other GEF projects that are part of the same program are the greatest incentives for joining it (Figure 7). Against these perceived advantages, transaction costs are perceived to be higher in programs, while there are greater challenges

²³ For these comparisons programs have been classified as low and high complexity programs, as explained in Box 1.

to manage efficiently. These stakeholders also highlighted that designing for the long term is an important factor contributing to better program results.

Figure 7: Incentives and disincentives to join a program | n=155



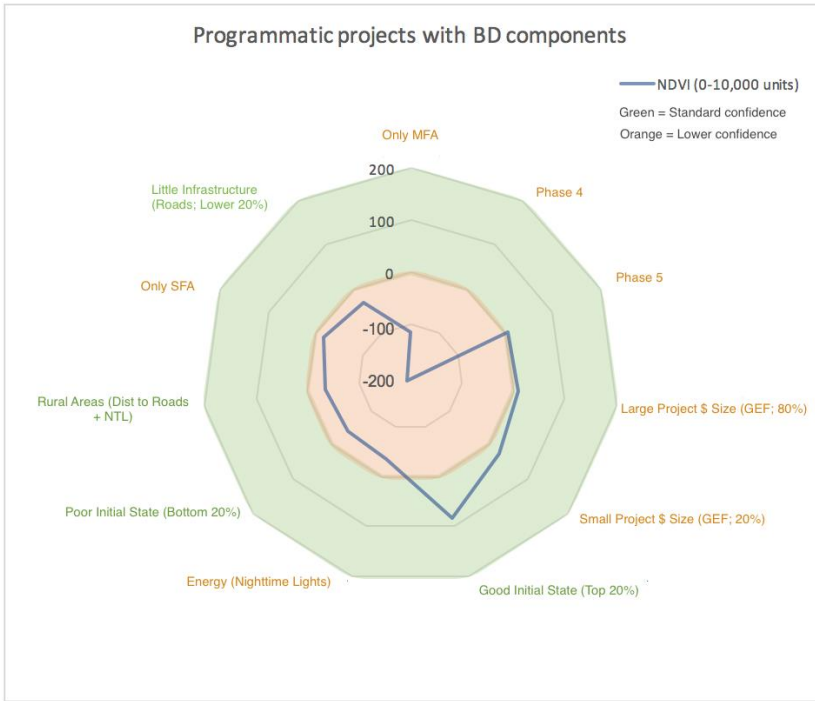
33. Of the GEF focal areas, biodiversity and land degradation offer a strong opportunity to biophysically measure results, since they have indicators that can be assessed across large scopes and a substantial body of programs and projects. The analysis conducted to measure geospatially observed physical changes in terms of vegetation productivity (NDVI) and forest cover indicates that at the global scale, child projects (whether biodiversity or land degradation) have resulted in improvements to local environmental conditions (as compared with the null case counterfactual), at a statistically significant level, which enables them to be attributed to the project. Of the different cohorts used for programmatic versus stand-alone comparisons, single programmatic biodiversity projects provided more benefits than their stand-alone comparators, at a statistically significant level.²⁴ Other findings from this analysis, which are discussed in the paragraphs below, are context-sensitive.

34. Figures 8, 9, 10 and 11 summarize findings on differences among projects in GEF-4 and GEF-5 cohorts, in each case choosing the best available counterfactual set. For example, multifocal child projects with land degradation components are contrasted to stand-alone projects with a similar multifocal component. The red areas in these figures indicate those aspects in which child projects underperformed stand-alone ones, while green areas indicate the contrary. Also in the figures, the green labels of the dimensions analyzed indicate standard confidence, akin to traditional significance testing: this descriptor is used if either the traditional, linear parametric model or the causal tree indicates significance or robustness (respectively), and the models agree on the findings. Lower confidence, in orange labels, indicates cases where the models agree in findings, but neither model ascribes clear significance or robustness. As an example of these illustrations, while it was found that – on average – child projects provided more benefits in the context of single focal biodiversity projects, mixed results were found for multifocal child projects with a biodiversity component; these figures summarize the dimensions along which all programmatic biodiversity implementations varied in their success relative to stand-alone projects.

35. As the figures illustrate, for projects in GEF-4 and 5 there was considerable heterogeneity in the conditions under which child projects gave more benefits than stand-alone ones. The only sub-set of child projects with biodiversity components that clearly related to improved vegetation were those operating in areas that already had relatively good initial vegetation productivity (Figure 8).

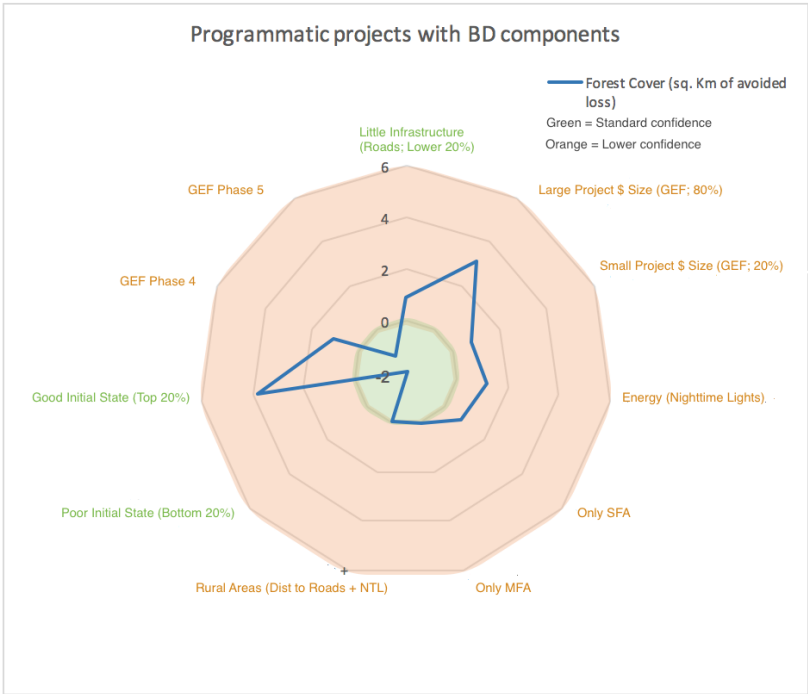
²⁴ In this case, the models used in this analysis provided standard confidence that improved environmental outcomes are attributable to child contrasted to stand-alone projects.

Figure 8: Heterogeneity in remote sensing findings along relevant dimensions (I)



36. In terms of avoided forest loss there were very few dimensions along which child projects with biodiversity components outperformed the stand-alone projects (Figure 9). The greatest additional positive contribution was evident in those areas where the initial state of forest cover was poor.

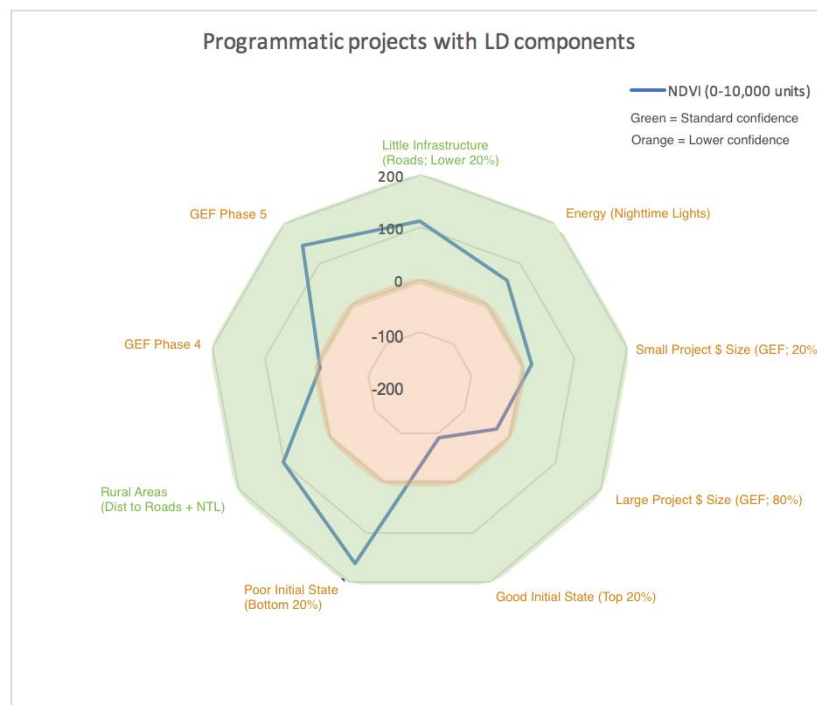
Figure 9: Heterogeneity in remote sensing findings along relevant dimensions (II)



37. With regard to vegetation cover, land degradation components in child projects outperformed comparable stand-alone projects on several dimensions (Figure 10). Substantial additional results were

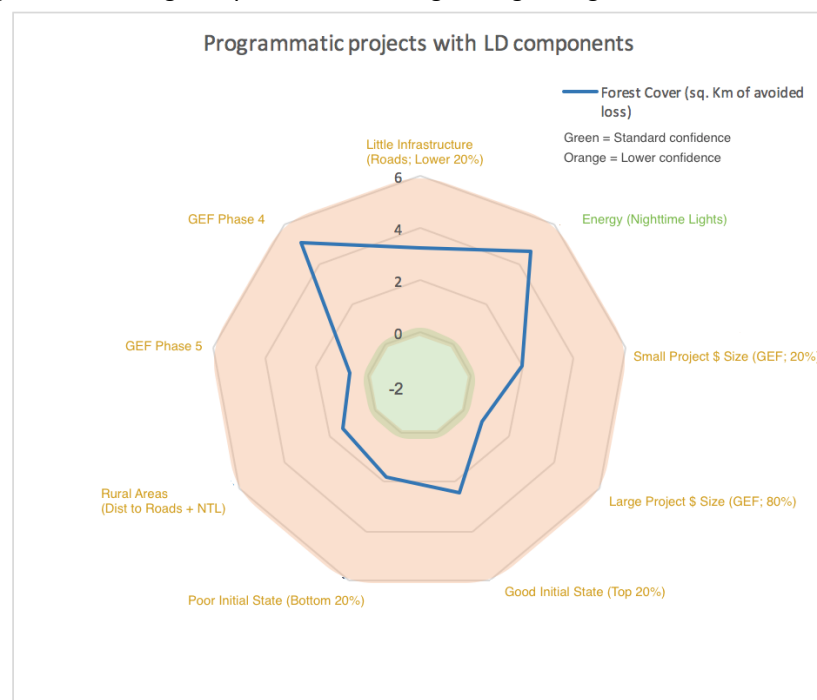
found in rural areas, those with little infrastructure and where the initial state of degradation was poorest.

Figure 10: Heterogeneity in remote sensing findings along relevant dimensions (III)



38. Unlike the situation with regard to vegetation productivity, forest cover effects of child projects with land degradation components are slightly worse than stand-alone ones for all dimensions (Figure 11). In other words, being part of a program doesn't seem to bring better results than stand-alone.

Figure 11: Heterogeneity in remote sensing findings along relevant dimensions (IV)



39. Except in the case of comparisons with the null case counterfactual, these findings based on comparison of remotely-observed physical indicators present a challenging picture to interpret due to their heterogeneity. At the global scale, considerable complexity exists when seeking to identify where programmatic implementations may be favorable as contrasted to stand-alone ones. The type of child project (multifocal as contrasted to single focal), geographic location, monetary size of the child projects, and targeted outcome of interest all contribute to the relative value-add of programs.

40. However, a global trend emerged suggesting that additional program complexity can mediate observed results in negative ways. Two pieces of consistent evidence support this finding²⁵. First, attributable evidence from separate causally identified studies suggest that the average single focal biodiversity project implemented under a program tended to outperform those not under a program; this was not true for multifocal biodiversity projects. Second, descriptive evidence suggests that as project scale increases, the attributable effect of GEF projects does not increase in a linear fashion (and, in fact, larger funding does not necessarily result in more positive outcomes along the outcomes analyzed here). This is consistent with a much broader finding emerging from this evaluation, namely that the challenges posed by additional forms of complexity to program performance and results have not yet been overcome by GEF implementing and executing partners – and require detailed strategies for improvement, if future intentions to expand programmatic approaches are to deliver the expected enhanced global environment benefits.

41. To delve deeper, the evaluation examined the same parameters in three of the four case studies. The results confirmed the global trends described in the previous paragraphs. In the child project areas of case studies in India, China and Jordan vegetation productivity improved as compared to the null case, while in India, where the program analyzed is a single biodiversity focal area one, it also improved against a counterfactual project.²⁶ In India, the vegetation in the Coringa Wildlife Sanctuary reached a higher level for the project period 2011-2015 when compared to the pre-project period 2007 – 2009. Compared to 2009, inside the India project site the vegetation condition for 2015 shows an improvement (Figure 12). In Jordan, Al Huseinieh Reserve shows a significant increase in vegetation cover since 2013 (Figure 13). More details are provided in the case study reports.²⁷

Figure 12: Spatial distribution of NDVI around the Coringa Wildlife Sanctuary, India (2015)



Figure 13: Comparing the average May-Aug NDVI Al Huseinieh Reserve, Jordan (2013 – 2015)



²⁵ See Figure 6 and related discussion in the Geospatial Impact Analysis in TD2, Volume II.

²⁶ In the cases of the PRC-GEF Partnership on Land Degradation, China and the MENA-DELP program no suitable counterfactual projects could be identified, so the comparison could only be made with the “without project” scenario.

²⁷ See TD4 – Volume II.

B. Broader and Longer-term Programmatic Results

Finding 2: Program design for broader adoption has improved substantially over time across focal areas, but actions were limited.

International waters are the only exception, having shown well-designed programmatic thinking from the early GEF phases. Data as to whether this improved design has translated into broader results are not yet available. While child projects rated higher than stand-alone projects on design for broader adoption, they demonstrated less concrete action for achieving broader adoption of outcomes during implementation.

42. Broader adoption in the GEF is said to have taken place when stakeholders adopt, expand and build on the initiatives that GEF funds, during the project period or afterwards, as a result of initial project successes. Broader adoption takes place mainly through five transformational mechanisms, namely mainstreaming, replication, scaling-up, sustaining, and market change, leading to progress along the path from outcomes to environmental impact (Box 3).²⁸

Box 3: Mechanisms for Broader Adoption of Project Outcomes

Mainstreaming: when information, lessons, or specific aspects of a GEF initiative become part of a stakeholder's own initiatives, such as laws, policies, regulations, and programs. This may occur through governments and/or through development organizations and other sectors.

Replication: when a GEF-supported intervention is copied at a similar scale, often in other locations.

Scaling-up: when a GEF-supported intervention is implemented at a larger geographical scale, often expanded to include more political, administrative, economic or ecological components. This allows concerns that cannot be resolved at lower scales to be addressed, and promotes the spread of GEF contributions to areas contiguous to the original project site.

Sustaining: when a GEF-supported intervention or outcome is continued by the original beneficiaries without GEF support so that they can keep reaping the benefits.

Market change: when a GEF-supported intervention influences economic demand and supply shift to more environment-friendly products and services.

43. Broader adoption is a central concept in GEF programs. One of the anticipated advantages of programmatic approaches as expressed in the Council document *'The GEF Programmatic Approach: Current Understandings (GEF/C.17/Inf.11)'*²⁹ is that they will deliver results, which are both broader and longer term than those obtained from stand-alone projects. This was confirmed by the expectations expressed by stakeholders interviewed. To help assess the extent to which this has been achieved, comparisons were conducted between the OPS5 cohort of TEs of stand-alone projects used for broader adoption analysis (pre-2008), TEs of post-2008 child projects and a post-2008 multifocal project TEs cohort.³⁰ These show differences in the extent to which projects were designed to achieve broader adoption as well as in the amount of concrete action taken for this purpose, as reported in their respective TEs.

44. On design for broader adoption, recent child projects scored higher than both earlier (pre-2008) projects analyzed for OPS5 and recent (post-2008) stand-alone projects. However, they showed less concrete action for broader adoption during implementation than their stand-alone counterparts. As shown in Figure 14, 31 percent of child projects intended to promote broader adoption, but only 13

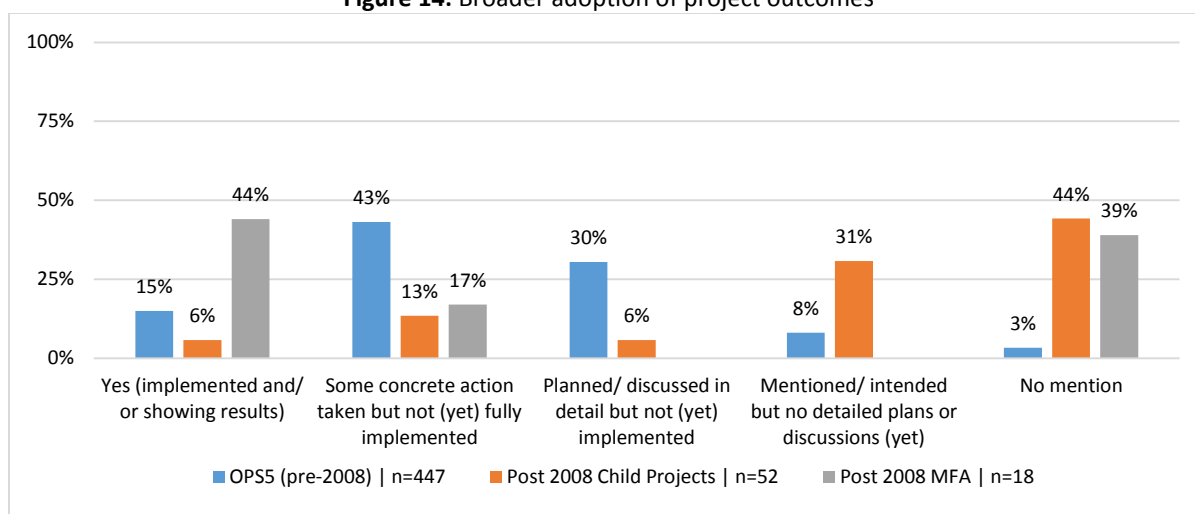
²⁸ For a detailed description of the GEF generic theory of change framework, see section 7.3 of OPS5 (GEF IEO 2014b).

²⁹ Ibid.

³⁰ See Section B in Chapter I for a description of these cohorts of comparators.

percent took some concrete actions towards this and 6 percent implemented elements of broader adoption. The inverse is true for pre-2008 OPS5 stand-alone and the post-2008 multifocal cohorts, in both of which projects have taken more concrete action for broader adoption than have the recent child projects. In the cases where actions for broader adoption were taken, it was related in most cases to management approaches and knowledge management initiatives that were already foreseen at design. It is noteworthy that the TEs of 44 percent of child projects and 39 percent of multifocal projects did not mention environmental change beyond project outcomes. No judgement could therefore be made on the extent to which broader adoption has been achieved.

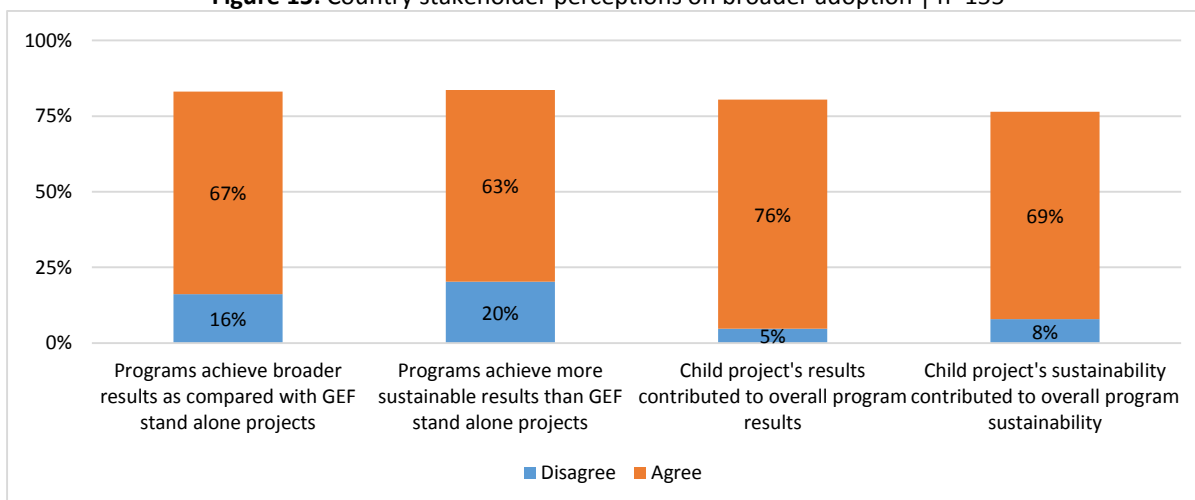
Figure 14: Broader adoption of project outcomes



Note: Multifocal n=18 out of 68 post-2008 multifocal (26 percent), OPS5 n=447 out of 2371 pre-2008 (19 percent), Post-2008 Child n=52 out of 301 (17 percent)

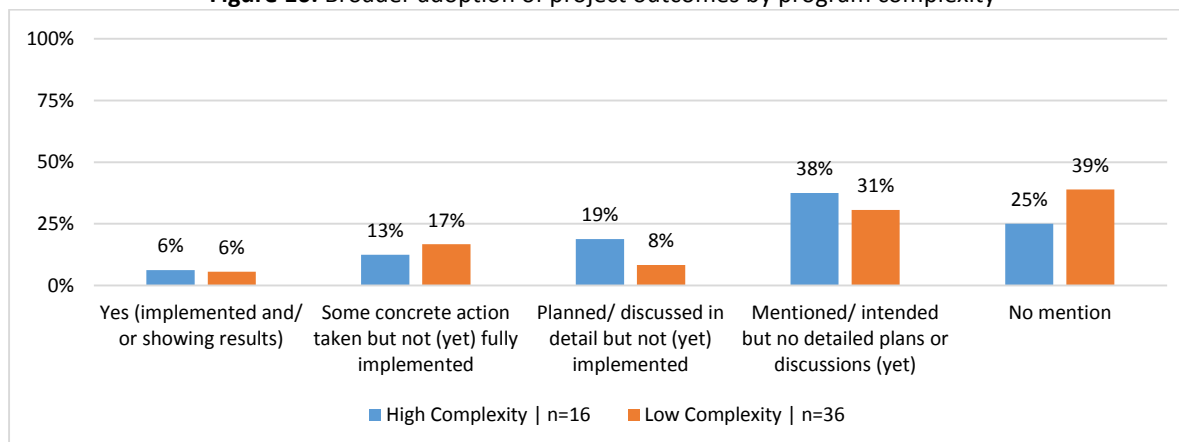
45. Compared with the empirical position revealed by examination of TEs, stakeholders indicated a highly optimistic perception of the contribution of child projects towards sustainability and broader adoption of program results (Figure 15).

Figure 15: Country stakeholder perceptions on broader adoption | n=155



46. While child projects under both the low and high complexity cohorts showed comparable results, those in highly complex programs had more mention and planning for broader adoption, while those in simpler programs had more concrete action on this dimension (Figure 16). This indicates an intention of a long-term perspective as complexity increases, while actual implementation reduces.

Figure 16: Broader adoption of project outcomes by program complexity



47. The most frequently observed mechanisms of broader adoption were mainstreaming, mentioned in one third of the TEs analyzed and replication, reported in 21 percent of the cases. Scaling-up was lower, at 6 percent, and market change was virtually absent. It is too early to say more on achieved programmatic broader adoption, except for acknowledging that it takes longer to operationalize a complex set of issues than more simple approaches.

48. Detailed case studies of specific programs provided examples where mainstreaming has occurred. The India Coastal and Marine Program and the PRC-GEF Partnership on Land Degradation in Dryland Ecosystems in China showed substantial progress in mainstreaming new approaches towards environmental management, particularly at systems and institutional level (Box 4).

Box 4: Evidence of Mainstreaming in two Country Programs

The India Coastal and Marine Program

In terms of delivering broader scale and longer-term environmental outcomes compared with stand-alone projects, both child projects have made progress towards stronger institutional and systemic frameworks for environmental management, whereas the stand-alone project did not. In Godavari, the EGREE Foundation (which the project established) has had substantial success in bringing private sector bodies into the environmental protection arena; whilst in Maharashtra the Coastal and Marine Biodiversity Foundation is less advanced, but expects to target similar stakeholders. In the counterfactual project area, the Gulf of Mannar Biosphere Reserve Trust was established, but has not played a strong role in environmental management and has been largely ineffective since project closure. Both the child and stand-alone projects have devoted much of their attention to strengthening community level livelihoods, but the counterfactual project has not gone far beyond this while both child projects have focused on a much broader range of stakeholders in the production landscape.

Both child projects have informed national policy actions. The Godavari project promoted the inclusion of a Chapter on Coastal and Marine Protected Area Conservation in the national Wildlife Action Plan (2016 – 2030), while the Malvan project influenced changes in the Marine Protected Area Law. This ability to influence national level instruments is attributed to the ability to escalate knowledge of project approaches and results through influential members of the National Steering Committee.

Partnership on Land Degradation in Dryland Ecosystems, China

GEF Agencies' staff in Beijing confirmed the Partnership achievements in mainstreaming of Integrated Ecosystem Management (IEM) in provincial level policies and plans. After an IFAD-led child project ended, IEM was integrated in 26 more counties in Gansu, and 6 more National Nature Reserves (NNR) in other provinces. Another implementing agency, ADB noted that through the Partnership, for the first time an integrated ecosystem approach was introduced in China, pointing to the IEM feature of involving technical and financial resources from different sectors and government ministries and departments at national and local level. Under IEM, a set of principles was developed to involve local stakeholders (local governments, local research institution, and universities) to build capacity to combat land degradation through a bottom-up approach. IEM was mainstreamed into the five-year plans in 4 out of the 6 provinces involved in ADB projects. The World Bank also confirmed that participatory natural resource planning and management was introduced and adopted in other projects.

49. These cases showed lower achievements with regard to replication and scaling-up. The RIE case study found that the South East Asia energy efficiency program is leading to substantial energy saving gains in both Vietnam and Indonesia and that these gains are attributable to the UNIDO program, set in the context of national legislation, ISO 50001 and other factors.³¹ However, replication of these direct effects to other high energy consuming enterprises was not yet certain, according to the assessments of a global panel of energy efficiency experts. This suggests that the UNIDO approach is relevant for enterprises directly engaging in applying Energy Management Systems (EnMS) as a result of taking the UNIDO training, but weak in generating replication beyond these enterprises. The limited progress in terms of scaling-up highlighted earlier in the quantitative broader adoption analysis is illustrated by field-based evidence from the MENA-DELP case study (Box 5).

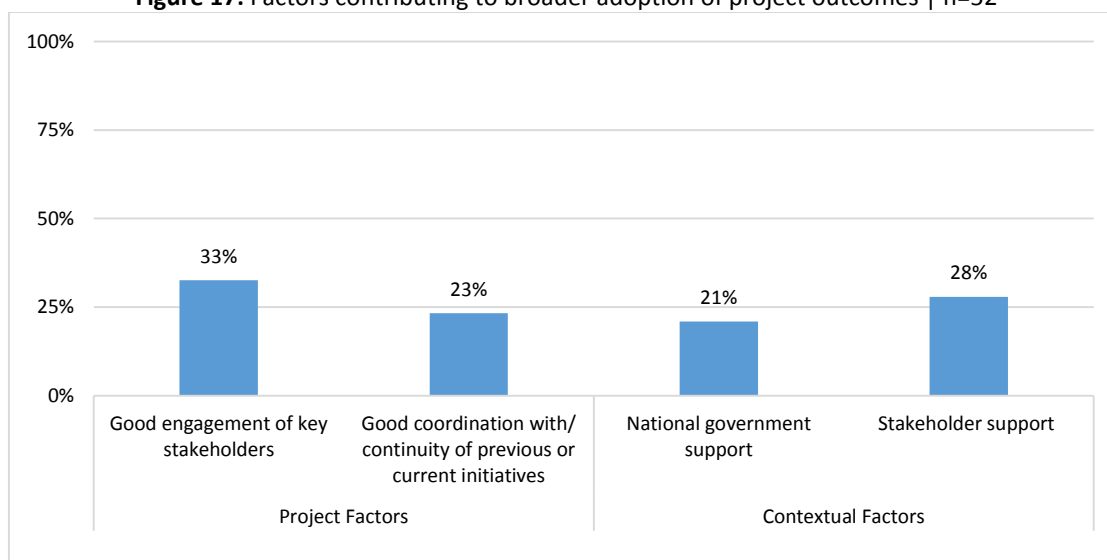
Box 5: Absence of scaled-up results of the MENA-DELP

The MENA-DELP is a collection of individual national projects, loosely related to each other through a regional umbrella project. Their coherence in the program in terms of environmental objectives is very generic – they are all attempting to harness an arid or semi-arid landscape for environmentally sustainable development. There is no evidence that there are any multiplicative benefits from their participation in the regional program. The outcomes and potential impacts of MENA-DELP are therefore not different from those of the national projects, apart from some aggregate M&E data and experience sharing and lesson learning among the program participants.

Overall, the results of MENA-DELP are not demonstrably broader scale or longer term than they would have been through the implementation of a set of stand-alone projects.

50. Whether project environmental outcomes are achieved or intermediate outcomes are broadly adopted is dependent on both project- and context-related factors. The most prominent factor impacting outcomes is strong national ownership, promoted by good engagement with, and support of, key stakeholders such as national governments, civil society organizations and the private sector (Figure 17).

Figure 17: Factors contributing to broader adoption of project outcomes | n=52



³¹ See the RIE case study report in TD4 – Volume II.

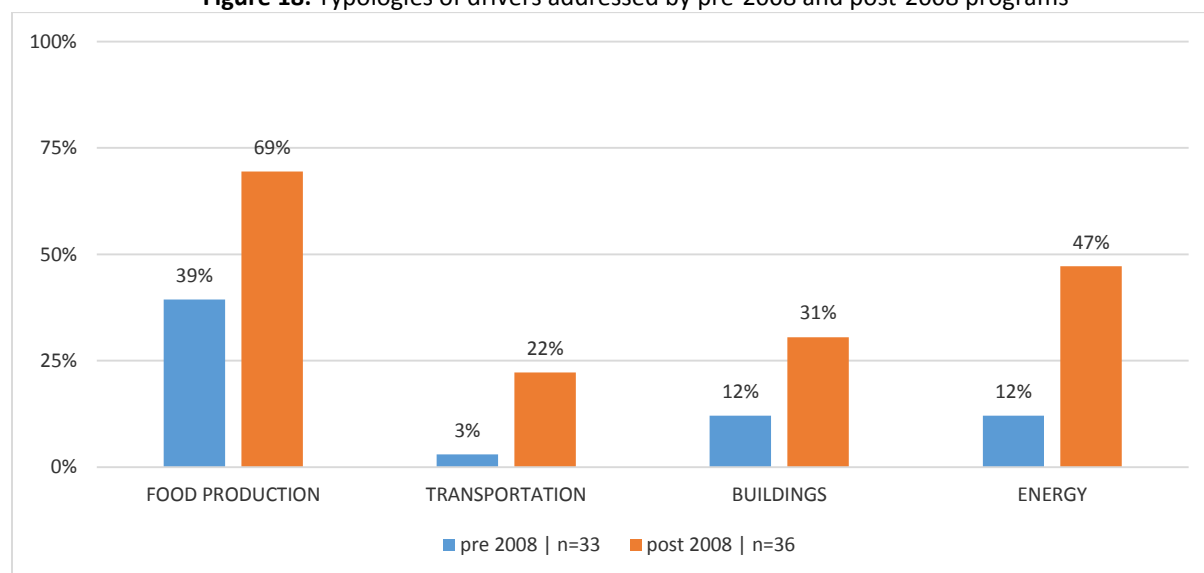
C. Addressing Drivers of Environmental Degradation through Programs

Finding 3: Programs represent a shift toward a more integrated systemic approach to address drivers.

Programs have evolved from a narrow approach, largely focused on mitigating the negative effects of food and energy production on biodiversity loss, land degradation, and climate change towards applying an integrated approach encompassing a wider set of drivers such as food and energy production and consumption, buildings and infrastructure, construction and transportation.

51. The meta-analysis conducted on pre-2008 programs³² indicates that 39 percent of those programs addressed food production as the driver of environmental degradation. Addressing food production as a driver significantly increased to 69 percent of (36) post-2008 programs (encompassing 282 projects). This cohort also shows a move towards addressing other drivers besides food production, in particular energy production, which was dealt with in 47 percent of the cases, and to a lesser extent buildings and infrastructure (31 percent) and transportation (22 percent) (Figure 18). This analysis indicates that while addressing drivers as a specific concept has been newly and explicitly articulated in the GEF 2020 Strategy, historically GEF interventions have attempted to do this to some extent. However, they did so without an explicit analysis of this dimension in their project documents.

Figure 18: Typologies of drivers addressed by pre-2008 and post-2008 programs



52. Central level stakeholders from STAP and the Secretariat highlighted the need for a comprehensive theory of change to address global drivers. To them, bigger investments are in principle more likely to address drivers, whereas projects tend to address symptoms. Drivers tend to cut across national boundaries, which explains why regional and global programs are necessary.

53. Case study programs in India and MENA-DELP showed a clear focus on drivers of environmental degradation (Box 6).

³² See Section B, Chapter I.

Box 6: Tackling Drivers of Environmental Degradation in the India and MENA-DELP programs

Both child projects targeted key drivers of environmental degradation directly. Godavari had a particular focus on private sector industry (energy and agriculture related), while Malvan targeted agriculture, fisheries and tourism. The counterfactual project targeted community level food and energy production activities, but did not interact with the important industrial level stakeholders in these sectors.

Insofar as the MENA-DELP addresses drivers, only the food production sector is a major focus. In Morocco, environmental effects of olive and cactus production are addressed through a value chain approach in the ASIMA project. Tunisia has some focus on food production and by-products through the Oasis project, while the Jordan Badia project targets benefits for the livestock sector through better water management. Both Jordan and Tunisia have an ecotourism focus, although it is not clear where this fits in the GEF classification of drivers. In terms of child project design, the drivers addressed are those recognized as most urgent in each country and/or those which are regarded as inadequately addressed to date.

54. In summary, pre-2008 programs focused mostly on environmental degradation caused by food production activities, although there were also substantial areas of intervention related to energy. This somewhat narrow approach changed in post-2008 programs, which focused on mitigating the negative effects of both food and energy production on biodiversity loss, land degradation, and climate change. A new emphasis on drivers was introduced with the GEF 2020 Strategy. GEF-6 programs, including the three IAPs, apply a system approach that potentially encompasses a wider set of drivers such as food and energy production and consumption, buildings and infrastructure construction, and transportation.

D. Ownership

Finding 4. Program ownership at the country level is highly linked to the degree of alignment with national priorities.

With the notable exception of programs addressing transboundary issues (notably International Waters), GEF programs have progressively shifted over time from a country to a multi-country focus.³³ STAR funds are a substantial share of total program resources, regardless of geographic scope. In general, the smaller the country's STAR funds, the higher the share of its total STAR allocation is given to a program. Central- and country-level stakeholders noted that country programs have stronger ownership than regional/global ones, as they tend to be more closely aligned with national priorities.

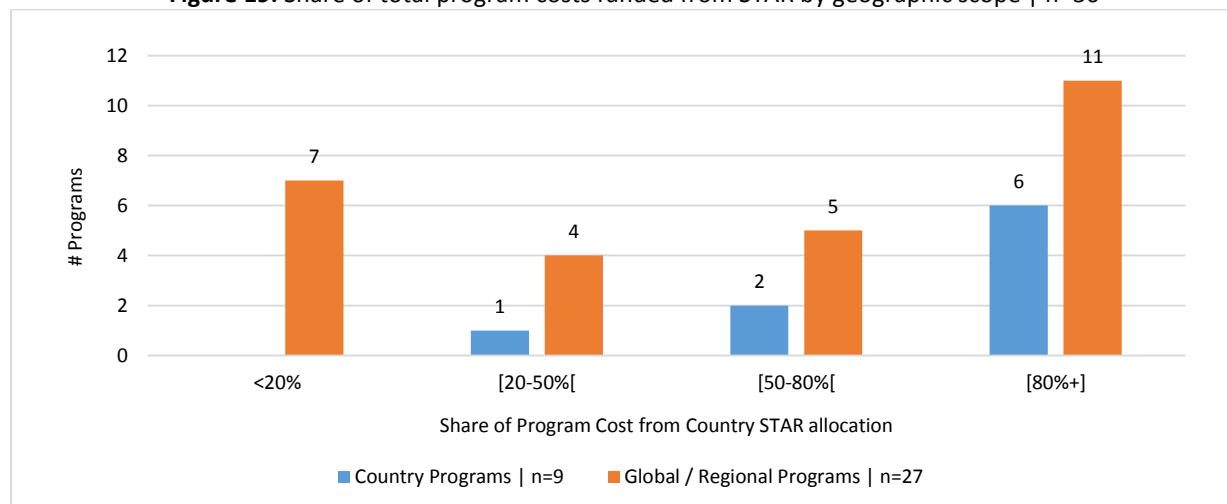
55. The evaluation measured the use of program grant amounts coming from country STAR allocations as a quantitative approach to assess program ownership by participating countries. In this analysis, the use of STAR funds in programs is expressed in two ways: (i) in terms of percentage of total program costs financed through STAR by program geographic scope; and (ii) share of individual country STAR allocation used for programs and their respective child projects.

56. The large majority of country programs use a substantial amount of STAR funds: for six out of the total of nine country programs in the post-2008 cohort, over 80 percent of the total program cost was funded from STAR allocations (Figure 19), suggesting a high level of national ownership. Regional/global programs tend to have a lower proportion of total program costs from the STAR allocation than do country ones. These programs relied on set-asides and seven of them were funded exclusively from that funding source. Two of the seven were persistent organic pollutants (POPs) and

³³ In GEF-4, seven out of 20 were country programs. Those programs progressively decreased in GEF-5 (two out of 14) and GEF-6 (none). At the same time, regional and global programs increased from 13 out of 20 in GEF-4 to 12 out of 14 in GEF-5. All GEF-6 program are either regional or global (Table 2 in Section A, Chapter II).

three were international waters programs: neither of which focal areas have STAR allocations. Another two were climate change regional programs (Asia, Africa).

Figure 19: Share of total program costs funded from STAR by geographic scope | n=36



57. Country programs, in both GEF-4 and GEF-5, are in countries with large STAR allocations, of over \$100 million (Figures 20, 21), with one exception.³⁴

Figure 20: GEF-4 - Program Use vs. Country allocation (\$ million)

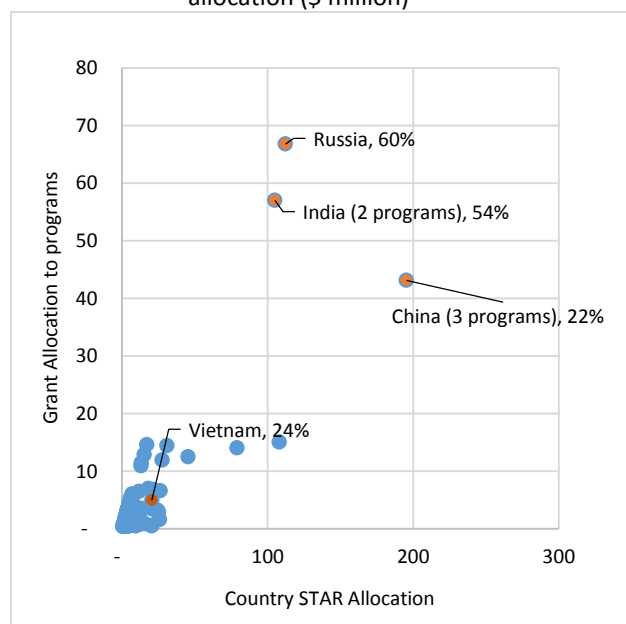
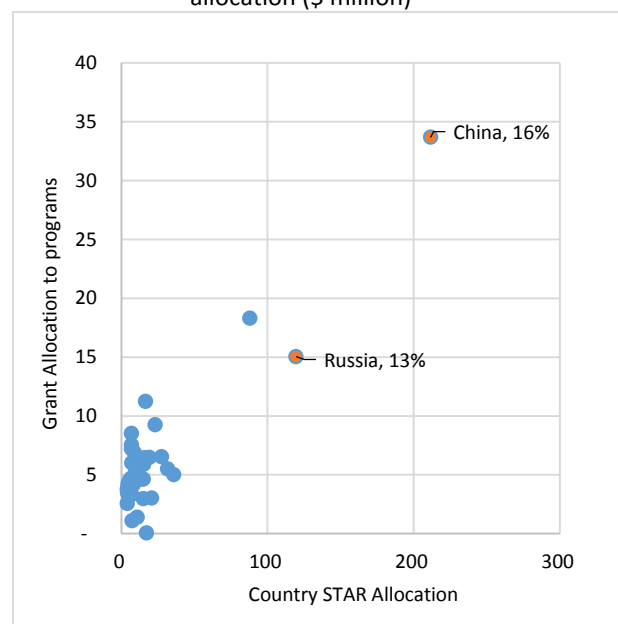


Figure 21: GEF-5 - Program Use vs. Country allocation (\$ million)

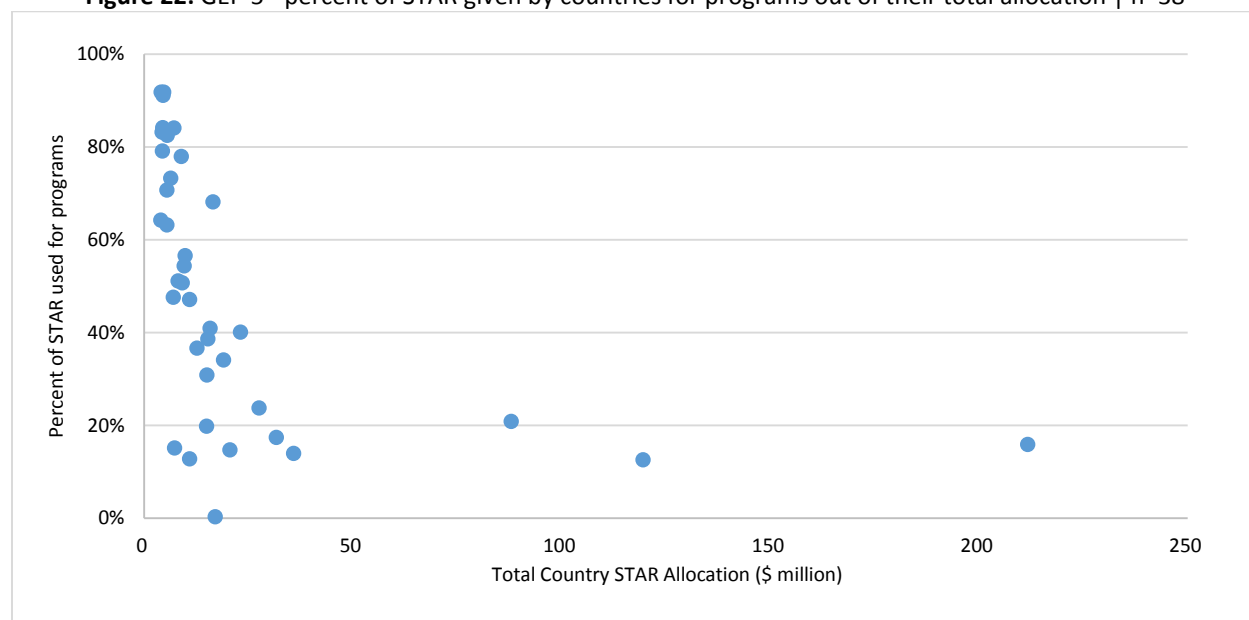


58. Figure 22 depicts the share of the total country STAR allocations used for programs. In GEF-5, of the total of 38 countries involved in GEF programs (be these country, or regional/global ones), those with a small STAR allocation tended to use the majority of it for those programs. The larger the STAR envelope, the lower the share of STAR funds used by countries towards programs. The same trend is observed in GEF-4 (108 countries). When small countries agree to commit most of their STAR funds to a

³⁴ A GEF-4 country program in Vietnam with a STAR allocation \$20.5 million (Figure 20).

regional/global program, they do so towards a child project. Their ownership relates more to that particular project than to the program of which it forms a part.

Figure 22: GEF-5 - percent of STAR given by countries for programs out of their total allocation | n=38



59. Country programs have stronger ownership than regional/global ones. This was confirmed by interviews conducted with national stakeholders during country case studies, who consistently indicated that alignment with country priorities is the most important factor for agreeing to participate in a program. This interpretation was supported by stakeholders in the GEF Secretariat, STAP and GEF Agencies, particularly Multilateral Development Banks (MDBs). Country stakeholders interviewed also highlighted that, when compared with stand-alone projects, GEF programs bring about a broadening of ownership within the countries from one body (often where the OFP is located) to a range of government ministries and/or departments and often also to private sector institutions. For example, the PRC-GEF Partnership for Land Degradation case study showed an increased scope of ownership of programs both at national and provincial level. Government and GEF Agency country-level stakeholders in China also stated that ownership of programs is incentivized by a longer-term perspective than available with stand-alone projects and by synergies among components, in addition to the potential of leveraging a high level of additional funding. In this case, program ownership is demonstrated by the substantial amount of provincial co-financing committed and by the widespread replication to other provinces of the integrated ecosystem management approach, which the program introduced. By contrast, in the case of MENA-DELP, program ownership by participating countries is limited to the respective child projects. The varying experience of ownership in national and regional programs is explored in Box 7 below.

Box 7: Country Ownership in National and Regional Programs

The Godavari Child project of the India Coastal and Marine program has realized several important gains from being part of this national program. The Program National Steering Committee has key Government stakeholders on it, who feed its experiences into high level policy discussions. These have even covered India's international obligations, for example with regard to the CITES. Similarly, The Malvan Coast Project has demonstrated good country ownership, since National, State and local governments are all engaged with it. The Project design is in line with national priorities, notably addressing one of the five most important marine and coastal areas in the country. The Minutes of the national and state level Steering Committees clearly demonstrate that the Project is nationally owned and in line with country-wide and local priorities. The Counterfactual project in the Gulf of Mannar involved several organizations and many communities, which brought a strong level of ownership from these stakeholders. However, the Government of Tamil Nadu altered the project's focus to fit its own ideas of what was needed. Its TER (P64) emphasized "the livelihood and protection aspects at the expense of the higher-level policy and institutional changes that were necessary and expected, and the management actions that could have encouraged the sustained use of marine resources have been largely overlooked".

The National Steering Committees for the program's child projects have not only linked them up to national level institutions and policies, but have also given greater visibility at District level. This has been instrumental in generating support from key officials, such as District Collectors. Ownership has therefore been strong at several key levels. For the Gulf of Mannar counterfactual, the State Government established ownership, but did so by focusing on livelihoods benefits to the virtual exclusion of the project's environmental objectives.

In contrast to States participating in the national project in India, MENA-DELP participating countries have only a modest degree of ownership of the program. They claim benefits from such aspects as participating in international workshops (particularly on M&E, tracking tools of the GEF system and income generating approaches in participant countries) and from shared experiences generated by study tours. Although the child projects are relevant to national needs, this derives from their original design processes as national projects, rather than from any direction generated by the program. In the case of Jordan, for example, the Badia project was already designed as part of a much larger Government program and was later fitted into the MENA-DELP on request of the World Bank.

GEF strategies in China are strongly aligned with those of the country. Environment protection is included in the 13th PRC Five-year Plan. The Integrated Environmental Management (IEM) concept has been integrated into policies and laws at different levels, from provincial to national. The GEF Operational Focal Point (OFP) noted that participation in the program has helped different national agencies work together. This collaboration is not limited to Partnership activities, but spreads more widely. A recent example is the collaboration between the Ministry of Agriculture, State Forestry Administration and the International Commerce Department on GEF-6 projects concerning alien species quarantine. The strong relevance to the national policies and plans for combating desertification in Western China contributes to strong ownership of the Partnership.

60. In summary, program ownership in countries tends to be closely related to the degree of alignment with national priorities. Further, the more a program is country focused, the more STAR allocations are used. Country programs are reported to bring about a broadening of national ownership from central control through the OFP, usually in the environment ministry/department, to a broader range of government and sometimes private sector institutions. It has also been noted that GEF funding is often not accounted for in national budgets, which is likely to weaken its place in national planning processes.

E. Coherence

Finding 5: Program/child project coherence has improved in recent programs

Program objectives are now better defined than in earlier GEF funding periods; child projects have improved in design and are now better linked to the overall program. This improved coherence of programs and the associated child projects is notable in the design of highly complex programs, under which projects more specifically address program outcomes.

61. Analysis of post-2008 child projects shows that 89 percent of them indicate clear linkages with their respective programs at the design stage. This high level of coherence is confirmed by the fact that 43 percent of child projects address through their activities all of the program components (Figure 23). Figure 24 below shows that projects under higher complexity programs tend to address more of the program components than those in less complex programs.

Figure 23: Program Components addressed through Child Projects | n=237

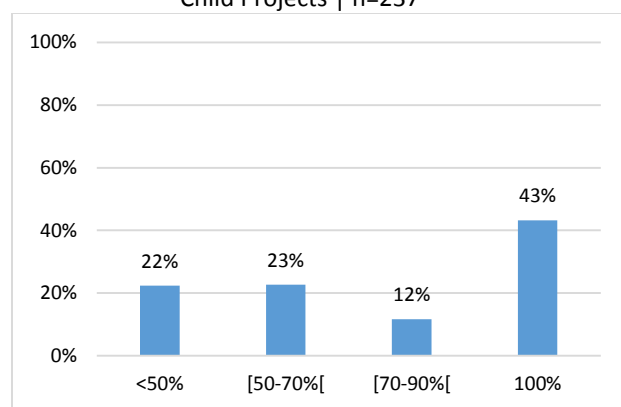
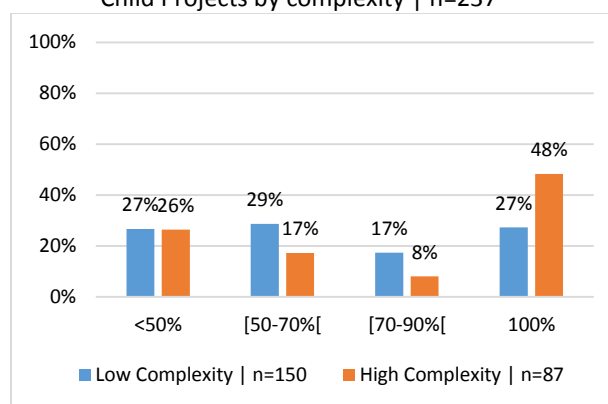


Figure 24: Program Components addressed through Child Projects by complexity | n=237



62. For 53 percent of the total of 237 child projects there is no clear indication as to how project indicators are intended to contribute to program reporting.³⁵ For the remaining 47 percent, the higher the program complexity, the more likely it is that a clear explanation of how project indicators will contribute to program reporting will be provided in the documentation. Projects under high complexity programs are therefore better designed to address how they will contribute to the program outcomes: 63 percent of projects under high complexity programs had a clear indication of how project indicators would contribute to reporting on program objectives, as against 37 percent of projects under low complexity programs. Importantly, the majority of highly complex programs has been approved towards the end of GEF-4, in GEF-5 and in GEF-6: improved reporting systems suggest learning from experience on this dimension. These findings are further highlighted in the geospatial impact analysis (see Figures 8, 9, 10 and 11). In three of these four cases, evidence suggests that the attributable value-add of GEF programmatic approaches relative to stand-alone approaches increased between GEF-4 and GEF-5.

63. As seen in the previous paragraph, coherence has significantly improved over time. Homogeneous programs, in which the project objective and outcomes are by definition coherent with those of the overall program, tend to perform better as a program, as observed in this evaluation's country case studies in the Energy Efficiency Program in South East Asia and the India Coastal Marine Program.

³⁵ Program-level M&E is discussed in Section H, Chapter IV.

64. The appearance of program/child project coherence in GEF-4 and some subsequent programs was facilitated by vague criteria and a tendency towards excessively broad objectives in program design, which made it easy to fit in a diverse set of very loosely inter-related child projects. This is evidenced by the India Coastal Marine, MENA-DELP and PRC-GEF Partnership in Land Degradation programs (Box 8).

Box 8: Coherence between GEF Programs and their Child Projects

The India Coastal Marine Program aimed to mainstream coastal and marine biodiversity conservation at three levels: systems, institutions and community. Overall, both of the program's child projects had objectives which were coherent with the overall program direction. The earlier counterfactual project had similar objectives and provided some inputs relevant to the later program design.

The MENA-DELP program outcome level objectives are broad and comprehensive, including the following: (i) Improved agricultural management; (ii) Sustained flow of services in agro-ecosystems; (iii) Increased investments in SLM; (iv) Increase in sustainably managed landscapes and seascapes that integrate biodiversity conservation; (v) Promote investment in renewable energy technologies; and (vi) Reduced vulnerability to climate change in development sectors. It has therefore been possible to integrate objectives of the child projects into those of the program in a coherent manner.

Interviews with GEF Agencies in Beijing broadly confirmed the alignment and coherence between the PFD and its child projects, explaining that it was a specific requirement at the design stage. In describing how the idea of the Partnership came about in 2002, ADB notes that the programmatic concept is very much embedded at the design stage. However, after several years of implementation, projects were largely implemented as stand-alone. Collaboration among GEF Agencies involved in the Partnership is not strong, not programmatic and is limited to rare round-table meetings.

65. These findings on coherence can be placed in perspective by those already discussed under results and broader adoption. Whilst programs are becoming more complex and have increasingly explicit connections between program and child project objectives at design stages, they have so far been delivered less well and with less concrete actions promoting broader adoption than earlier and simpler child projects. This indicates that implementation practices and supervision have not been able to keep pace with the challenges posed by managing and coordinating increasingly complex interventions.

66. The complexity of design documents varies considerably between program types. A broad range of interviewed stakeholders involved in design processes suggested that IAPs have greater design coherence than previous types of programs, which were often just a set of projects bundled together. Similarly, other GEF 6 programs have explicit design systems to filter for coherence of child projects. For example, the China National Protected Area System Program (GEF ID: 4646) is said to be more complex than previous major GEF investments in the country, since provinces drove these with no national framework or procedures for a coherent National Protected Area System. Under the new approach, coherence is ensured through a more detailed approach to design, which explicitly relates child projects to the broader program objectives.

67. A program PFD should set out its overall objectives and how the program will contribute to meeting them, including the specific contribution of its child projects. GEF Secretariat stakeholders reported that, when programs are under development, there is often incoherence between child and program objectives, which leads to much back and forth between Agencies and the GEF. There have even been occasions when Council has had to get involved. However, MDBs see this issue from a different perspective, under which the GEF Secretariat sometimes applies a narrower lens on coherence than they do. For example, for one GEF-6 programmatic approach proposal, the MDB involved as GEF Agency wanted to mainstream chemicals and POPs into its broader development activities, but the GEF Secretariat did not accept the linkages. This led to a 1-½ years delay and may have lost the opportunity to blend GEF activities with the co-financing brought from the concessional loans provided by the MDB in question, which had to follow its own internal financing cycle. It therefore appears that any design

complexity experienced as a result of program participation is mainly felt by the GEF Agency or Agencies involved, rather than at country level.

68. According to a variety of stakeholders, when child projects are designed by a diverse set of agencies there can be substantial challenges to develop a coherent program. In such cases, complexity may arise at program level, since it is difficult to draw inter-relationships between projects, which have no common design basis and which are expected to fit within the management systems of different implementers. These aspects are discussed in detail in the following sections.

F. Cost-effectiveness, Efficiency and Coordination issues

Finding 6: Cost-effectiveness and efficiency decline as programs become more multidimensional.

Child projects do not differ much from stand-alone projects in terms of project cycles. Overall, child projects scored higher on effectiveness and efficiency, and leveraged higher co-financing than stand-alone projects, but efficiency ratings declined with increased complexity. Due to their diversity in terms of mandates and operational approaches, GEF Agencies often struggle to work together in the way envisaged by the GEF. Program coordination is an added cost that increases with complexity.

69. The GEF defines cost-effectiveness in Council Paper GEF/C.25/11 of April 29th 2005, which formally addresses the requirement for cost-effectiveness analysis in GEF projects at design. The paper indicates that: *“...in GEF project preparation and review, cost-effective analysis requires that a range of alternative paths to achieve a stated objective are considered and evaluated, with the most effective and least-cost approach being selected. Such work requires knowledge of lessons learned through past programming experience”*.

70. This evaluation took a more holistic approach in addressing cost-effectiveness analysis, by looking at three factors: (i) program and child project approval times as per GEF Project Cycle; (ii) program financing and co-financing; and (iii) program effectiveness and efficiency, expressed by the APR ratings.³⁶ In this approach to cost-effectiveness analysis, program-level results simply represent the sum of project-level results. If the costs of a program are less than the alternative (stand-alone project/cluster of projects), then the program is still more cost-effective. Whilst the ideal cost-effectiveness comparison between programmatic and stand-alone projects would look specifically at the relationship between costs and results, much of the available data focuses on either one or the other. The evaluation has therefore analyzed data sets on costs and on results separately, but also considered the relationship between the two on the basis of its own fieldwork and interviews in several countries and on telephone interviews with a range of stakeholders in the GEF Secretariat, STAP and GEF Agencies.

71. The *GEF Project and Program Cycle Policy*, indicates that programs and related child projects follow a similar approval process to that of full-size projects. The difference is that project concepts are submitted as part of the PFD, eliminating the need to submit a PIF for each individual project. This process consists of two main steps: *“(i) Council Approval of a Work Program that includes a Program Framework Document (PFD) together with any Child Project titles or concepts; and (ii) CEO*

³⁶ See Section A, Chapter III.

*endorsement/approval of Child Projects under the Program.*³⁷ Removal of the requirement to submit detailed child project PIFs separately to Council may have offset some of the additional time required for the PFD. Otherwise, child and stand-alone projects show similar time lines to clear GEF processes (Table 3).

Table 3: Program and Child Projects document requirement vs. Stand-alone projects

Document	Program / Child Projects	Stand-alone Project
Program Framework Document	Submitted by Agency/Coordinating Agency to be included in the Work Program	N/A
Endorsement Letter from the GEF Operational Focal Point	Required from all countries included in the program	Required from all countries included in the project
Full-sized Project – Project Identification Form	Child Project titles/concepts submitted with program for Work Program Inclusion	Submitted for Work Program Inclusion
Medium-sized project – Project Identification Form	Not required	Not required (unless for 2-Step MSPs)
Full-sized Project – Final Project Document	Fully prepared document submitted for Chief Executive Officer (CEO) Endorsement	Fully prepared document submitted for CEO Endorsement and circulated to Council for comments. 4-week web posting period
Medium-sized project – Final Project Document	Fully prepared document submitted for CEO Approval	Fully prepared document submitted for CEO Approval
Resource Commitment by Trustee	Trustee sets aside entire amount of resources for each child project	Trustee sets aside entire amount of stand-alone project

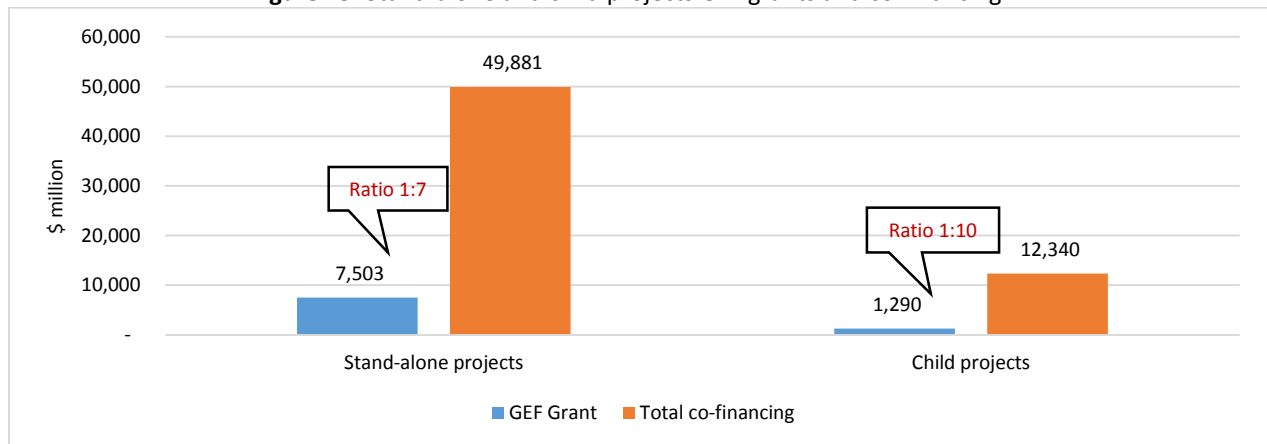
72. This evaluation extracted and analyzed project milestone data for the entire post-2008 cohort of full-size stand-alone projects (n=1748) and compared it with the post-2008 programmatic cohort (n=301). This analysis shows no major differences between child and stand-alone projects. Sixty-five percent of post-2008 stand-alone full-size projects fail to meet the GEF standards for moving from Council approval to CEO endorsement, and 31 percent fail to meet the same standards from CEO endorsement to project start-up. In the case of programs, 67 percent of full-size child projects – which are the large majority in the programmatic cohort – fail to meet the standards from Council approval to CEO endorsement, while 36 percent fail to meet the standards for moving from CEO endorsement to project start up. According to interviewed stakeholders, under some programs, child projects have been delayed as GEF Agencies waited for all of the child projects to be ready for CEO endorsement and start implementation at the same time.

73. From May 2008 to date, the GEF has allocated a 1/6th of all GEF grants to programs and their child projects. Programs, like GEF stand-alone projects, also have an agency fee for the services they provide. This fee is marginally less for programs at 8-9 percent of program total grant, than for stand-alone projects at 9-9.5 percent.³⁸ This fee reduction is linked to the removal of the requirement for programs to submit detailed child project PIFs separately to Council. Higher program cost-effectiveness is also evidenced by the potential for leveraging additional co-financing. In fact, co-financing ratios for child projects are higher than for stand-alone comparators, at 1:10 vs. 1:7 (Figure 25).

³⁷ Global Environment Facility. *GEF Project and Program Cycle Policy*. (GEF/C.50/08). 2008.
https://www.thegef.org/sites/default/files/council-meeting-documents/EN_GEF.C.50.08_GEF_Project_and_Program_Cycle_Policy_3.pdf

³⁸ Global Environment Facility. *Fee Policy for GEF Partner Agencies* (FI/PL/03). August 2012.
https://www.thegef.org/sites/default/files/documents/Agency_Fee_Policy-2012_0.pdf

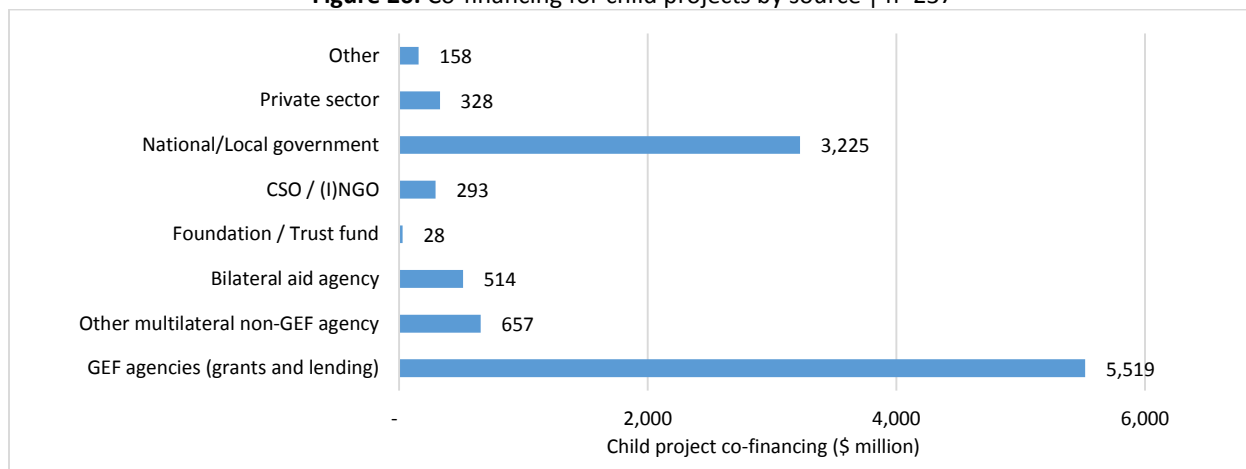
Figure 25: Stand-alone and child projects GEF grants and co-financing*



*\$1,290 million is the total grant amount for programmatic approaches excluding the Project Preparation Grant (PPG) and the GEF Agency fees; total GEF grant including PPG and Agency fees is \$1,486 million as indicated in the introduction of this report.

74. Co-financing either comes from governments or from GEF Agencies as grants or concessional lending by MDBs, suggesting increasing ownership and adoption of long term program perspectives (Figure 26).

Figure 26: Co-financing for child projects by source | n=237

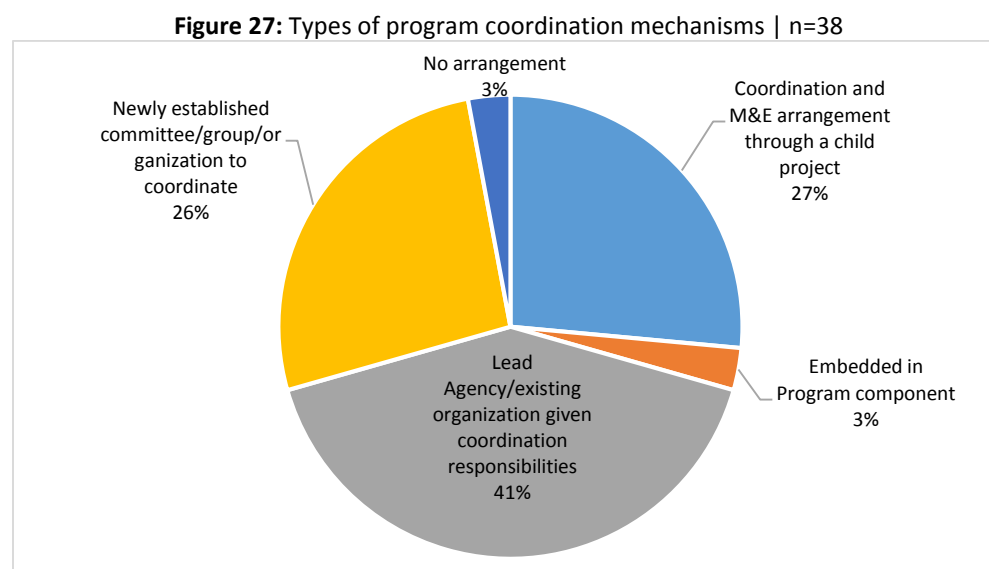


75. As seen in Section A of Chapter III, post-2008 child projects were rated higher on overall efficiency than stand-alone projects in the APR database. However, child projects effectiveness and efficiency ratings decreased with increased program complexity. GEF Agencies consider more simple programs – especially the homogeneous ones – as having lower transaction costs and being easier to manage than more complex ones. According to them, country programs are the most manageable; but only those countries with large STAR allocations are able to handle such programs. Most of the post-2008 programs involve more than one GEF Agency. As reported earlier, due to their diversity in mandates and operational approaches, GEF Agencies often find it difficult to work together. However, child projects tend to be implemented by one agency, removing this obstacle at field level.

76. As seen, at initial approval stages, programs and stand-alone projects require similar documentation and processing times (Table 3). A program however would have less initial costs and its child projects tend to leverage higher amounts of co-financing than stand-alone projects. Assuming the same results, any given amount of money spent under a program should therefore be more cost-effective than the same amount spent through stand-alone projects. This however does not take into account

program coordination costs associated with executing a program. On average, coordination costs are a little over \$1 million per program and coordination budgets increase with program size and complexity. Program coordination, institutional and management arrangements are discussed in detail in the following paragraphs.

77. While all but one program out of the 38 included in the programmatic approach cohort have some form of coordination, only eight of them have a dedicated coordination budget allocated from the program itself. Only two have specified coordination arrangements embedded in the program, with an allocated budget as part of the program cost. Half of the programs are coordinated through efforts by the lead GEF Agency (Figure 27).



78. Regardless of the type of coordination arrangement, seven of the 38 had plans for program coordination meetings, and 25 of them had national and/ or local governments involved in program coordination. Nine of the 34 programs analyzed had coordination and M&E funded through a child project. Those child projects (often referred to as ‘glue projects’) were originally mainly of medium-size, with a budget of up to \$1 million. In GEF-5, two full-sized ‘glue projects’ were endorsed with a value of \$4.6 million and \$5.5 million respectively. In GEF-6, the glue project for the Wildlife Program (GEF ID: 9071) has a \$7 million budget. The Amazon Sustainable Landscapes Program (GEF ID: 9272) has a glue project with a total budget of \$5 million, and the one for the Leapfrogging Markets to High Efficiency Products Program (GEF ID: 9083) has a \$3.1 million budget.³⁹ The three GEF-6 IAPs also have large budgets for their so-called ‘hub projects’, demonstrating that program designers have become aware of the high costs of effective coordination of programs. In cost-effectiveness terms, these newer programs will therefore have to deliver substantially improved results in order to justify these high costs. This counteracts one of the original expectations of programs, namely that they would lead to decreased management costs.

79. Successful program coordination depends on various factors including, as we have seen, the availability of funding specified for this purpose. Regional programs, often based on a homogenous model, attach less importance to coordination, reflecting what are often substantial budget constraints in this area. Other factors besides the availability of coordination funding are relevant here and the evaluation

³⁹ ‘Glue projects’ for the Amazon Sustainable Landscapes Program and Leapfrogging Markets to High Efficiency Products Program are still at PIF stage (i.e. they are not CEO endorsed yet). Their indicative budgets are extracted from their respective PFD.

has seen cases in the field of both strong (India Coastal and Marine Program) and weak (PRC-GEF Drylands Partnership for Land Degradation, China) national program coordination. Some of these differences among case study programs are illustrated in Box 9 below.

Box 9: Coordination between Child Projects and Programs

Overall, child projects under the India GEF Coastal and Marine Program have performed substantially better in terms of meeting their environmental objectives than did the counterfactual project. The most important factor in this has been the attention and coordination between high level National Steering Committees, which have helped the projects to stay on track and have linked their successes to national level arenas, including policy and strategy formulation. In comparison, the counterfactual project was taken over by State Government and was effectively re-oriented to become a livelihoods project with minimal environmental linkages or results. It can also be observed that the presence of these National Committees has raised the importance of the projects with the GEF Implementing Agency, as compared with the counterfactual project that (according to its TE) received inadequate Agency supervision, which allowed it to divert away from its GEF objective and outcomes, contributing towards its overall poor performance in terms of the original environment objectives, which attracted GEF funding.

The MENA-DELP program has a Steering Committee with representatives from each of the national implementing institutions. There is relatively little need for operational coordination, since the child projects are nationally managed and have no specific relationship to each other. Further, there are no regional program funds, which might need coordinated management. The program is almost entirely a collection of very loosely related national projects and performance is therefore managed at country level. In terms of performance delivery, there is little role for management at program level, while routine management is ensured through the World Bank project management system.

GEF Agencies in China have a shared perception that program coordination is heavy. IFAD goes further and indicates that the Partnership was designed more to share the available financial envelope, than to coordinate projects. The State Forestry Administration was quite active in coordinating with other government ministries and departments the projects falling under its responsibility. For other projects, the Ministry of Finance was more active. As for the Partnership, as a whole, there was no formal arrangement of coordination, but a few meetings were organized by the government. The Asian Development Bank confirmed that the Partnership had no funds specifically earmarked for coordination.

80. Some of the programs developed under GEF-6 (including the IAPs) have substantial resources and attempt to address major issues, including those that are global in scope. Although their funding for hub activities may look large, it is often quite small in the context of the number of countries and issues involved. Further, in the case of large programs such as Global Wildlife, it is difficult to envisage a set of stand-alone projects, which could attempt to address the same issue. Management costs, therefore cannot necessarily be demonstrated to be less than those for stand-alone projects on the grounds that: a) there could not be a set of stand-alone projects addressing the same issue; and b) the major programs are delivering a different range of benefits, such as those addressing global value chains at several different points.

81. For such large-scale programs, the GEF Secretariat often initially takes a long time to sort out the governance structures. It may meet all the Agency stakeholders several times, which has been shown to avoid management issues down the line. Once the roles are set, programs seem to settle down well. Where institutions have already worked together outside of the GEF, things tend to work more smoothly. These programs often have substantial development costs for Agencies concerned, which may not be directly recoverable and some Agencies expressed a reluctance to participate in such interventions in future, unless there is some way of meeting these costs.

82. At country level, challenges were noted as a result of including nationally designed projects in broader regional programs. For example, one NGO implementing partner in the MENA-DELP program reported considerable confusion over how its engagement would work. It initially believed that it would have a direct funding link to the World Bank to implement the activities outlined in its proposal to participate. It was later told that its proposal was included in a broader national project, which was in turn part of a regional program. As a result of the confused development, approval and start-up processes its

four-year project lost one year and the NGO was told that there could be no extension to allow this time to be replaced, even though it was not in any way responsible.

83. The GEF Secretariat is aware that programs with many partners have faced significant liaison and coordination problems, both between GEF Agencies and among in-country stakeholders. These problems are perceived to have weakened ownership. Since the new generation of programs tends to be even more complex and rich in partnerships, this will be a major challenge to overcome.

84. Some Agencies still prefer homogeneous programs, since they find these easier and more cost-effective to manage than major multifocal or multi-Agency programs. In these programs, the management costs may be relatively low, since child projects are similar and can be essentially self-managed. As seen, coordination funds are usually limited to the 'glue project'.

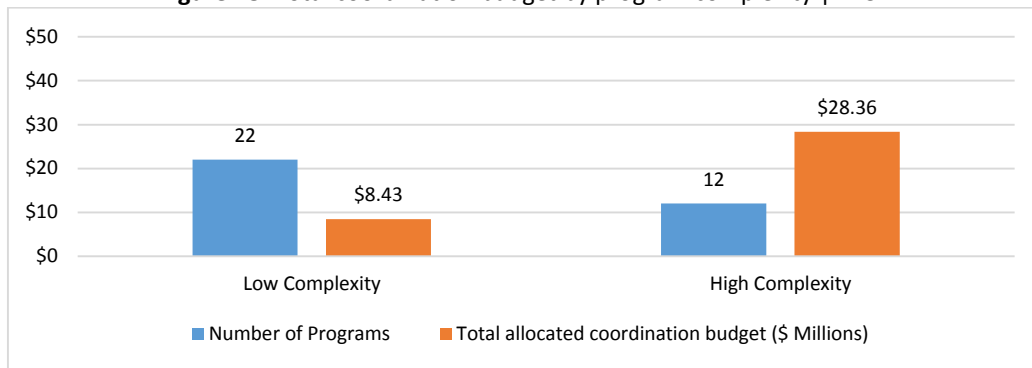
85. One approach emerges as the preference of several of the GEF's development bank partners. This is for an initial concept note followed by a flexible investment program developed over time. These partners do not favor forcing countries into programs with a blueprint design (as in many aggregation programs), which cannot easily be modified. However, the MDBs see GEF becoming increasingly prescriptive, pushing for early design of child projects, even in countries, which have had no prior commitment in the area. MDBs prefer to slowly build on prior national activities and commitments, to give a stronger platform for bigger and more sustainable results.

86. GEF Agencies perceive that in program design and implementation, a balance must be struck between being too open and allowing controlled flexibility. This is important because each country is different and needs to take its own approaches to reach the program objectives. For example, with the Food Security IAP one country may focus on agricultural extension while another does something completely different. This needs good program design with flexible entry points per country and generic options for national participation, which can be shared and discussed through hub services.

87. Knowledge management and adaptive learning, both of which are key aspects in program coordination, provide the basis for decisions on what should be replicated and what should not. The concept of knowledge sharing has received heightened importance within programs. However, there is broad skepticism among stakeholders that this is largely because funds are available, while there is no historical evidence of results or sustainability in this area and there is, so far, no clear indication of how programs will be better in this respect than stand-alone projects. The MENA-DELP is a typical example of the 'glue project' approach to knowledge sharing and M&E. Its regional medium-sized project has a \$1 million budget, with \$800,000 for knowledge sharing, \$100,00 for M&E and \$100,000 to cover management costs. The difficulty of such an approach is that the funds for program M&E are so small that there will be little credible information for knowledge sharing.

88. Interviews in case study countries revealed a perception that funding for coordination is under-resourced. Desk analysis showed that more complex programs tend to have better designed and resourced knowledge management and coordination systems than simpler ones. However, 45 percent of projects under low complexity programs had M&E allocations, compared to only 26 percent for those under highly complex programs. At the program level, five out of the 12 high complexity programs had a coordination budget, as against only two of the 22 low complexity ones (Figure 28). This suggests that while high complexity programs tend not to have a specific M&E allocation for their projects, that cost is likely to be covered under their larger budget for overall program coordination. M&E is discussed more in detail under Section G below.

Figure 28: Total coordination budget by program complexity | n=34



G. Monitoring and Evaluation

Finding 7: M&E has improved in design of recent programs, but still faces challenges.

Child projects achieved higher ratings for M&E design compared to stand-alone projects, indicating that child projects tend to be more cognizant in designing their M&E frameworks. However, these projects also show weaker implementation of M&E than their stand-alone counterparts. Highly complex programs have similar M&E ratings to simpler ones, but their ratings drop from design to implementation stages. When present, program M&E and RBM strategies are coherent with those of their respective child projects. Little evidence of program M&E has been found. When present, it is most likely because of individual GEF Agency requirements. Early evidence from the Food Security IAP suggests a strong attention to M&E design, but the systems have yet to be implemented⁴⁰.

89. M&E is mainly undertaken for projects and is performed reasonably well at that level. On both M&E design and implementation, post-2008 child projects were rated higher than stand-alone projects in the APR database (see Figure 5, Section A of Chapter III). Projects under high complexity programs show the largest drop in ratings from M&E design to M&E implementation, suggesting that their designs are unrealistic in the light of resources devoted to this activity (Figures 29 and 30).

Figure 29: M&E Design Ratings*

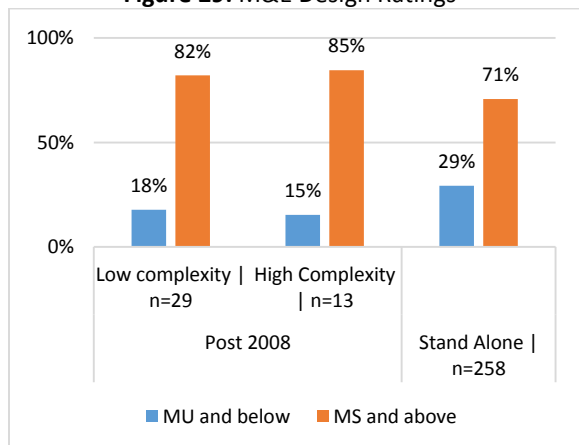
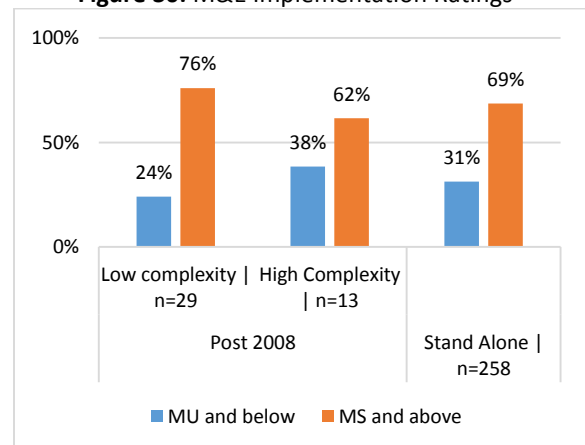


Figure 30: M&E Implementation Ratings*



* MU: Marginally Unsatisfactory; MS: Marginally Satisfactory

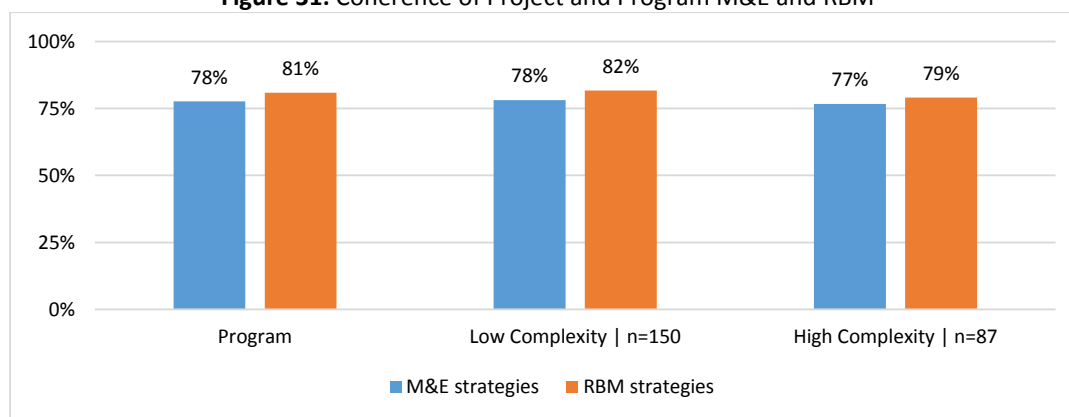
⁴⁰ M&E is currently being systematically assessed in the ongoing reviews of the three IAPs.

90. For each project under a program, the evaluation assessed if program documents indicate how that project's M&E and Results-Based Management (RBM) strategies and indicators will contribute to those of the overall program, whether the two approaches were coherent and at what level. Overall, the analysis shows that roughly half of the projects' M&E strategies are intended to contribute to program M&E. There was no variation between low and high complexity programs on this count. With regard to RBM, 61 percent of low complexity programs indicate how the project level system will contribute to that at program level, while 43 percent of high complexity programs covered this factor in their design.

91. An M&E strategy was presented for 71 percent of programs analyzed. Figure 31 below shows that over 78 percent of those with an M&E Strategy, showed coherence and alignment between program and project levels, with no variation between low and high complexity programs. However, this finding does not hold true when looking at M&E indicators. High complexity programs showed a higher level of coherence and alignment between project and program M&E indicators than the low complexity ones. Only two programs have a specific budget allocation for program level M&E.

92. The coherence of program and project RBM could only be assessed for the 17 programs with an RBM framework. The results are similar to those for the M&E strategy. Figure 31 shows that for the 17 programs with an RBM strategy 81 percent of these showed coherence with project strategies. This proportion decreased to 35 percent when looking specifically at coherence among indicators. Here, no variation was observed between programs with different degrees of complexity.

Figure 31: Coherence of Project and Program M&E and RBM



93. Despite the presence of program level M&E strategies, implementation of program level M&E is extremely rare. When it is present, this is probably because of individual GEF Agency requirements. For example, the lead agency of the PRC-GEF Drylands Partnership for Land Degradation in China, ADB, commissioned an independent program level evaluation. To note, in response to a specific Council request, the GEF Secretariat is assembling a compendium of existing and new guidelines on various GEF project and program cycle related policies. The draft version of this document is silent on responsibilities for compliance with the third minimum requirement of the GEF M&E Policy, entitled '*Project and program evaluation*'. Also, the RBM section of draft guidance document is not specific on how to aggregate program level M&E from child project level M&E up.

94. This evaluation found evidence of some cases of successful project level adaptive management, responding to M&E inputs. The M&E undertaken by case study child projects (Box 10) supports the broader data suggesting that program level M&E is rarely a systematic activity, which is well-supported by child projects and that attention to M&E activities is most effective at the level of individual projects.

Box 10: Monitoring and Evaluation at Program and Child Project level

For the India child projects, monitoring and evaluation information is reported to the National Steering Committees, which have played an important role in keeping them on-track. It appears that progress and results of activities are considered in terms of their contribution towards broad project objectives, rather than through consistent use of the project log frames or documents. The counterfactual project was governed at State level through a Project Board, which paid little attention to either the original project document or to the Mid Term Review. As a result, the project deviated substantially from its original intentions and did not deliver its environmental objectives.

In the MENA-DELP program, projects have largely used M&E for their own internal management purposes. Although they have also sent data to the “glue” project office, they claim not to know whether or how these data have been used and there has been no useful feedback from the evidence provided. So, from the perspective of child project participants, it is not clear the program has made any flexible use of M&E outputs to better achieve its outcomes and impacts. Indeed, it is not clear to the child projects that there are any tangible program level outcomes.

There is no evidence of systematic use of M&E for management purposes at program level. Project M&E systems are focused on reporting to national systems and to the WB/GEF as individual projects. The regional project has devised a program M&E system to which the individual projects should contribute. To date, there has been little enthusiasm from projects to regularly input data into this system; and it is not clear how aggregation of data from a limited set of small projects around a vast region will generate information, which could be useful for management purposes. Projects are subject to regular World Bank management missions and reporting, including Mid Term Reviews and eventually GEF Final Evaluations. However, review of documents to date shows that each project is assessed as an individual entity and that there is no monitoring or reporting on how it has contributed to the program or vice versa. There appears to be no provision in the standard World Bank progress reporting TORs to assess the role of projects in programs.

With regards to the China Partnership program, GEF Agencies expressed doubts as to whether results would have been different if the same funding was executed through stand-alone projects. However, they recognize that an important value addition brought in by the Partnership is knowledge exchange. While the only information exchange between GEF Agencies happens at design stage and concerns funding and geographic targeting, the actual knowledge exchange on lessons learned happens at local level, between counties. The Partnership has been evaluated both at the program (by ADB and SFA) and at the child projects level (by ADB, IFAD and WB). However, as seen earlier, no program level aggregation of child project M&E data was conducted in the State Forestry Administration Assessment Report. At the child project level, several completion reports, ICRs and evaluations point at the weaknesses in land degradation monitoring and assessment.

95. The GEF Secretariat has indicated that M&E poses a very important challenge for programs, requiring the development of new types of results frameworks and tracking tools, to trace different types of outcomes and impacts over an extended time. This is new for the GEF and its Agencies as well as for countries. GEF-6 IAPs explicitly address these challenges by building program M&E into a specific program component, the operationalization of which has proved challenging to date. For example, the applicability of the multifocal tracking tool of the Food Security IAP being developed by the GEF Secretariat is still being questioned by GEF Agencies and countries alike, delaying its operationalization.

96. Challenges in M&E have been present since the early generations of programs. The evaluation has found that for the majority of GEF-4 programs the common approaches to M&E, if any, were super-imposed on top of the project-specific systems, which are considered more important by the projects and countries. Rather than reducing variation, they pull out some comparable elements and enable a low level of program-wide reporting.

97. M&E therefore poses a very important challenge for the development of programs. This is not just a question of M&E, but requires the development of new types of results frameworks, which need to track and aggregate different types of outcomes and impacts over an extended time. This is new for the GEF, Agencies and countries. The M&E systems from early programs rarely look at the contribution of projects to the program or vice versa. These programs were largely developed to secure GEF grants and speed up fund flows, but there was usually no deep commitment to program level results as opposed to those of the child projects.

98. Any set of projects can yield a thematic evaluation, so the consideration now concerns how programs can have a bigger impact than projects. What is the biggest narrative that can be reported? GEF-6 has not resolved how to address the challenge of program M&E.

99. Another important aspect concerning M&E is that this is conducted very differently by various GEF Agencies. In some cases, program evaluation is independent. In others, the implementers hire the evaluators. This can cause substantial disagreements, since staff who commission evaluations say that these will affect their performance ratings, with associated reduction of independence. In the case of large multi-agency programs, M&E so far tends to be parceled out to one agency. This may not be the best approach, since the way this agency approaches M&E may not be compatible with how the other agencies work in this field.

100. One emerging approach holds that GEF needs to promote large-scale hypotheses to address the global and regional forces driving the types of environmental degradation, which it seeks to address. For example, reducing deforestation in Brazil needs a hypothesis and theory of change, which can encompass the working of the global soya market. In such programs, lesson learning is complex and should be based on progress of the contribution of different program elements towards the objective defined in the theory of change.

H. Governance

Finding 8: The roles for partners in program design have evolved with changing focus on global programs and IAPs.

Recently, the development of the IAPs and other global programs is being actively managed by the GEF Secretariat with a view to pilot new initiatives and include the newer GEF Agencies. However, this increased role of the GEF Secretariat in program design is being perceived by a number of stakeholders in the partnership as a shift in responsibilities from what was typically under the purview of the Agencies towards the Secretariat.

101. The “Instrument for the Establishment of the Restructured Global Environment Facility” (GEF, 2015) delineates at paragraph 21 the main functions of the GEF Secretariat. On the specific role the GEF Secretariat should play in program formulation, the Instrument is rather open: point (b) generically refers to an overall coordination and oversight role: “*coordinate the formulation and oversee the implementation of program activities pursuant to the joint work program...*”. In its efforts to promote programmatic approaches that are innovative and inclusive of those GEF Agencies that have joined the partnership only recently, the GEF Secretariat has somehow taken a more proactive role in program formulation.

102. A number of interviewed GEF Agencies’ representatives expressed concerns on the increased role of the GEF Secretariat in program design.⁴¹ Some have noticed this increasing engagement materialize through centrally conceived programs. They made the point that the GEF Secretariat is promoting the programmatic approach to Agencies with insufficient evidence that this is the best approach to address the relevant issue. To them, this reduces the role of GEF Agencies. According to

⁴¹ Five out of the 11 GEF Agencies involved in post-2008 programs. A stronger role of the GEF Secretariat in indicating how much STAR was to contribute to programs was also mentioned by the OFP in one of the seven countries visited during this evaluation. The OFP also stressed the importance of the alignment of programs to national priorities. In another country, the OFP said that the GEF is supposed to be country-driven. While for stand-alone projects this is easily done, it is difficult for countries to be involved in program design. Decision making at that level needs to take into account national priorities.

others, in the last generation of GEF programs the GEF Secretariat has increasingly stepped into Agency roles. For example, the Secretariat is actively participating in the selection of countries to be included in a program. Some Agencies do not consider this as GEF Secretariat's role. Although they are aware of GEF's discussions with countries, they are often not part of these meetings and hence consider the processes to be non-transparent. As a result, they question the impartiality of the Secretariat in deciding whether to advance programs to Council, given its involvement in their design and in the selection of program partners.

IV. CONCLUSIONS AND RECOMMENDATIONS

103. The evidence and findings presented in the previous chapters allowed the evaluation to reach five conclusions and to formulate three recommendations.

A. Conclusions

Conclusion 1: GEF programmatic approaches have promoted projects that are better designed to produce broader and more sustainable results than stand-alone ones.

104. At the project level, the evaluation findings indicate that the GEF program support modality in general provides better and larger scale results than project support. Programs provide a long term perspective and enable through their projects integrated solutions to the environmental challenges the GEF has been tasked to address.

105. There are several implementation challenges that need to be overcome before such results can be consistently (or at all) delivered. Importantly, program complexity has increased over time and has been associated with improved design. However, better designed and more coherent programs have also required longer times to produce results, which may not be measurable by project closure.

Conclusion 2: The multidimensional nature of programs has generated greater need for coordination and management, with implications for efficiency, results and performance.

106. The evaluation clearly shows that complexity is the most significant challenge to program performance. Simpler programs show better results. Furthermore, complex programs require much larger resources to coordinate and manage. Although designs have progressively improved, management and supervision systems have not kept pace with the increasing demands and remain focused at the level of individual projects.

107. In particular, multi-agency programs face major obstacles, posed by their different mandates, operating practices and M&E systems. Unless management and supervision systems for programs are substantially improved and more appropriately resourced, program implementations are unlikely to perform as anticipated.

Conclusion 3: Alignment of program support with country priorities has generated strong program ownership.

108. The evidence indicates that regardless whether a program is country, regional or global in its geographic scope, the more it is aligned with country priorities, the more STAR as well as other national and sub-national financial resources are committed to it. This increases the likelihood to improve performance and produce better programmatic results that are sustained in-country.

109. Although there has been a progressive shift in GEF programs from country towards global and regional level, national ownership has remained stronger for country programs. This has been overcome in situations where wider programs are strongly aligned with national priorities. In such circumstances, ownership often shows a broadening from one government department or ministry to several and, in some cases, even to private and non-governmental bodies.

110. The earlier tendency to bundle sets of loosely related country-level projects into regional programs typical on the GEF-4 period has not generated strong ownership of programmatic results, even though child projects were well-implemented. This approach is widely understood as a mechanism for financial convenience, rather than being truly programmatic and should be reduced to preserve scarce funding for more coherent programs.

Conclusion 4: Program design has improved, but M&E systems have not adapted to measure and demonstrate program level results and additionality.

111. While established project reporting systems are relatively strong, there is little progress towards assessing the additionality of programs to global environmental benefits.

112. Projects under programs are not seen differently by countries when it comes to implementation and also M&E is performed at project level. Although coherence of program design is improving, there is still inadequate attention to demonstrating the added value of a program over a set of projects. Program level evaluations would help in this sense, but are currently largely absent. Initial steps to this end have been taken through the establishment of program-wide theories of change in some recent IAPs and other global interventions.

Conclusion 5: Decision making on program design needs to reflect greater transparency and clear roles for all players in the partnership.

113. Programmatic thinking is increasingly done centrally and program designs are more and more developed in the GEF Secretariat. The development of the IAPs and other global programs is seen by a number of stakeholders as marking a trend towards centralized planning under direct management of the GEF Secretariat. They see this trend as a substantial revision of the division of responsibilities between the GEF and its Agencies, which they believe has not yet been fully articulated and assessed in terms of the requirements of the GEF Instrument.

B. Recommendations

114. The evidence collected and analyzed by this evaluation supports the following three main recommendations.⁴²

Recommendation 1: The GEF should continue with appropriate programmatic interventions, addressing issues that are likely to impede outcomes and performance, efficiency, and management, as they become multidimensional.

115. The GEF should emphasize programmatic approaches by deploying its resources catalytically to mobilize larger flows of funding that achieve impact at scale. However, the GEF is promoting increasingly complex programs, while simpler programs have shown better results. Furthermore, complex programs require much larger resources to coordinate and manage. Importantly, the GEF shows an increasing

⁴² Some of the issues raised by these recommendations point in the same direction of those provided in a recent study conducted by the World Resource Institute (WRI) on the future of the institutional architecture of multilateral climate finance. This study analysed seven major environmental funds active in the international climate change arena, the GEF being one of them. World Resource Institute. *The Future of the Funds – Exploring the Architecture of Multilateral Climate Finance*, 2017. http://www.wri.org/sites/default/files/The_Future_of_the_Funds_0.pdf

preference for multi-agency programs, although the evidence shows that these are the most difficult to implement and evaluate. Since this aspect reduces both efficiency and cost-effectiveness, program complexity will need to be better managed to ensure good results.

Recommendation 2: The GEF should continue ensuring that programs are relevant to the specific national environmental priorities of the participating countries while meeting the requirements of the Conventions.

116. The GEF should continue to ensure that finance is being channelled to support nationally determined priorities (inclusive of broad stakeholder engagement) in line with the requirements of the multilateral environmental conventions, and strengthen national capacities to plan, coordinate, implement, and monitor environmental change actions. The GEF should continue to promote multi-partner platforms in-country, provide incentives for longer term investments, strengthen national capacities and involve country partners early in the programming process, to ensure that it can respond effectively to country priorities. It should do so in all its programs, be these global, regional or national in their geographic scope.

Recommendation 3: M&E should be implemented at the program levels, with a clear demonstration of the additionality of the program over projects.

117. Program additionality over a set of projects needs to be demonstrated through a well-developed program theory of change, as well as through better information sharing on programs to enhance program monitoring, midterm reviews and terminal evaluations, which are currently largely absent. Importantly, the four M&E Minimum Requirements in the current GEF M&E Policy (2010) already apply to both projects and programs.⁴³ As programs become even more prominent in the future, the GEF Secretariat should endeavour to strengthen RBM and monitoring to better capture program results over and above the aggregation of project level results.

⁴³ In the GEF ME& Policy, program evaluations are defined as evaluations: “... of a set of interventions to attain specific global, regional, country, or sector objectives; these include evaluations or studies of the GEF focal areas, programmatic approaches, and GEF corporate programs.”