Comparative analysis of IFIs and DFIs Impact Measurement and Project Assessment Tools

Final Draft

Commissioned by GIZ Sector Project “MDBs for GPGs”
Contract no: 81251241
Project processing no: 17.2169.5-001.00

Version: 12 October 2020

About Oxford Policy Management

Oxford Policy Management is committed to helping low- and middle-income countries achieve growth and reduce poverty and disadvantage through public policy reform.

We seek to bring about lasting positive change using analytical and practical policy expertise. Through our global network of offices, we work in partnership with national decision-makers to research, design, implement, and evaluate impactful public policy.

We work in all areas of social and economic policy and governance, including health, finance, education, climate change, and public sector management. We draw on our local and international sector experts to provide the very best evidence-based support.

Table of contents

List of tables

List of figures

List of abbreviations

ADOA	Additionality and Development Outcomes Assessment
AfDB	African Development Bank
AIIB	Asian Infrastructure Investment Bank
AIMM	Anticipated Impact Measurement and Monitoring
AsDB	Asian Development Bank
ATQs	Assessment of Transition Qualities
BCR	Benefit cost ratio
BMZ	German Federal Ministry for Economic Cooperation and Development
CEDVP	Economics and Private Sector Development Vice Presidency
CPSD	Country Private Sector Diagnostics
CS	Country Strategy
DAC	OECD-Development Assistance Committee
DEA	Development Effectiveness Analytics
DEG	Deutsche Investitions- und Entwicklungsgesellschaft mbH
DEIS	Development Effectiveness Indicator System
DELTA	Development Effectiveness Learning, Tracking, and Assessment
DERa	Project Appraisal Report
DFI	Development Finance Institution
DI	Development Impact
DOTS	Development Outcome Tracking System
EBRD	European Bank for Reconstruction and Development
ECVP	Vice-Presidency for Economic Governance and Knowledge Management
EIB	European Investment Bank
EPG	Economics, Policy and Governance
ERR	Economic rate of return
ESAP	Environmental and Social Action Plan
ESG	Environmental, Social, and Governance
ESM	Early Screening Meeting
ETI	Expected Transition Impact
EvD	Evaluation Department
FCS	Fragile and Conflict-affected States
FI	Financial Intermediary
FIIRM	Financial, Impact, Innovation and Risk Management
GIIN	Global Impact Investing Network
GIZ	Gesellschaft für Internationale Zusammenarbeit GmbH
GPR	Corporate-Policy Project Rating
HIPSO	Harmonised Indicators for Private Sector Operations
IADB (IDB Invest)	Inter-American Development Bank (Inter-American Investment Corporation)
IEG	Independent Evaluation Group (World Bank)
IFC	International Finance Corporation
IFI	International Financial Institution
IMM	Impact Measurement and Management
IMP	Impact Management Project
IS-FSD	OECD Impact Standards for Financing Sustainable Development
KPI	Key performance indicators
MA	Managing Authority
MDB	Multilateral Development Bank

MIGA	Multilateral Investment Guarantee Agency
MSME	Micro, Small & Medium Enterprises
NPV	Net Present Value
NSO	Non-Sovereign Operations
OECD	Organisation for Economic Co-operation and Development
OPM	Oxford Policy Management
PAR	Project Appraisal Report
PAT	Project Appraisal Team
PCN	Project Concept Note
PCV	Pacific Community Venture
PINS	Private Sector Development Department
PRC	Project Review Committee
PSO	Private Sector Operation
REA	Rapid evidence assessment
SCBA	Social cost benefit analysis
SDG	Sustainable Development Goal
SROI	Social return on investment
SSI	Semi-structured interview
TI	Transition impact
TIMS	Transition Impact Monitoring System
TOC	Theory of Change
TOMS	Transition Objectives Measurement System
ToR	Terms of reference
UNPRI	UN Principles for Responsible Investment

1 Executive Summary

Each year Development Finance Institutions (DFIs), including Multilateral Development Banks, commit just below US$90 billion in non-sovereign private sector development financing. **This is a huge volume of finance, equivalent to about 60% of total Official Development Assistance.** We believe DFIs could have a greater development impact if they were better able to estimate the impact of their projects when making the funding decision. The **purpose of this study is to provide recommendations for six leading DFIs to strengthen their ex-ante impact measurement tools**. We worked with the International Finance Corporation (IFC), Multilateral Investment Guarantee Agency (MIGA), African Development Bank (AfDB), Inter-American Development Bank (IDB), Deutsche Investitions- und Entwicklungsgesellschaft (DEG) and the European Bank for Reconstruction and Development (EBRD).

In addition to analysing these target DFIs, the study also **examined the way development impact is measured amongst a heterogeneous comparator group**. This group comprised the European Investment Bank, two bilateral DFIs - CDC and Proparco - and five private and philanthropic sector impact investors: Actis, LeapFrog, Acumen Fund, Root Capital and Bridges Ventures. We do not cover the Asian Development Bank and the Asian Infrastructure Investment Bank, as they are either reviewing or have not yet developed respective tools.

We structured our investigation along **four dimensions,** namely, adequacy of scope in the definition of impact, robustness of tools, ease of use and extent of integration.

1.1 Adequacy of scope

Development impacts are defined and measured differently in the DFI tools. In one case the measurement of development impact is such a new endeavour that it is only at pilot stage. DFIs all adopt a partial view of impact and, in some cases, an extremely narrow approach is used, which precludes any indirect or systemic impact. IDB Invest aside, there is a scarcity of rigorous results chains in most DFI investments. This contrasts strongly with most private impact investors where impact is comprehensively defined, and each investment has a Theory of Change detailing how the investment will achieve the impacts envisaged.

Some tools assess impacts at end-beneficiary and system level - such as IDB Invest and IFC. Others such as AfDB, lack this strategic focus and focus on a partial view of direct effects. There is relatively limited evidence of how DFIs engage more actively with investees. Compared with private sector investors, they appear to learn less about investees' customers, the ultimate beneficiaries.

Some DFIs have strong strategic focus, where the overarching aim and cumulative impact of numerous investments is explicit and clear. Elsewhere, it is not immediately obvious that the portfolio of investments amounts to much more than a collection of projects rather than incremental steps towards a greater, overarching goal.

Some projects are only considered eligible for preparation as potentially fundable projects when they have already demonstrated they are financially viable. There is some evidence of DFIs awarding additional points for projects from prioritised sectors or countries with high fragility. However, this process is not fully transparent, and we found **little evidence of how DFIs address the trade-off between commercial and development returns**. There are some examples of 'good practice' from the private sector, notably Acumen, Root Capital and Leapfrog.

Making a robust case for additionality by being able to demonstrate positive net impact is fundamental to the rationale for DFIs and, ultimately, to justify their use of public funds. If DFIs displace appropriate commercial funding for projects, it is not only a waste of public money, but it can also distort existing capital markets and impede development. Of the tools reviewed, IDB Invest is closest to assessing additionality with sufficient rigour, followed by IFC and EBRD. The other DFIs' assessment of additionality is relatively weak and, in the case of AfDB, appears to be designed in a way that can completely cloud the issue of financial additionality.

1.2 Degree of robustness

A robust case for investment should involve a clear diagnosis of the problem to be resolved, the intervention that will resolve the problem and the consequences of this with an emphasis on systemic effects and impacts on end beneficiaries. Although some DFIs do measure impact at end beneficiary and systemic level, **most DFIs, with the exception of IDB Invest and IFC, lack clear Theories of Change for specific investments.** Some do not even consider most of their impacts at beneficiary or system levels. Without a basic results chain, the impact of DFI investments are not evaluable.

The **apparent absence of rigorous consideration of the counterfactual scenario is a serious weakness in DFI impact assessment methodology**. Without a clear counterfactual analysis DFIs cannot robustly estimate additionality or attribution – and, therefore, net development impact.

The use of harmonised indicators such as HIPSO and IRIS+ allows for comparability of investment performance within and between DFIs and helps assess their contribution to global ambitions. However, **few indicators relate to indirect outcomes, most DFIs only assess indicators quantitatively** and there is limited evidence of indicators being used sensitively to assess development impacts on end beneficiaries.

Evidence from impact investors demonstrates consultation with investees and their clients is important for the investor to understand their likely development impact and even whether the investment will work commercially. While most DFIs carry out varying types of due diligence with DFI investees, most of these focus on risk reduction rather than identifying non-financial investee support or understanding development impact, in particular the investee's outreach and the interface with their customers.

The emphasis on quality assurance of impact measurement tool outputs varies considerably, from 'light touch' to rather intensive. However, even where quality assurance is intensive, it is rarely triangulated with assessments made at the critical interface between the investee and their customers.

1.3 Ease of use

All DFIs, except AIIB, use tools to estimate ex-ante impact and the output of these tools has an influence on the stop / go investment decision. **Tools can improve the impact of projects funded**. They also result in dropping projects before the investment decision stage and in informing the investment decision by the management or Board. Tools vary both in terms of how their outputs are presented to Boards and in how effectively they influence investment decisions. Some tools are highly quantitative and require all issues of impact to be distilled down to a single number.

The positioning and human resource requirements of different tools vary significantly. This is a consequence of the scope and robustness associated with the tools. It also reflects the extent to which the tool is a conduit for data already collected by a range of DFI and investee staff not responsible for the tool.

1.4 Extent of integration

The only way to demonstrate the robustness of the tools' ex-ante impact assessments is to find out how the projects approved perform during and after implementation. **Some DFI impact tools have no role after the investment decision**, which prevents the dissemination of learning between the design and implementation phases of a project. This also undermines the ability of staff to assess and improve the impacts forecast by the tool.

Most DFIs monitor the impacts estimated by tools into the implementation phase of investments through their routine monitoring system. However, **DFIs use impact measurement tools less frequently than impact investors to actively encourage learning and manage impact.**

DFIs are capitalised with public money and benefit from a shareholding structure of sovereigns. However, **DFIs show limited accountability to the public for the use of their funds**. Most DFIs were hesitant to share information about how they assess development impact ex-ante. While some information on appraisal reports is made publicly available, neither the tools *per se* nor the results generated by them for individual investments are available to the public.

1.5 Recommendations

Five themes emerge that cut across all DFIs: transparency, additionality, measurement of development impact, a portfolio approach and customer centricity.

Transparency and accountability – as publicly-owned development institutions, DFIs should be at least as accountable to the public who own and finance them, as the 'best in class' in the rest of the development sector. This implies publishing the impact measurement tool and methodology for public scrutiny. Producing information in the public domain on each investment (impact score, monitoring results and evaluations) is also recommended.

Additionality – at the very least all DFIs should make their assessment of financial additionality more explicit. In addition, given the centrality of financial additionality to the rationale for using public funds to finance DFIs, the ambition should be to improve the quality of additionality assessment. This is necessary to demonstrate convincingly that DFIs are having net development impact, are not displacing other investors and explain the effectiveness of the non-financial contributions they bring to the investee.

Definition and measurement of development impact – all DFIs should clearly define development impact and have a Theory of Change for each investment to allow the project to be evaluable. They should then measure the contributions to (the action theory) and the consequences of these contributions for the direct and indirect effects (the change theory) on end-beneficiaries, as well as on the broader market system making explicit the assumptions.

A portfolio approach – to help make decisions in favour of investments that maximise development impact, while maintaining financial sustainability, all DFIs should develop respective frameworks. Among the comparator group, the Omidyar Network's "Returns Continuum" and Root Capital's "Efficient Impact Frontier" implement the portfolio approach.

Doing so will allow them to consider investments with different financial returns depending on the expected development impact.

Customer centricity – there are compelling commercial as well as developmental reasons for investors to understand the end-customers of the products and services, which they are supporting. There is considerable scope for all the DFIs reviewed to significantly improve their customer centricity.

2 Introduction

2.1 Purpose, intended users and scope of enquiry

The German Federal Ministry for Economic Cooperation and Development (BMZ), representing Germany, as an important shareholder of IFIs, has the potential to provide key impetus for reform of multilateral organisations. The following section provides a summary of the purpose, intended users and scope of enquiry[1].

The key deliverable of this study is to develop a specific and persuasive set of recommendations tailored to the targeted Development Finance Institutions (DFIs). DFI is a collective term, which can be broken down into two distinct categories.[2]

1. **Multilateral DFIs** are private sector arms of international financial institutions (IFIs or Multilateral Development Banks) whose shareholders are mainly national governments. They finance projects mainly through equity investments, long-term loans and guarantees.

2. **Bilateral DFIs** are either independent institutions, such as the Netherlands Development Finance Company (FMO) and the UK's CDC or part of larger bilateral development banks, such as the German Investment and Development Company (DEG), which is part of KfW.

These recommendations, informed by identifying lessons and practices in ex-ante development impact measurement from targeted DFIs, aim to promote specific improvements in the way target DFIs manage for impact when preparing, appraising and making decisions on their investments.

The purpose of this study, therefore, is to identify and compare relevant tools and practices used among the six targeted DFIs: IFC, MIGA, AfDB, IADB (IDB Invest), EBRD and DEG. In addition to these, the study's scope includes two other categories or tiers for the purpose of comparison with the target DFIs: a lighter and more descriptive account of institutions in which Germany is a shareholder (in particular EIB; AsDB and AIIB were excluded); and further evidence, which draws on two bilateral DFIs, private sector investors previously researched by OPM and new case studies on Leapfrog and Actis.

Table 1 - List of selected institutions

Target DFIs
1. IFC's AIMM
2. MIGA IMPACT, DEIS and PER
3. AfDB's ADOA
4. IDB's DELTA
5. DEG's DERa
6. EBRD (TOMS)
Comparator Group I (funded by Germany)
7. EIB
8. AsDB

[1] For a more in-depth information, please read the Concept Note approved by GIZ at the end of March 2020.
[2] https://www.oecd.org/development/development-finance-institutions-private-sector-development.htm

Target DFIs
9. AIIB
Comparator Group II (not funded by Germany)
10. CDC
11. Proparco
12. Actis
13. Leapfrog
14. Acumen Fund
15. Expected Impact Rating, Root Capital
16. Impact Radar and Scorecard, Bridges Venture

For each Target DFI, the study provides a detailed analysis broken down into four categories, namely (1) adequacy of scope, (2) degree of robustness, (3) ease of use, and (4) extent of integration. These are described further below. Organised in this way, these four criteria and associated sub-questions form the basis of our research framework.

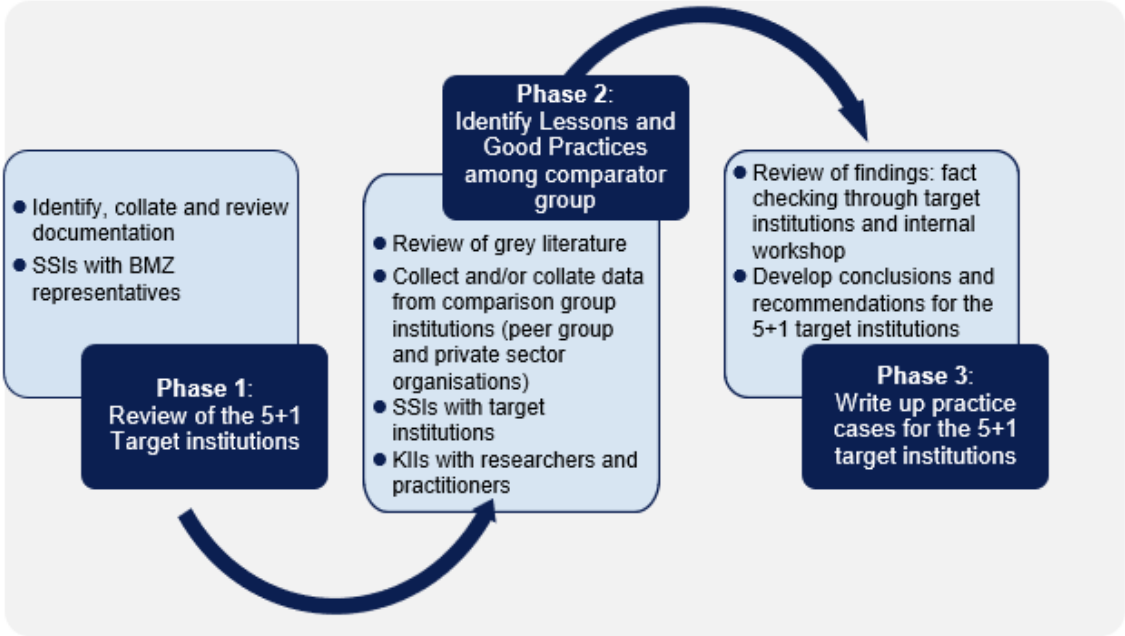
We summarise the scope of questions under each of these four categories, details of which are presented in the assessment matrix approved by GIZ in March 2020.

1. **Adequacy of Scope:** Is there a clear definition of 'development impact' based on the organisation's mandate in relation to social, environmental and economic dimensions and a Theory of Change/results framework? To what extent are the Impact Management Project's five dimensions of impact considered? Do indicator typologies cover both financial and non-financial returns? How prevalent is the use of harmonised indicators and how well are risks and assumptions treated?
2. **Degree of Robustness:** Is there use of quantitative and qualitative methods (including establishing a valid counterfactual)? Does the tool allow for respective comparability across different investment sectors and geographies? Are there ways to synthesise the evidence they generate (for instance, in an index, or a single score)? What arrangements are in place for quality control? To what extent is there consultation with the investee?
3. **Ease of Use:** How and how well are the tools used to estimate development impact and to make investment decisions? Is there clear guidance and/or training material and in what format? Is the tool resource intensive? How is the interplay between financial returns, additionality and expected development impact assessed in the project selection process? Are there clearly defined thresholds across different elements of impact for a go/no go investment decision to be made? And how well is the evidence presented to decision-makers?
4. **Extent of Integration:** To what extent and how do DFIs integrate tools and implement them throughout the project cycle, including identification, design, monitoring, learning and evaluation? Who is responsible for collecting, managing, synthesising and communicating the tool's results? On client engagement, how well are the tools linked to appraisal and implementation performance of the investee (e.g. ESG assessments and action plans)? How do the tools and the evidence they generate feed into broader results management and IT systems? To what extent do tools and evidence play a part in broader stakeholder engagement and external accountability mechanisms? To what degree are project assessments made publicly available? How are claims of impact monitored and what are the consequences in case the expected impacts are not met?

2.2 Approach, methodology and limitations

Our approach is broken down across three phases (see Figure 1).

Figure 1 - Overview on project phases



In Phase 1, we focussed on the target DFIs by carrying out an initial search and collated all relevant publicly available material. Informed by the need to prioritise information needs and fill any gaps across the four criteria, we held semi-structured interviews (SSIs) with BMZ representatives among the Tier 1 DFIs (please see **Annex I** for a full bibliography).

Following these, we held another round of semi-structured interviews with those responsible for the tools among the Targeted DFIs. (see **Annex II** for the list of interviewees and the associated checklists).

For Phase 2, we identified and collated material for the peer-group of DFIs and private/social sector impact investors in both Comparator Groups that relied on a desk review of available material. This included Interviews with key informants from the Impact Management Project, New Philanthropy Capital, the SVT Group, Actis and Acumen (see **Annex III**) to help the team identify issues, source further material on the state of play and spot examples of good practice among comparator investors from the private and philanthropic sectors for the two case studies: Leapfrog and Actis (see **Annex IV**).

In Phase 3, the findings for the target DFIs were fact checked through the interview partners and reviewed in an internal OPM team workshop, where the set of recommendations presented later were developed by the team.

Throughout the implementation of the project, bi-weekly meetings with the GIZ ensured smooth project delivery and immediate response to queries from both sides as well as mitigation of below mentioned limitations to the study.

Several limitations to this study need to be taken into account:

1. No information on AsDB's ex-ante impact assessment tools is publicly available, and these are currently under review. Therefore, AsDB was dropped from the comparison and replaced with EBRD (originally included as a comparator DFI).

2. AIIB does not have a fully developed ex-ante impact assessment tool. Therefore, AIIB was dropped from the comparison.

3. While publicly available information on MIGA's tools is limited, MIGA was still included.

4. The budget and schedule laid out in our Proposal and agreed upon in our Concept Note were defined before the COVID-19 outbreak. As a consequence of the lockdown, we experienced delays in securing the semi-structured interviews, which resulted in a delay in organising the research process and in synthesising the study's results within and across the tiers.

2.3 Definitions

In this analysis it is important to have a common understanding of key terms like 'impact', 'attribution' and 'additionality' to reduce the chance of misunderstandings based on different understandings of these key terms. Our work in this section has been drawn from a careful literature review process.

2.3.1 What is development impact?

The Practitioner Community of over 2000 enterprises, private sector investors, DFIs and MDBs, participating in the Impact Management Project (IMP) defines Impact as:

> *"A change in an outcome caused by an organisation. An impact can be positive or negative, intended or unintended".* [3]

IMP identifies 5 dimensions to help understand what impact is and provides guidance on the type of data needed to assess impact performance. The five dimensions are:

- **What**. This dimension captures what stakeholders experience as a result of an organisation's activities (i.e. outcome). For example, a stakeholder might experience a change in his/her monthly income.
- **Who**. Who refers to which stakeholders experience the outcome and how underserved they were in relation to the outcome prior to the organisation's activities? For example, stakeholders might be the poor and vulnerable.
- **How Much**. This dimension measures the significance of the outcome, considering how many stakeholders experienced the outcome (i.e. reach), what degree of change they experienced (i.e. depth), and how long they experienced the outcome for (i.e. duration). For example, the organisation activities reached 100 poor and vulnerable people, increasing their monthly income by 5% by the end of the intervention (or over longer periods of time).
- **Contribution.** This dimension helps organisations assess their contribution to the outcomes that people experience, relative to what the market or social system would have done anyway. Contribution is closely related to the concept of attribution.
- **Risk** considers the likelihood that impact will be different than expected. What are the consequences for stakeholders if the organisation does not implement its activities on time?

While the IMP's five dimensions set out a comprehensive framework to consistently defining and measuring impact, their translation into a workable system is not straightforward. The

[3] https://impactmanagementproject.com/impact-management/impact-management-norms/#anchor2

2019 Global Impact Investing Network (GIIN) global survey captures data from 278 impact investors, including private sector investors, foundations and government-backed development finance institutions. The survey findings indicate that there is a large variation in the dimensions investors examine. Most investors' impact tools specify the outcome they want to achieve, identify the target group they are seeking to reach, and estimate the number of end beneficiaries reached. However, only 40% of institutions seek to quantify the significance of impact on beneficiaries, 30% consider contribution and just about 25% consider risks to not achieving the anticipated impact or how long the effects of their investments last (GIIN 2020).

As a consequence, the evidence used to assess impact at the aggregate level is not very robust. If tools do not measure scale of impact beyond the number of beneficiaries reached it is not possible to estimate the aggregate benefit of the investment. Similarly, neglecting risks may jeopardise the outcomes that beneficiaries experience. And not assessing contribution simply nullifies claims of impact, given that the observed changes cannot be attributed to the investment, but might have been caused by other factors.

Reflecting the large variation in the way impact is understood and examined, some investors focus more on the direct impacts of the proposed investment on the client, whilst others seek to also estimate indirect impacts and broader societal, economic, and systemic impacts.

The direct impacts "…can be inputs or outputs that arise from the day-to-day activities of a company, such as the creation of jobs within the firm, or the sale of a product, or the adherence to a certain code or standard. Direct impacts are to a large extent within the control of a company." (WBC and IFC 2008). The Indirect Impacts "…are not in the enterprises'/investees' control but within their influence. They can also be characterised as "knock-on effects" of the direct impacts. They may include the creation of jobs within the supply chain or a change in the quality of life for the consumers who buy a product or service. It can also be seen as the additional value derived by other firms (small and large) that deal with the company." (WBC and IFC 2008).

As a consequence of the different definitions of impact amongst investors, specific measures of direct and indirect impact also differ: some tools define a limited range of measures, such as direct and indirect jobs created together with tax revenues generated and consumers reached. These measures are often based on a menu of globally harmonised indicators, which are limited to mainly direct effects.[4] Other tools use more bespoke indicators, and the scope of indirect impacts that they measure go beyond job creation. For example, they consider catalytic effects such as effects on communities, markets and financial systems. It is worth noting that many Multilateral Development Banks (MDBs) and DFIs distinguish between direct and systemic impacts, either through a logic model or Theory of Change (IFC 2019). This aligns with the mandate of these institutions, usually pursuing development objectives beyond the project outcomes.

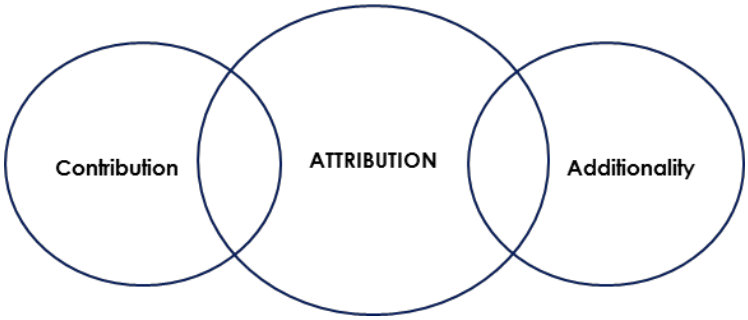
Investors' impact objectives primarily reflect the sectoral concerns of the institutions, although most investors report to use the SDGs as a framework to guide their impact measurement practice. Findings from the GIIN 2020 survey of private and philanthropic impact investors indicate that the majority aim to generate employment (71%), and about 60% of respondents focus on improving the agricultural sector and access to financial services (63% and 62%, respectively). Respondents also demonstrate strong commitments to providing basic services, such as health (60%), education (56%), and energy (56%), as well as addressing climate issues (54%). Even though the primary stakeholder groups targeted by the investors' interventions are individuals of a certain socioeconomic status, and in some cases middle- or low-income individuals, investors' impact on poverty and shared prosperity is rarely assessed (GIIN 2020). This is surprising given the poverty alleviation mandate of several investors,

[4] https://indicators.ifipartnership.org/indicators/

especially DFIs and Multilateral Development Banks (MDBs). It also indicates that the SDGs are not equally considered in setting impact objectives of investors.

The term contribution has been interpreted in slightly different ways depending on the nature of the organisation and the stage of the investment cycle. A similar observation has been made by the Donor Committee on Enterprise Development. [5]

Figure 2 - Concepts related to attribution



Source: Attribution in results Measurement: Rationale and Hurdles for Impact Investors, DCED, July 2017

Contribution is a preliminary step to get to attribution. Like the IMP, several investors either confuse contribution with attribution or consider the two concepts being the same thing. These investors usually trace the change in effects at invested companies since the start of their intervention and consider all these effects linked to their intervention. They consequently consider and communicate their intervention to have contributed to the total effects (without taking into account the contribution of other investors). The second related concept is that of additionality, which refers to the extent to which the input of an investor fills a market gap.

The assessment of contribution ex-post implies an impact evaluation analysis, comparing actual results of the investments with a counterfactual scenario without the investment.

Establishing a valid counterfactual with a control group for the development impact, as opposed to direct outcomes on an investee, of a specific project with or without the investor's support is challenging (Kenny and Moss 2020). Some authors tried to assess the impact of investments by measuring associations between the investment and some development outcomes (Velde 2011; Massa et al 2016; Broccolini et al 2019). Nonetheless, Carter et al. (2019) suggest the approaches used do not allow for causal statements of impact of the investor's support. A recent rapid evidence assessment (REA) on the impact of DFI's investments on raising incomes and increasing access to goods and services corroborates this conclusion. Much of the literature does not acknowledge the counterfactual (i.e. what would have happened without DFI investment) and counts impact as if it were wholly derived from DFI investment. The REA concludes that more and higher quality studies are needed to better understand the impact of DFI investments (Attridge et al 2019).

There exist many sets of indicators to help impact investors' measure their impact. The most commonly used are SDGs indicators, followed by indicators explicitly designed to social, environmental and financial performance of an investment: the IRIS Catalog of Metrics and the IRIS+ Core Metric Sets. Many DFIs use the Harmonised Indicators for Private Sector

[5] https://www.enterprise-development.org/wp-content/uploads/DCED-Report-on-Attribution-in-Results-Measurement-for-Impact-Investors.pdf

Operations, or HIPSO[6] (GIIN 2020) and some of the Target DFIs are on its steering group: the African Development Bank, IDB Invest and IFC.[7] Most of these indicators reflect direct, not indirect outcomes and, by way of illustration, there are few cases where SDGs 1 and 2 are used to establish a business case for the investment. Harmonisation may well bring opportunities for comparability among DFIs for direct outcomes - notably the services and products investees deliver - yet those related to HIPSO, for example, have a limited offering regarding development impact – the knock-on effects among ultimate beneficiaries and market systems.

An important aspect of the framework to understand and measure development impact revolves around sufficient involvement of the investees, potentially affected groups and other relevant stakeholders throughout the project cycle. Consultation practices are organisation specific and very little is known about stakeholder engagement processes. Recent data from the GIIN 2019 survey find that around 40% of respondents set targets in line with their investees' objectives or their investees' baseline or historical performance. Similarly, around 40% report to revise impact targets reacting to feedback from their stakeholders. Regarding the selection of impact metrics, around one third of investors consider the investee's perspective. This is surprising given that the development impact of the investors' will ultimately depend on the investees' operations. All of this reveals that investors focus primarily on their priorities, rather than the perspectives of the investees and people most affected by their investments.

2.3.2 What are impact assessment frameworks?

The IFC 2019 report on the impact investing industry outlines three dominant framework archetypes for impact measurement emerging in the market (IFC 2019):

- Impact target framework
- Impact rating framework
- Impact monetisation framework [8]

These frameworks are not mutually exclusive but might be combined within an investor's impact management system. For example, IFC's Anticipated Impact Measurement and Monitoring (AIMM) system combines aspects of impact ratings and impact targets (IFC 2019). A brief description of each framework is provided below. The description encompasses the weakness of each approach, although assessing the robustness of each framework is not possible. This section ends with a review of how investors measure contribution, the indicators used and the extent of consultation between investor and investee.

The impact target framework guides the investor to identify in what areas a specific project will generate impact and the related indicators and targets. Identification of impact areas and targets may build on a project specific or an overarching sector Theory of Change. Targets can cover multiple aspects of impact (i.e. depth, duration, type of target beneficiary), although the most common indicators measure the reach of the intervention. A typical target indicator would be the number of individuals provided with improved access to a service. Baseline data are collected at appraisal and monitored throughout the course of the investment, with impact assessed based on the progress against the targets. Because different types of investments will have different targets, this approach does not easily provide a basis for comparison between different investments, and particularly across different geographies and industries.

[6] SDGs indicators https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202020%20review_Eng.pdf; IRIS https://iris.thegiin.org/ and HIPSO https://indicators.ifipartnership.org/

[7] https://indicators.ifipartnership.org/partners/?mgc_26=45/steering-group

In addition, credibility of impact claims may be a challenge. As outlined above, many observed target frameworks tend to focus primarily on measuring number of beneficiaries and coverage of good and services provision, ignoring the benefits accrued to beneficiaries and the extent to which the investment contributed to these benefits.

The impact rating framework relies on an overarching impact scoring or rating system that can capture multiple dimensions of an investment, including direct and indirect impacts, and type of beneficiaries. Evidence of the impact assessment is synthesised in a score and rated using a numeric scale or qualitatively. Benchmarks and guidelines are established against which the specific investment may be assessed. During the due diligence phase, investments are rated based on how well they perform in terms of impact potential and the score is used to inform a go/no go decision. For approved investments, the impact score may be monitored throughout implementation and computed again at project completion. Setting thresholds for scoring often requires data on benchmarks, which may not exist and/or may be difficult to access, especially in new and innovative markets. Ensuring objectivity of the scoring process might be difficult, with clear implications on the robustness of the framework.

The **Impact monetisation framework** goes under different names, including social return on investment (SROI), benefit cost ratio (BCR), social cost benefit analysis (SCBA), and economic rate of return (ERR). The framework expresses different types of benefits (environmental, social, etc.) in monetary terms, allowing a comparison between benefits and financial costs. Given the technical rigor required, and the data needed to convert different types of benefits into an economic number, this framework can be technically difficult to implement and might rely on strong assumptions, thus challenging to use and interpret.

2.3.3 What is additionality?

For DFIs and MDBs, funded by the taxpayer, additionality is a critically important issue and refers to the principle that DFIs and MDBs should finance projects that the market would not fund strictly through private actors. Proving additionality of investment is key to MDBs and DFIs. According to the OECD-Development Assistance Committee (DAC), *"additionality of the activities should be a primary requirement"* of DFIs (OECD 2015). Without additionality, there is a risk of DFIs crowding out the private sector, which would translate into misallocation of public resources (Kenny & Moss 2020). Additionality is assessed ex-ante, and it's a key criterion for investment decision-making.

There exist at least two definitions of additionality (Heinrich 2014). A narrow interpretation, also known as financial additionality, refers to the fact that impact investors' interventions result in investments that could not have been made without the investor's financial support, and DFIs and MDBs financing is not used to displace the private sector (Kenny & Moss 2020; Eurodad 2015; IFC 2019). A more extensive definition acknowledges that additionality can also be non-financial in nature, and implies the provision of risk mitigation, improved project design, better development outcomes, better environmental, social, and governance standards, as a result of the investor's support (Kenny & Moss 2020; Heinrich 2014; IFC 2019).

Despite the importance of measuring additionality, the literature shows that additionality is hard to demonstrate at project appraisal level (Kenny & Moss 2020, Heinrich 2014). Heinrich (2014) argues that many agencies rely on assessment practices with limited or vague criteria, and inadequate assessment processes based on poor internal guidelines on how additionality should be considered ex-ante.

It is important to mention that, in an effort to develop a harmonised approach to additionality, in 2017 MDBs formed a task force with the mandate to develop more detail on the principle of additionality; common definitions; guidance on a common approach to the governance of additionality; and guidance on types of evidence that help demonstrate the presence of

additionality (MDBs 2018).[9] The approach outlines that to demonstrate additionality, a project description should include three elements: i) a narrative that describes the types of additionality present, supported by information that demonstrates the judgement based on market knowledge and/or available information; ii) a focus on key sources of additionality, identifying the most significant and relevant sources; iii) a description of why additionality matters to the project's success. The task force also compiled a list of possible evidence to demonstrate financial and non-financial additionality. The list aims at helping MDBs to assess their additionality more robustly and support their additionality claims.

The term additionality is less commonly used across private sector actors, and ex-ante assessment of contribution is not systematic. Notwithstanding, some private investors consider this dimension at appraisal. For example, Bridges Venture assesses each investment based on the investments' specific logic models, which have bespoke definitions of four criteria: 1) Social or Societal Outcomes, their scale, depth and systemic change, 2) Additionality or whether the target outcomes would have occurred without the investment; Root Capital's measurement of impact gives higher weight to investments that would not have occurred without Root Capital's intervention (OPM 2019). See section 3.3 for a description of these institutions.

[9] The Task Force included AsDB, AfDB, AIIB, EBRD, EIB, IDBG, IFC and ICD.

3 Main Findings

The main findings draw upon the research collected through document review and semi-structured interviews. A detailed overview of the findings and the analysis is annexed in **Annex III - Detailed overview and analysis on Target DFIs and EBRD**. A summary of the main findings is provided in this chapter.

3.1 Target DFIs

3.1.1 IFC

3.1.1.1 Introduction to the Tool

IFC uses the Anticipated Impact Measurement and Monitoring (AIMM) project assessment tool to estimate and manage the expected development impact of their investments. AIMM was developed in 2017 following IFC's new corporate strategy (IFC 3.0) that focuses on creating markets and mobilising private sector resources with increased support to the most difficult geographies. Following a first phase of rating projects beginning in 2017 and several refinements to the tool, AIMM is now used to assess all investment projects at IFC, with some further elements expected to be further improved in the near term. Note that IFC are currently piloting a framework to assess advisory projects, but this is beyond the scope of this study and the below findings focus on AIMM for investment projects.

AIMM is an integral function in IFC's decision-making and supports an end-to-end system for impact management. It allows IFC to measure the anticipated development impacts of all its investments across all sectors and countries at origination as well as to monitor the achievement of these impacts throughout project implementation. AIMM adopts the Impact Rating Framework archetype and the assessment of the impact is synthesised into a single rating score.

The appraisal of each investment takes place over four stages: Concept Review, Appraisal, Post-Appraisal, and Investment Review. The ex-ante assessment starts at Concept Review with investment officers (operational teams) working with sector economists and specialists from the Economics and Private Sector Development Vice Presidency (CEDVP) to establish an initial AIMM rating. This informs an initial decision on whether to proceed with the investment or not. The AIMM assessment and rating are then further refined throughout the next stages as more data is collected and analysed and the CEDVP team provides guidance on how to strengthen the potential impact. At each stage, management is involved in making a decision on whether to proceed or not. At the Investment Review stage, a final AIMM rating is determined by CEDVP management and if the project is approved, the output from AIMM is incorporated into a Board Paper that is submitted to the board for final approval. After the board approves a project, it transitions into the monitoring phase, which lasts until the end life of the project.

3.1.1.2 Adequacy of scope

IFC defines their intermediate objectives as generating positive economic, stakeholder and environmental impacts while creating markets that are sustainable, inclusive, resilient, integrated and competitive. The ultimate goal of their projects is to contribute to the achievement of the World Bank Corporate Goals (end poverty and build shared prosperity in a sustainable manner) and the SDGs.

In line with IFC's intermediate objectives and the IFC 3.0 strategy, AIMM assesses development impact along two dimensions:

- Project outcomes that measure direct and indirect effects linked to the project. These are categorised into three components: direct stakeholder impacts (on customers, suppliers, employees, local community and government); indirect or induced economy-wide impacts (value-added and employment, including job quality, as well as foreign exchange effects on balance of payments); and environmental and social impacts (emissions reduced, carbon footprint, hazardous waste, reduction in the number of workplace accidents, but also improvements in climate adaptation).

- Contributions to market creation that measure catalysed changes induced in the functioning or structure of the market beyond those brought about by the project itself. These changes are categorised into five components, called 'market attributes' by IFC: competitiveness, integration, inclusiveness, resilience, and sustainability.

The focus ex-ante is therefore on measuring outcomes as opposed to measuring impacts. For each project, the assessment is done on both project and market outcomes and these are weighed equally in the final rating. However, projects are not necessarily assessed on all of the above components under the two dimensions as the assessment is guided by the outcomes that are most relevant for each individual project. Generally, all projects would incorporate at the least an assessment of effects under project outcomes and up to two components under market outcomes.

As well as measuring the potential development impact that a project can achieve, AIMM also assesses the likelihood that the predicted outcomes materialise at both project and market level. This is done by conducting a risk assessment that identifies and analyses the significance of a range of operational, sector, country and political risk factors. The potential AIMM score is then discounted by the likelihood of impact not materialising and the final output of AIMM is a single risk-adjusted rating of the expected development impact of the project.

AIMM measures the magnitude of the anticipated effects mostly using breadth indicators (that capture the magnitude of change such as percentage decrease in power outage or percentage decrease in tuition fees) and where applicable using reach/scale indicators (that measure the number of stakeholders reached, such as number of new users of power, farmers reached, etc.) Here, AIMM assesses intensity by normalizing "gross data". In turn the breadth and reach are measured by combining estimates of the current gap of these outcomes with the intensity by which the project is addressing the relevant gap.

When assessing the direct stakeholder impacts, an identification of the different stakeholders of the project is done, including looking at whether the affected populations are typically underserved (e.g. women, low-income people, youth, refugees, etc.). For the outcomes that are most significant to a project's impact, the type of stakeholder is consistently included. The impacts on all types of stakeholders including those who are not necessarily underserved are accounted for in AIMM. However, AIMM does incentivise projects that have a specific focus on reaching underserved populations by providing an uplift to the specific claim, if this focus can be demonstrated with evidence. This uplift not necessarily affects the final score, as the overall assessment takes into account other claims on project outcomes.

AIMM also considers potential material negative effects of the project and if identified as present, these then have consequences on the AIMM score (by discounting it). According to IFC, among refinements on which IFC is currently working is guidance on the use of economic-rate-of-return analysis for specific negative impacts (economic distortions, carbon pricing, and other negative effects. As rating teams have the discretion to select the outcomes and indicators that are most relevant to their project, the internal validation panel brings these negative effects into the assessment whenever they are not identified beforehand.

While additionality is not part of AIMM, project teams conduct an assessment of additionality on all projects and this is one of the components upon which a decision to invest is made. Additionality is measured based on financial and/or non-financial aspects.

3.1.1.3 Degree of robustness

AIMM allows for measurement that is project specific while at the same time grounded in systematic frameworks that allow comparability across projects. The underlying analytical tools are sector frameworks that define theories of change for each sector and provide sector-specific guiding principles for the selection, measurement and scoring of the project and market level outcomes. They include a set of core project and market outcomes that are relevant for the sector; set of breadth and reach quantitative and qualitative indicators to measure each outcome, some of which are aligned with harmonised frameworks (e.g. HIPSO); benchmarks to determine the relative size of the gap and intensity of impact; market typologies that define stages of market development within each sector; and other useful guidance.

For each project, the rating team would have to define a development impact thesis specific to the project and, based upon that, determine the core project and market outcomes (from the list available in the sector framework) related to that project. Only those outcomes are then measured. While AIMM's Internal Validation Panel is structured to be independent of operations, it is unclear if there are any unwarranted incentives about outcome/indicator selection in order to maximise the AIMM score.

For each outcome identified under project outcomes, one or more indicators are selected to measure that outcome (from the Sector Framework). Two measurements for each indicator are then established: one that defines the current development gap of that indicator in the country and sector of operation, and one that defines the magnitude of the intensity with which the project will address the gap. The combination of these two measurements provides an assessment of the project's contribution to that specific outcome relative to the country context (see figure below**Error! Reference source not found.Error! Reference source not found.** for an example). The measurement of the intensity is normalised taking into account the size of the project so as not to penalise small projects.[10] The evidence relies on a range of data sources and analytical methods (such as quantitative multipliers based on Input-Output/Social Accounting Matrix models and other economic assessments) as well as on the judgement and structured dialogues between the investment teams and the Sector Economists. The benchmarks in the sector frameworks are used to provide a rating for the size of the gap and intensity. Based on the gap and intensity ratings for all selected indicators, the team provides a judgement on the overall rating for project outcomes potential. The bigger the gap and stronger the intensity, the stronger the potential for impact.

Figure 3 - AIMM example of the assessment of project outcomes' potential

Project Dimension		Gap Addressed			
		Low	Medium	Large	Very Large
Impact Intensity	Significantly Above Average	Strong	Strong	Very Strong	Very Strong
	Above Average	Moderate	Strong	Strong	Very Strong
	Average	Moderate	Moderate	Strong	Strong
	Below Average	Marginal	Moderate	Moderate	Moderate

Source: IFC, AIMM General Guidance Note, March 2019

[10] Normalisation by country or market is done for development gap indicators.

For each component identified under market creation, three qualitative indicators are required: one that defines the current stage of market development with respect to that component; one that determines the magnitude of expected advancement of the market resulting from the investment; and one that defines the channel through which the project advances the market. Essentially, the assessment takes the form of a narrative for how much the project is advancing the specific aspects of the market guided by these three indicators. The market creation assessment is based on the market typologies defined in the Sector Frameworks and Country Private Sector Diagnostics that are conducted at country level using evidence-based assessments, among a broader set of market-relevant information used by IFC's sector economists, as well as investment team's knowledge of the specificities of the project. An overall rating for market creation potential is determined based on aggregating the ratings of the individual components (see Figure 4 for an example).

Figure 4 - AIMM example of the assessment of market creation potential

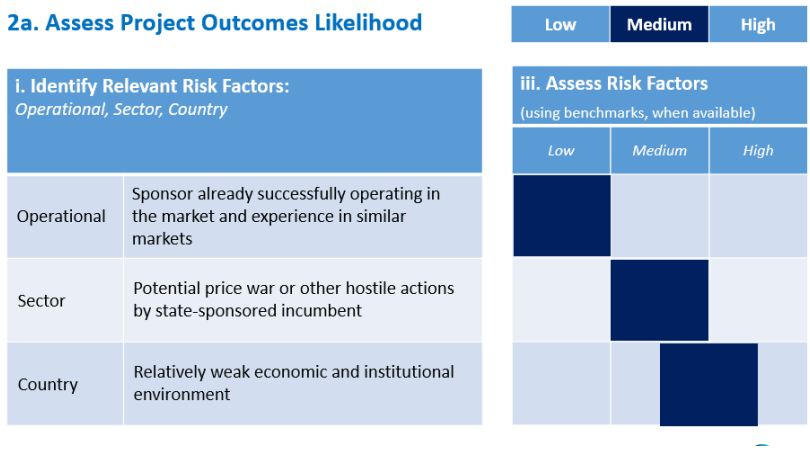
Market Dimension		Market Typology			
		Highly developed market	Moderately developed market	Underdeveloped market	Highly under-developed market
Market Movement	Highly Significant	Strong*	Strong*	Very Strong	Very Strong
	Significant	Moderate	Strong	Strong	Very Strong
	Meaningful	Marginal	Moderate	Moderate	Strong
	Marginal	Marginal	Marginal	Marginal	Marginal

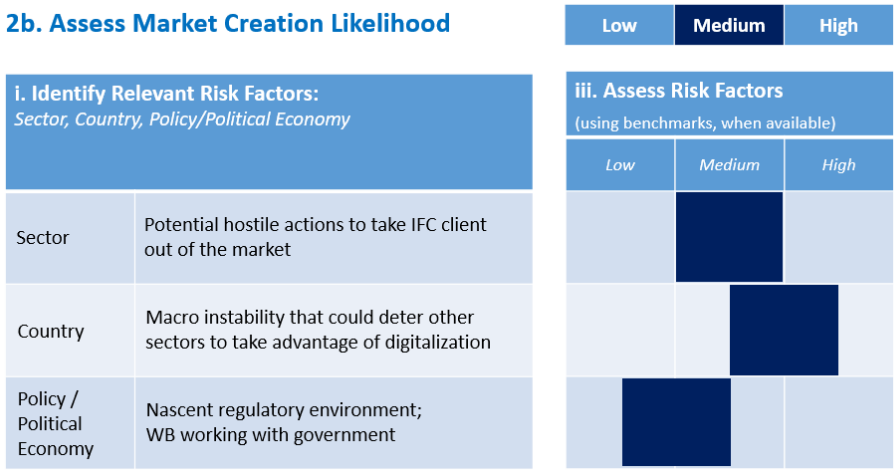
*In exceptional cases where the stage of market development is clearly documented as "Moderately developed" or "Highly developed", the intervention is the first in the market, and specifically targets IFC's underserved groups or a vast swath of customer segments, a "Very Strong" rating may be considered if the impact intensity is highly significant. Note, a market can move beyond the "highly developed market" typology as it is highly developed in the context of other emerging markets, and move into the market typology that characterizes developed economies for that specific market.

Source: IFC, AIMM General Guidance Note, March 2019

The project outcome rating is then discounted by the likelihood of project outcomes materialising, and similarly for the market outcome rating (see Figure 5 for an example). These are then added to result in a final and single risk-adjusted rating/score.

Figure 5 - AIMM example of the assessment of project and market outcomes' likelihood





Source: IFC, Presentation on AIMM, 29 March 2018

The individual AIMM scores for all projects can be aggregated into a portfolio score, which is the simple unweighted average of all AIMM scores of projects in that portfolio (regardless of the investment amount). This portfolio score is the average rating of expected impact of all projects. For example, an average of 50 at the portfolio level would mean that on average, projects in the portfolio have a 'good' impact score.

Furthermore, there is a good internal quality assurance process in place. Sector economists are involved in the assessment from the start, reviewing and validating the work by the investment officers. Projects that have a high ex-ante AIMM score will have to be reviewed and signed off by an internal AIMM panel before it goes to management for approval – and this panel is at arm's length from the operations team. An independent quality assurance firm also provides quality assurance on AIMM data annually.

Additionality relies on a qualitative evidence-based assessment. IFC have developed an additionality framework that sets guidance on how additionality should be assessed and articulated for each project and how the additionality claims need to be associated with indicators that can then be measured during monitoring.

3.1.1.4 Use

AIMM is a critical tool for decision-making. Besides providing a project-level rating of potential impact, AIMM scores can be aggregated at different portfolio levels. This is a simple and unweighted average of the AIMM scores across all projects signed within that portfolio. Aggregation can be done at different levels including region, country, sector, corporate and departmental.

The output of AIMM is presented in a Board Paper that includes a structured narrative for the development impact section. This includes a description of the main project and market level outcomes, the AIMM score, and the list of indicators used to substantiate the impact claims and that will be monitored during project implementation. The Board Paper also includes summary of other information such as, additionality, environmental and social risks, the fit of the project with country, regional or IFC strategies and description of market, sector, operational and credit risks. This forms the basis for the review and approval of the project. As mentioned in the introduction to AIMM, management is involved in decision-making throughout all stages of project appraisal, while the Board reviews and makes the final decision on a project at the end of the Investment Review stage. The Board reviews all projects with the exception of sub-projects within frameworks that are delegated to management (the frameworks though are approved by the Board).

The AIMM score is a key consideration when making a decision on whether to approve investment in a project. There is no minimum threshold on the AIMM score defined at project level to screen out investments. Rather, IFC adopts a portfolio wide approach. There are two targets at the institutional level for the development impact: average AIMM score of 50 across all projects signed in that year, and at least 15% of projects rated as 'very strong' on the market creation potential. The decision therefore is informed by the relation of the project AIMM score to the institutional targets. There are other considerations that affect the decision to approve, key among them profitability and additionality of the project. IFC are exploring ways in which they can visualise the financial and development impact returns of all the projects in their portfolio to allow for better management of the portfolio and to facilitate decision-making at the project level by balancing these two key dimensions.

By its nature, the AIMM score can be compared across different projects in different sectors providing a relative assessment of the strength of development impact. However, given that the AIMM assessment is sector specific and each project selects relevant core outcomes, the AIMM scores of projects across and even within sectors are not based on the same dimensions and this should be kept in mind when comparing scores. For example, if one project has a higher market creation score than another project, it would not follow that the former project is inducing more competitiveness in the market because it may be that only one of these projects was assessed against competitiveness or neither. The comparison of AIMM scores across projects is useful to signal which projects have a higher potential for impact.

AIMM provides a steer to invest in countries and sectors where the development gaps are widest, which are typically in low-income countries and Fragile and Conflict-affected States (FCS) environments. This is not done through applying a direct bonus or uplift to the AIMM score for low income or FCS countries but rather indirectly through the AIMM scoring mechanism. For each outcome and indicator assessed, the bigger the size of the gap or lower the current level of the market, the higher the AIMM score. Even though the likelihood of outcomes materialising in those countries is typically lower, the uplift to the AIMM score resulting from addressing wide development gaps / underdeveloped markets outweighs the discount to the score resulting from the likelihood assessment.

The AIMM tool is accompanied by extensive guidance including all the Sector Frameworks, sector guidance notes and a general guidance note, among others. The CEDVP team have delivered over 100 training and briefings on the application of AIMM during the first year of its roll-out alone, and continue to work very closely with the investment teams on the assessment of development impacts. When it comes to the resources required to conduct an AIMM assessment, besides the investment team, sectoral and economics expertise is required and the tool relies on a number of extensive analytical frameworks (sector frameworks, market typologies, country diagnostics and economic models).

3.1.1.5 Integration

AIMM is integrated into the monitoring cycle. All indicators defined ex-ante are tracked annually and the rating is recalculated and compared to the ex-ante score. Given AIMM is new, it is not yet fully integrated with the monitoring system, with indicators related to market creation mostly missing. However, IFC is working to update the monitoring system so it is fully integrated with AIMM and tracks all the indicators assessed ex-ante for both project and market outcomes. The results can also be aggregated into a portfolio score (simple unweighted average of real-time scores across all projects in a portfolio) to give a sense on how well the portfolio is delivering on its anticipated outcomes. The comparison provides an opportunity to portfolio managers to take corrective actions if needed.

In addition to monitoring the indicators related to the ex-ante impact claims, IFC also tracks a number of other standard indicators for reporting purposes such as payment to government, wages and benefits, number of female employees, number and volume of loans, etc. Results from a selection of indicators – mostly reach indicators – are aggregated at portfolio level and

mapped to the different SDGs to illustrate IFC's contribution to the SDGs. The CEDVP team also maintains a database that includes AIMM scores for all projects (both ex-ante and actual) and monthly reports on these scores are shared with sector teams. The results from AIMM and the monitoring in turn feed into targets on development impact in corporate level scorecards. We have seen evidence that these corporate level scorecards have been submitted to the Board on a quarterly basis in the last financial year; however, we do not have evidence on how often they are available publicly.

AIMM is also integrated within the overall results measurement system at IFC. Front-end diagnostics such as Country Private Sector Diagnostics inform the ex-ante development impact assessment conducted using AIMM, which feeds into the monitoring of results.

IFC also conducts a number of evaluation assessments; however, it is not clear what the formal link is between these activities and AIMM. These involve:

- A self-evaluation conducted by CEDVP team typically five years after approval. However, this is only done on a random sample of projects). All of these self-evaluations are validated by the World Bank Group's Independent Evaluation Group (IEG).
- IFC and IEG also conduct demand-driven evaluations that are often thematic based as opposed to evaluating the effectiveness, impact or sustainability of individual investments.

Internal stakeholder engagement and accountability on IFC's management of development impact is quite strong. All projects are approved by the board and the ex-ante scores as well as quarterly reviews of average portfolio AIMM scores are shared with the board. While a list of all projects is disclosed publicly and published on the IFC website, the AIMM scores are not shared and there is only a brief narrative on the main expected development impact for each project. The AIMM scores are also not consistently shared with clients. In their annual report, IFC aggregate some development impact results and show some of the portfolio level AIMM scores (but not all).

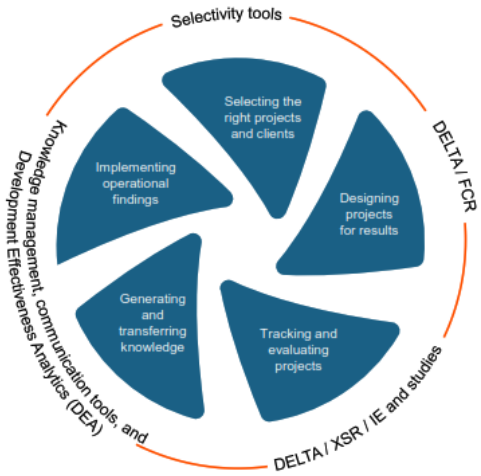
3.1.2 IDB Invest

3.1.2.1 Introduction to the Tool

The IDB Group is comprised of two separate legal entities: The Inter-American Development Bank (IDB) and the Inter-American Investment Corporation (hereafter, IDB Invest). IDB Invest is the private sector arm of the IDB Group. The Development Effectiveness Learning, Tracking, and Assessment tool (DELTA) is an integral function of IDB Invest's Impact Management Framework.

The Impact Management Framework is an end-to-end system that includes a series of tools and practices to support the complete project lifecycle (see Figure 6). The framework is rooted in a portfolio approach, which seeks to balance impact and financial sustainability. The DELTA is a fact-based scoring system that assesses the expected impact of each investment at origination and tracks development performance throughout implementation, facilitating decision-making and portfolio management (DEO 2019).

Figure 6 - IDB Invest' Impact Management Framework



Source: Development Effectiveness Learning, Tracking, and Assessment Tool – DELTA. Audiovisual presentation – Technical Briefing Washington DC, July 2018

The DELTA has been developed by the Development Effectiveness Division of IDB Invest. Development effectiveness officers are appointed for each IDB Invest project and are accountable for applying the DELTA tool. Together with the investment officers and risk officers, they make up an integrated team. The teams assess both financial and development impact returns of the investment.

As part of the assessment, the teams produce the two parameters that make up IDB Invest's portfolio approach (i.e., Portfolio 2.0): the DELTA Project Score and the Financial Contribution Rating (FCR), which measures the financial contribution of each operation to IDB Invest, based on the risk-adjusted return on capital. The output of the DELTA assessment and the FCR are included in the transaction documents and presented to the Board. The Board approves new projects based on the interplay of the DELTA Project Score and FCR, both at the project level and in relation to the overall portfolio performance (Semi-structured interview).

3.1.2.2 Adequacy of Scope

There is no explicit definition of development impact made by the DELTA. The tool understands impact as the development outcomes (i.e. benefits) generated by the investment (DELTA Audiovisual presentation 2016). Benefits are described as the effects of the investment on the economy and society, with particular focus on the development priorities identified by the Sustainable Development Goals (SDGs), as well as IDB Invest's institutional priority areas (i.e., climate change, gender equality and diversity, MSMEs, and serving the region's smaller economies and small and island countries). The DELTA assesses both direct and systemic effects, as well as the characteristics of the stakeholders affected by the investment.

The tool defines three types of direct benefits (DELTA Audiovisual presentation 2018):

- environmental benefits,
- increase in productivity, and
- improvement of products and services to better address beneficiaries' needs.

In addition, the DELTA identifies systemic effects, which indicate the extent to which the investment produces effects beyond its end beneficiaries. The DELTA understands systemic effects as (DELTA Audiovisual presentation 2018):

- The project improves or expands market linkages.
- The project involves innovation and/or leads to transferring of knowledge, practices or technology.
- The project improves policies, regulations, legal frameworks, or business environment and tests new frameworks/regulations.

The DELTA stakeholder analysis explores to what extent the investment improved access to products and services for beneficiaries, as well as who benefits from the goods and services provided by the investee and the type of benefits the project generates. This assessment of beneficiaries takes into account the economic, and socio-demographic characteristics of the stakeholders. This is limited to four stakeholder groups identified by IDB Group's strategic priorities (the poor and the vulnerable, the excluded populations, women and MSMEs). Impact on beneficiaries is understood as improved access to/quality of good and services, and adequately captured by the project result matrix indicators. However, no information is available on the extent to which IDB Invest engages and solicits feedback from the beneficiaries.

The development impact assessed by the DELTA is aligned to IDB Invest's mandate to finance sustainable enterprises and projects to achieve financial results that maximise social and environmental development for the region.[11] It also aligns to the focus areas of the IDB Group's institutional strategy. The 2015 Update of the Institutional Strategy identified three development challenges – social inclusion and equality, productivity and innovation, and economic integration – and three cross-cutting issues – gender equality and diversity, climate change and environmental sustainability; and institutional capacity and the rule of law. Looking at the results from the stakeholder analysis one can align the project contribution to the three development challenges, as well as the three cross-cutting issues.

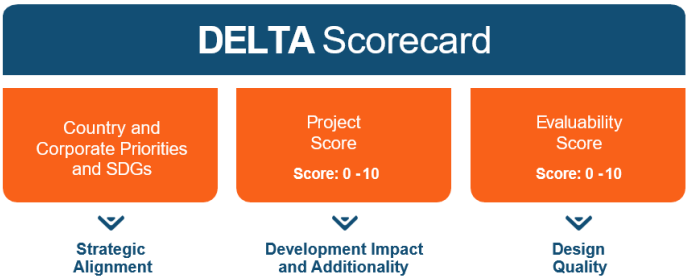
Both direct and systemic effects, as well as the extent to which the project reaches its beneficiaries are measured by indicators, for example number of jobs created and percentage of jobs filled by women. Indicators are project specific and reported in the project results matrix, which translates the project logical chain from inputs to outcomes.

The DELTA tool assesses three dimensions of IDB Invest's operations: Alignment with country and corporate priorities and SDGs, development impact potential including additionality (i.e., Project Score), quality of the design at entry to ensure evaluability of results throughout the life of the investment (Evaluability Score). See Figure on DELTA Pillars below.

The DELTA Project Score is based on the Development Outcome assessment and an Additionality assessment. The Development Outcome assessment looks at the direct and indirect benefits of the investment, the socio-economic characteristics of the stakeholders (i.e. investment contribution to social and economic development), the sustainability of the project and alignment to ESG requirements. The DELTA follows specific, data and fact-driven guidelines to rate the different development outcome categories from "Somewhat" to "Exceptional". Higher scores are driven by the level of impact and value added determined in each category, as well as the quality of supporting evidence provided. The rating system also reflects the duration of the change in terms of when effects occur and whether they are sustainable. A results matrix is presented ex-ante, as part of the eligibility proposal and its robustness is assessed by the evaluability component of the DELTA.

[11] https://www.idbinvest.org/en

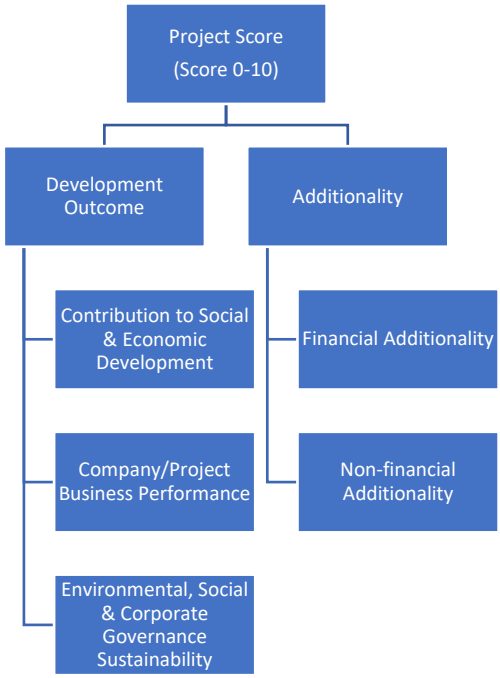
Figure 7 - DELTA Pillars



Source: Development Effectiveness Learning, Tracking, and Assessment Tool – DELTA. Audiovisual presentation – Technical Briefing Washington DC, July 2018

The development outcome assessment carries the most weight in the project score, to prioritise projects with high expected development outcome, over additionality of IDB Invest's support. Within the analysis of development outcomes, three aspects are taken into consideration: project contribution to social and economic development, financial sustainability (Company/Project Business Performance) and environmental, social, and corporate governance sustainability (i.e. compliance with ESG requirements). The tool takes into account financial returns, with higher development returns expected for projects with lower financial returns. Additionality takes into account financial and non-financial additionality.

Figure 8 – Elements of DELTA Project Score

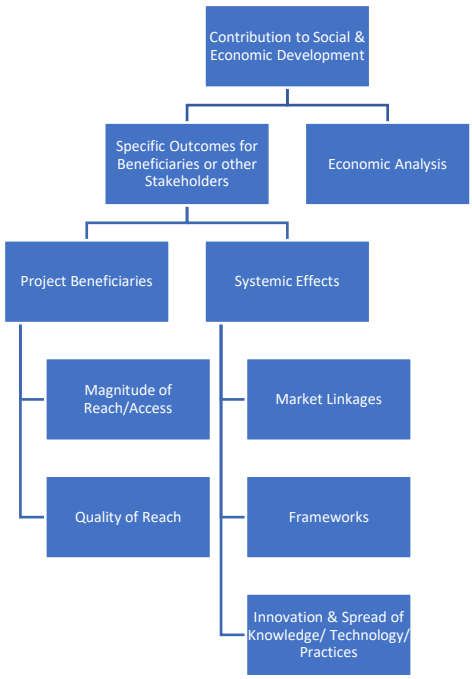


The evaluation of the development impact potential takes into account risks and assumptions underpinning the realisation of impact. The DELTA embeds risk assessment into the calculation of the project score, by running sensitivity analysis and developing multiple projections of impact based on the project logic and contextual factors. The final scenario reflected in the project score represents the most likely case from several impact projections and it is monitored and validated throughout implementation. If the project does not perform as expected, assumptions and risks are reassessed, and the project score is adjusted (SSI).

As part of the Development Outcome assessment, the tool implements an economic analysis measuring economic and social return of the investment (see Figure 9 below). The tool uses a social discount rate to determine the present value of the project benefits, and costs

occurring in the future. In addition, sensitivity analysis is performed to test the robustness of the social and economic rate of return against different scenarios (SSI).

Figure 9 - DELTA analysis of development outcome



Source: Based on Development Effectiveness Learning, Tracking, and Assessment Tool – DELTA. Audiovisual presentation – Technical Briefing Washington DC, July 2018

3.1.2.3 Degree of Robustness

The assessment of the anticipated impact undertaken by DELTA is robust and evidence based. It is informed by the project logic, as well as relevant sectorial studies and impact evaluations.

IDB Invest's projects present a results matrix as part of the project eligibility proposal articulating outputs and outcomes and the project development objective. The matrix includes specific indicators and targets relevant to the development objective, including indicators to track the project's contribution to identified SDG targets. The client usually provides output indicators (i.e. Intensity of treatment from the investment plan). The Development effectiveness unit estimates outcome indicators and targets based on project information and secondary sources such as sector studies, regional sectoral surveys and impact evaluations (IEs). The DELTA officers use econometric models and estimation methodologies, which differ by sector. Although counterfactual thinking underpins the DELTA, attribution is not systematically measured. IDB Invest's impact evaluation strategy calls for the use of IE only to fill knowledge gaps and test innovation, with particular focus on projects highly relevant and representative of existing or potential business lines of IDB Invest. Project level indicators broadly align with the direct and systemic benefits set out by the DELTA. The results matrix also reports indicators for type of beneficiaries, depending on the project targeting. The project logic draws on the evidence from the literature, with typical business lines relying on similar theories of change (SSI). The DELTA evaluability assessment evaluates the robustness of the project logic (from development challenges to impact), the evaluability of the indicators, as well as the project M&E plan (DELTA Audiovisual presentation 2016).

The assessment relies on a comprehensive set of indicators, and whenever possible, these indicators align with standardised metrics from the IRIS+ system and the Harmonised

Indicators for Private Sector Operations (HIPSO) (SSI). Core indicators align with the IDB Group Contributions to Development Results Indicators, included in the Corporate Result Framework 2020-2023. [12] There are 27 IDB Group Development results indicators, which track the magnitude of IDB Group contributions to the three challenges and cross-cutting issues of the IDB Group strategy.

A scoring system is used to assess the project development outcome. The scoring system is based on a discrete, non-linear scale using qualitative categories, which correspond to numeric scores from 0 to 10. Higher scores reflect higher development outcome, with the project contribution moving from somewhat, yes, to exceptional. The scoring process is based on data and fact-based guidelines. An overall rating is derived by aggregating the scores for direct and systemic benefits, the results of the economic analysis, as well as the business financial sustainability score and compliance with ESG requirements.

Consistent guidelines are available to support the assessment. The assessment is appropriately assured through several validation exercises with a qualified team of experts.

A similar approach is used to assess additionality. The DELTA rates additionality on a 3-point score: Exceptional, Yes and Somewhat. Financial and non-financial additionality is demonstrated through a qualitative assessment supported by evidence from multiple sources (DELTA Audiovisual presentation 2018). Guidelines are provided to assess all dimensions of financial and non-financial additionality. Guidelines drawn on international best practices across MDBs working with the private sector. [13]

While investees are consulted to prepare the project documents for approval, it is not clear whether consultations explore the investees' commitment to impact, and system of incentive tied to impact performance to ensure alignment of goals between IDB Invest and its investees.

3.1.2.4 Use

The DELTA has options that can be adapted by sector and financial instruments, for each of the DELTA modules. The DELTA formalises its assessment through a scorecard reporting the ratings for each of the DELTA modules. See Figure 10. The scorecard alongside the detailed assessment are also included in the transaction document as appendix.

Figure 10 - DELTA scorecard

Source: Based on Development Effectiveness Learning, Tracking, and Assessment Tool – DELTA. Audiovisual presentation – Technical Briefing Washington DC, July 2018

[12] Level 2 indicators – Annex A CRF 2020-2023
[13] MDB Harmonised Framework for Additionality in Private Sector Operations.

The project score is revised and reissued during the design stage as new information becomes available. A preliminary DELTA is included in the eligibility proposal and the investment proposal for the Investment Decision Meeting. The final assessment, encompassing the economic analysis is incorporated in the Investment Proposal used for approval (DELTA Audiovisual presentation 2016). The Investment Proposal used for approval includes the overall score, the development outcome score and additionality score. It also includes a description of the expected impact and additionality. A DELTA summary reporting the assessment by module and component including scores by subcomponent, is also attached to the document and made available to the Board.

While the project score is an effective synthesis of the development impact potential of an investment, its interpretation is not straightforward. The rating reflects different benefits and types of additionality depending on the nature of the project and context, with similar rating potentially driven by different dimensions of impact. This limits the comparability of the scores across sectors or geographies.

Decision-making is informed by a portfolio approach, which optimises impact and financial performance simultaneously, moving towards a portfolio of projects, which ensure financial sustainability and achieve the maximum development impact. Investments are selected based on their expected development impact and risk-weighted financial returns. Higher impact potential is required for projects with low financial contribution to IDB Invest. Investments need to meet certain project score and FCR thresholds to be approved. The minimum project score required for an investment to advance depends on the FCR; the lower the FCR the higher the required DELTA (SSI).

The Development Effectiveness team provides both internal and external capacity building, including training sessions to members of the Board. Technical briefings are regularly held to present the DELTA tools and its applications. The team also produces knowledge products, such as sector studies, and knowledge gap analysis. The level of effort in the ex-ante analysis depends on data availability and type of transaction. Conditional on data availability and its quality, the assessment may take a week, on average. Validation exercises and quality assurance may increase the level of effort (SSI).

3.1.2.5 Integration

Based on the documentation available, DELTA appears to be integrated in the project life cycle, with measurement of development outcome and additionality reassessed and validated throughout implementation[14].

The DELTA substantially contributes to the oversight and accountability mechanisms of the institution. DELTA metrics are collided in the Development Effectiveness Analytics platform and used to provide management with real time data to identify potential issues and take corrective measures.

> *"…DELTA analytics dashboards provide a real-time snapshot of the portfolio in terms of development impact" (DEO 2019).*

DELTA metrics also contribute to the IDB Group's Corporate Results Framework. There is also accountability to the public as the DELTA metrics are shared with the investees and reported in the IDB Group's Development Effectiveness Overview on an annual basis. However, only aggregate values are reported. Some project level information are reported for complete projects and from impact evaluations.

[14]Semi-structured interview: "The DELTA is fully integrated in our Impact Management Framework, which allows managing knowledge throughout the project cycle. Detaching the DELTA from the framework would reduce its potential."

Lessons learnt from previous projects are stored in the IDB Invest's knowledge management system and included in a repository accessible through the Development Effectiveness Analytics (DEA) platform. Lessons learnt feed into the DELTA assessment of new operations, helping to refine assumptions on development impact potential and improve risk assessment (SSI). Lessons learnt are not reported by the DELTA outputs, but captured by other documents, such as sector reviews of lessons learned, and loan proposals.

3.1.3 MIGA

3.1.3.1 Introduction to the tool

MIGA developed IMPACT in 2018, with full implementation expected in 2020. The tool mimics IFC's ex-ante assessment framework used in IFC's AIMM, although it has been adapted to MIGA's roles and mandates.

IMPACT is used at origination and diligence stages. The tool assesses a project's anticipated development impact, measuring both project-specific and beyond-project outcomes. The expected development impact rating as measured by IMPACT is then used as an input for project prioritisation and approval.

This report mainly builds on the findings from the first 10 months of piloting of IMPACT, as presented to the Executive Directors in April 2019. This review primarily focuses on the scope of the development impact measured by IMPACT and the robustness of the assessment undertaken by IMPACT.

At the time of writing this report, there is limited evidence on how IMPACT is used for decision-making and the extent to which the tool is integrated into the project lifecycle and more broadly into MIGA's accountability mechanisms. To our knowledge, MIGA's Strategy for FY 21-23 touches on some of these aspects, although details are not available yet.

3.1.3.2 Adequacy of Scope

MIGA understands development impact as MIGA's contribution to the SDGs and the World Bank Corporate Goals to end extreme poverty and boost shared prosperity in a sustainable manner. IMPACT measures development impact along two dimensions:

- Expected project outcome.
- Expected foreign investment contribution.

The Project Outcome dimension covers MIGA's project-level contribution to development impact while the Foreign Investment Contribution dimension covers MIGA's beyond-project level impact.

It is not clear how project outcomes and expected foreign investment outcomes contribute to development impact defined as MIGA's contribution to the SDGs and the WB Corporate Goals. It is plausible to assume that MIGA uses a similar logic to the one used by IFC. AIMM measures outcomes against objectives that are associated with the SDGs, helping IFC to track its impact in terms of contribution to the SDGs and the Corporate Goals (IFC AIMM brochure 2018). In a similar vein, IMPACT measures outcomes against expected results, which align to the SDGs.

The documentation reviewed reveal limited details on the selection and use of indicators. The tool measures both project-level outcomes and beyond-project level outcomes, i.e. the extent to which MIGA fulfils its objective "to encourage the flow of investments for productive purposes among member countries.

MIGA acknowledges that additionality is a core component of its operations and aims to support projects and stakeholders (i.e. risk classes, coverage amounts, tenors and member states) where few or no private insurers will provide coverage. Nonetheless, IMPACT does not seem to assess additionality of MIGA's projects.

Despite IMPACT assesses multiple dimensions of impact, i.e. financial, economic and social and environmental impacts, and systemic impacts (i.e. foreign investment contribution), no information is available to assess how IMPACT mediates trade-offs between these different dimensions.

IMPACT considers the risks that may limit the likelihood of development impact to realise. A risk factors matrix identifies relevant risks at sector, country, policy and political economy level. The likelihood of the project outcomes to realise, as well as the likelihood of the project contribution to encourage additional flow of investments is then assessed on a scale ranging from low to high.

The key difference between the IMPACT and AIMM lies in the way the institutions define beyond-project outcomes. MIGA's beyond-project level outcomes focus on the DFI's ability to encourage additional flow of investments, whereas IFC views beyond-project level outcomes as its contribution to market development.

3.1.3.3 Degree of Robustness

An assessment of the quality of the result framework underpinning MIGA's projects is not possible due to data limitations, as recent information on the methodology used by MIGA to assess the development impact of its projects is not available.

Drawing on the IEG assessment of the M&E system of MIGA and IFC (2013), MIGA uses the stakeholder analysis framework to assess the development impact of its projects, although it is not clear whether the methodology has changed since then. The stakeholder analysis is a common approach among private sector-oriented development finance institutions. The framework identifies the key actors affected by the project, as well as the magnitude of the project's impact on them. The assessment is completed by identifying counterfactuals: "What would the position of stakeholders be if the project did not exist?"

IMPACT uses findings from the stakeholder analysis to inform its assessment of expected project outcome and expected foreign investment contribution. It is not always clear how IMPACT processes these findings across its dimensions and related sub-dimensions, especially how private sector development effects feed into the assessment of the expected foreign investment contribution.

The IMPACT team formalises its assessment of ex-ante development impact through a final score that combines an assessment of the project outcomes and foreign investment contribution, as well as their likelihood of being achieved. No information is provided on the assessment process. Given the similarities between IMPACT and IFC's ex-ante assessment framework, we expect the IMPACT team to identify relevant components under each IMPACT dimension, determine the baseline stage of these components, and the magnitude of expected advancement. Based on these two parameters the team assigns a qualitative rating to each component, which are then combined to produce the project outcome and foreign investment contribution ratings. The ratings correspond to a qualitative scale including four categories: marginal, moderate, strong and very strong.

No information is available on the arrangements in place to quality assure the evidence generated by IMPACT, nor the extent to which investees and borrowers are involved in the ex-ante assessment process.

3.1.3.4 Use

IMPACT is used to assess and compare the expected development impact for each project and to inform project prioritisation across two review stages prior to the Board approval: Early Screening Meeting (ESM) and Project Review Committee (PRC). A note summarising the IMPACT ratings and providing evidence to support the assessment is produced and used to inform decision-making. The note describes the results of the IMPACT assessment along the three dimensions of expected project outcome (Business Performance, Economic Sustainability and Environmental and Social Performance), and relevant dimensions of expected foreign investment contribution. It includes an assessment of the likelihood of impact for both dimensions. Based on some extracts from this note, the content is mostly assertive in nature and provides limited explanation on how the impact is expected to realise, especially with regard to the project effects on the economy and society overall (Economic performance). Besides business performance indicators, indicators for economic performance and foreign investment effects are not always provided. The project documents present more prose on general development impact, although these have not been shared.

Findings from the IMPACT pilot phase indicate that the tool brings value to the process by steering decision-making towards high-impact projects. Three changes have been observed so far:

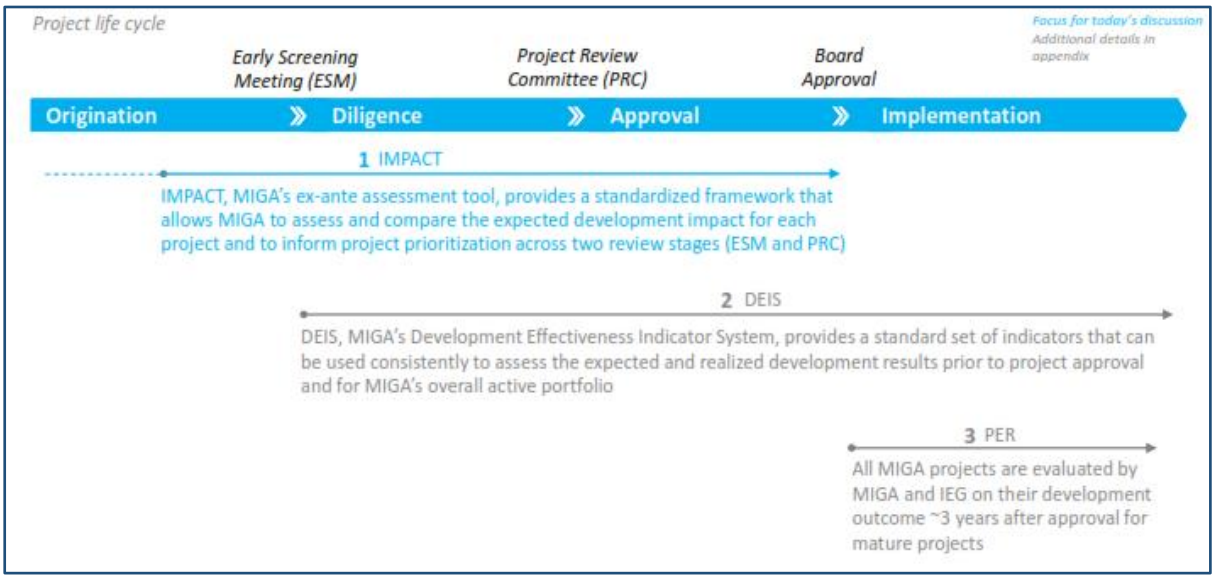
- At origination: projects with low expected IMPACT ratings have been deprioritised.
- At due diligence: IMPACT ratings inform opportunities to enhance development impact and structure projects accordingly. Level of engagement of clients has also increased.
- In approval phase: projects with low IMPACT ratings subsequently deprioritised from pipeline.

While the IMPACT rating does factor into the decision-making as reported by MIGA, it is not clear from the available documentation whether MIGA adopts a hard threshold for the IMPACT score to screen out investments. The tool brings value to the decision-making process by providing a case to prioritise high impact projects and steering decision-making towards projects in certain priority sectors or geographies for MIGA. The degree to which the IMPACT rating offers an adequate basis to inform decisions by the Board appears to be limited. The way the rating is portrayed in the project document is overly assertive, with little reference to indicators and how the impact is expected to realise. Nonetheless, our evidence base is thin, and it draws on an extract of an IMPACT note, which excludes some development impact considerations.

3.1.3.5 Integration

Since 2008, MIGA has invested in building a comprehensive results measurement system relying on largely similar frameworks, methodology and process used by IFC. Nonetheless, MIGA's results measurement system and M&E capacity is constrained by its business model as a political insurance provider. MIGA has more remote relationship to the project enterprise than IFC, which makes access to information not automatic. The system is made up of 3 pillars IMPACT, Development Effectiveness Indicator System (DEIS) and the Project Evaluation Report (PER), which provides an ex-post project-level assessment of development outcomes. The system has been designed to comprehensively assess development impact through the project cycle. See Figure 11.

Figure 11 - MIGA's results measurement system



Source: Executive Directors' Seminar - Progress against MIGA FY2020 Strategy and to review MIGA's Development Results Framework. April 29, 2019

Based on the documentation available IMPACT is used at ex-ante only and it is not clear whether IMPACT indicators are followed through by DEIS and reassessed ex-post. The Update on the review of MIGA development result framework in April last year reported that MIGA was evaluating the potential to evolve DEIS to incorporate features of the IMPACT framework. At the time of writing this report, no further information is available.

More broadly, insights from a semi-structured interview indicate that MIGA is moving toward a more integrated system although it is too early to evaluate any changes in accountability mechanisms and processes nurturing feedback loops between evaluation and learning.

3.1.4 AfDB

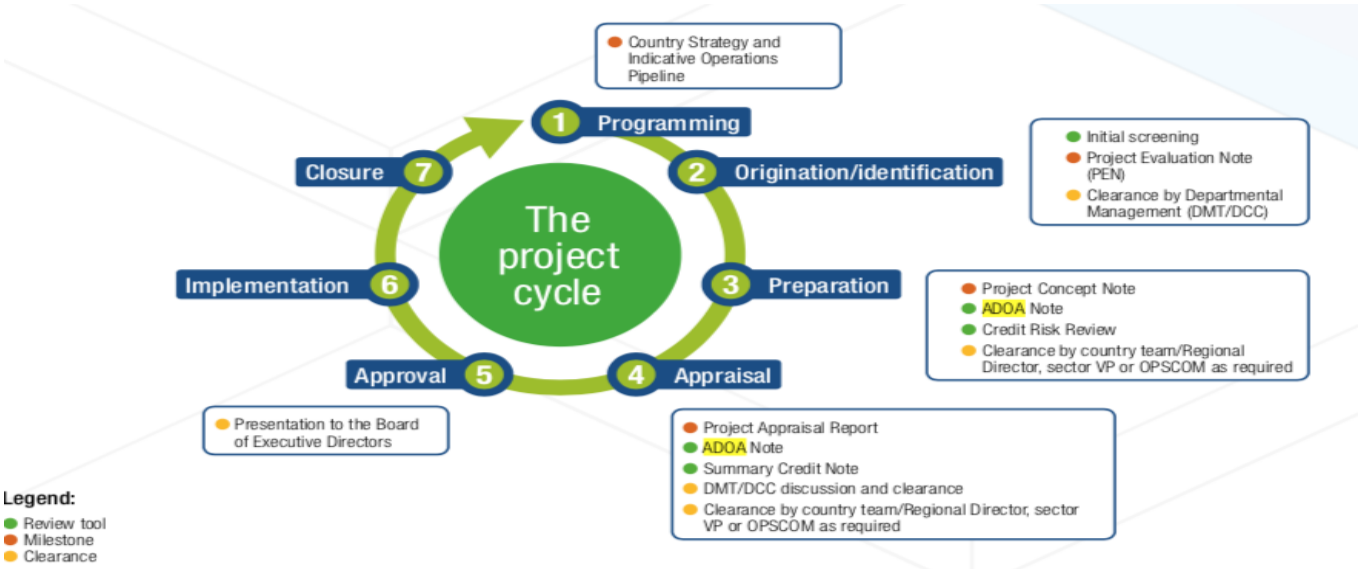
3.1.4.1 Introduction to the Tool

The African Development Bank's (hereafter, Bank) Additionality and Development Outcomes Assessment (ADOA) is an independent ex-ante decision-making tool based on an impact rating framework. All transactions entering the Non-Sovereign Operations (NSO) portfolio are subject to ADOA. Its purpose is to establish the added value and define the development effectiveness and measurement of NSOs at entry by providing answers to two questions: a) what do DFIs bring to the operation that commercial lenders cannot or do not bring?; and b) what are the expected development outcomes?

ADOA is an integral function in the NSO ecosystem that starts at Stage 3 (preparation) and completes at Stage 5 (approval) (Figure 12 below.) It is carried out by a dedicated Team in the ECVP who are accountable to and submit ADOA notes to the Board. The team provide what is best described as a challenge function to a Project Appraisal Team (PAT) made up of representation from the Credit, Social and Environmental Safeguard and Legal Units coordinated by a Project Officer and led by a Task Manager from the Private Sector Development Dept (PINS) (ADOA Operations Manual, 2016).

The PATs are established following approval of the Project Evaluation Note and start working together from Stage 3 in developing a Project Concept Note (PCN) and Stage 4 in developing a Project Appraisal Report (PAR). The final PAR is presented to the Board and includes a final ADOA note presented separately by an ADOA team member.

Figure 12 - The Bank's procedural framework for Quality at Entry – Non-Sovereign Operations



Source: Evaluation of the Quality at Entry of the African Development Bank Group's Sovereign and Non-Sovereign Operations (2013–2017) Summary Report, October 2018

3.1.4.2 Adequacy of Scope

While there is no explicit definition of development impact, the development outcomes are made explicit. They relate to the High 5's priorities - Light up and power Africa, Feed Africa, Industrialise Africa, Integrate Africa and Improve the quality of life for the people of Africa – reflected in Level 2 of the AfDB's 2016-2025 Results Measurement Framework - its contribution to Development Impact.

These Development Outcomes are best described as direct impacts with the exception of the last category, private sector development and demonstration effects (Figure 13)[15]. This is understandable: the positioning of the bank necessarily means the majority of its NSO operations target financial wholesalers such as regional funds and national financial institutions. Less understandable is they are heavily weighted towards macro and regional level factors with muted reference to end beneficiaries.

Figure 13 - Seven categories of development outcomes[16]

[15] Mention is made about expanding ADOA's Development Outcomes to include NSO's indirect impacts on jobs in the Bank's 2016-2025 Results Management Framework trialled by CDC (2016-2025 Results Management Framework and Semi-Structured Interview)

[16] These categories are also aligned with those developed by the Evaluation Cooperation Group.

CATEGORY	DESCRIPTION
Household Benefits and Job Creation	NSO's effects on the welfare of households such as job creation, introduction of new or higher quality products and improvements in the price/quality of goods and services.
Infrastructure	Improvements in the capacity and supply (e.g. power generation, increased capacity, etc.), as well as access to infrastructure services (number of connections to the grid,) due to the NSO.
Governance and Fiscal Effects	NSO's effects on the budget of concerned governments in RMCs including fiscal effects, public sector governance (e.g. compliance with investment codes) and accountability, and the fairness of concessions and public contracts awarded to private parties.
Regional Integration and Economic Resilience	NSO's contribution to regional integration (e.g. intra-regional trade or projects serving several countries within the region) and macroeconomic resilience such as effects on regional trade, external balance, forex generation and economic diversification.
Environmental Effects and Contribution to Green Growth	Captures all relevant environmental effects caused by a NSO (effects on local ecosystem) including effects on green growth (e.g. greenhouse gas emissions and sustainable use of non-renewable natural resources).
Gender and Social Effects	NSO's contribution to gender equity (e.g. employment, capacity building and ownership of means of production) and social inclusion (e.g. increased access to electricity for rural populations, financial inclusion for women and the poor, projects in fragile states, contribution to food security, social programs for local communities).
Private Sector Development and Demonstration effects	NSO's effects on private sector development (e.g. development of value chains) and the enabling environment for private sector to thrive. The category captures the alleviation of financial constraints and improved supply of business inputs as well as the spread of good corporate governance, managerial and technological know-how and local linkages and positive externalities.

Source: Policy on Non-Sovereign Operations, 2018

Drawn from the above categories, ADOA selects two deep core indicators:

- Numbers of Direct Full Time Equivalent Jobs created: a**)** Construction and b) operations and maintenance at project and, if applicable, at sub-project level broken down by Gender
- Payment to Government in Million USD NPV (Net Present Value)**:** by projects and, if applicable, by sub-projects

The reason to decide on these appears to contradict how ADOA places special focus on the Bank's role in mobilising co-financing that would not have been forthcoming in the absence of the Bank's and/or other DFIs participation in the operation and catalysing other investments in related sectors of the economy. They also appear to effectively preclude the possibility of any project, which impacts on low-income groups. The jobs indicator, although disaggregated by gender, is quiet on the qualitative dimensions. It does not specify what types of position among the investees' management and operations.

Its assessment of additionality provides answers across three dimensions, namely financial additionality, political risk mitigation, and improved development outcomes:

1. **Financial additionality** addresses the additionality brought by DFI financing by reducing commercial operators' exposure to credit, liquidity, or market risk, in ways that cannot be achieved using private sources and commercial players alone. Financial additionality depends on the overall reduction in commercial risk relative to the counterfactual scenario of no DFI participation

2. **Political risk mitigation** seeks to reduce the Private Sector Organisations' (i.e. the Financial Intermediaries) exposure to adverse government actions or political instability through financial and legal instruments given how such instruments are not available for many African countries. A dimension, unlike the other two, not included in the in the MDB's harmonised approach to assessing additionality. [17] and

3. **Improved Development Outcomes** (i.e. non-financial additionality) measures improvements in development outcomes, which can be attributed to DFIs' involvement in designing PSOs or putting in place necessary capacities, frameworks or systems. Such DFIs' involvement may positively affect the set of expected development

[17] https://www.adb.org/sites/default/files/institutional-document/456886/mdb-additionality-private-sector.pdf

outcomes, amplify their magnitude, or increase the likelihood that these expected development outcomes will be realised.

ADOA's categories of development outcomes and dimensions of additionality define a balanced scope of enquiry, including integration of gender and the environment. They also provide a consistent and clear basis for ADOA' independent assessment for projects in country and at regional level.

It is assumed that all projects assessed by ADOA meet the minimum requirements in terms of commercial viability as investigated by the Bank's credit risk department at the preparation stage. (ADOA Framework 2.0, 2017) That ADOA assumes commercial viability as a precondition could mean that projects with development impact are never seen if they fall just below the commercial criteria.

While ADOA evaluates the likelihood of achieving sustained development outcomes over time, such analysis lacks a framework and appears dis-connected from and so different to risks and assumptions set out in the results frameworks.

ADOA typically uses a seven-year time horizon for financing implementation covering the construction and operation phases for infrastructure, and goods and services companies under project finance. In the case of corporate loans or equity participation this is less clear-cut and there are a few exceptions.

3.1.4.3 Degree of Robustness

The Development Outcomes' indicators are harmonised, prioritised and broken down across sectors, the due diligence missions ensure some degree of scrutiny in the data presented in the financial appraisal on investee outputs and there is an internal quality mechanism in place. The 5 yearly independent evaluation of by the Bank's evaluation office seems light.

The ADOA team takes a lead role in the International Financial Institutions (IFIs) meetings to define Harmonised Indicators for Private Sector Operations (HIPSO). The ADOA team leads the infrastructure work stream. It has also played a key role in the MDB working group and indicators for all seven development outcomes are harmonised across the Bank.

> *"The main and most important strength of ADOA has been in establishing a harmonised set of indicators for the development outcomes" (SSI).*

The seven development outcomes are not presented within a clear and logical Theory of Change or results framework. There are results frameworks in the PAR, but ADOA does not assess the evaluability of these and they are not systematically informed by the ADOA analysis at screening.

ADOA's guiding principle on the counterfactual explain how the development effects attributed to the PSO are those incremental to the project vs. the "no-project" scenario. Yet the basis on how and how well this is done is not evident. The only exception to the tool's treatment of this is the *Environmental effects and contribution to green growth* category. Establishing a counterfactual is based on current practices and technologies specific to the sector. That for attribution is treated differently with a focus on avoiding double-counting the effects among the development outcomes.

There appears to be some inconsistency among the indicators ADOA sets for development outcomes and the consequences of the Banks efforts in improving these through the improved development outcome dimension of additionality with those set out in the project's results/logical frameworks in the PARs. This is acknowledged and appropriate adjustments are being made to the PCN and the PAR guidance by the Quality Assurance Team (SSI). Moreover, these results frameworks are not part of the agreement with borrowers. (SSI).

The overall rating for additionality is derived by taking the highest of the individual dimension ratings reflecting the principle that a DFI could enhance a PSO's design through financial additionality, political risk mitigation or improving upon its expected development effects, respectively. It is not always clear how well and robust the analysis necessarily leads to the assertion that the MDBs bring addition:

> *"……some board members consider this assessment to be too lenient" (SSI).*

Regarding client consultation, the ADOA officer does send a Due Diligence (DD) questionnaire to the Task Manager for onward transmission to the Project Sponsor ahead of the DD mission. However, the adequacy of client consultation is limited to internal aspects of the investee, does not include an adequate analysis of the investees' customers in terms of who they and the analysis it provides does not come through clearly enough in the ADOA notes (MOPAN 2016 and semi-structured interview). As noted in the NSO quality assurance plan (2019):

> *"there is a need to ensure more significant involvement of private sector investors in the process of defining relevant indicators and setting baselines, realistic targets and timelines in the view of greater ownership and higher probability of delivery."*

3.1.4.4 Use

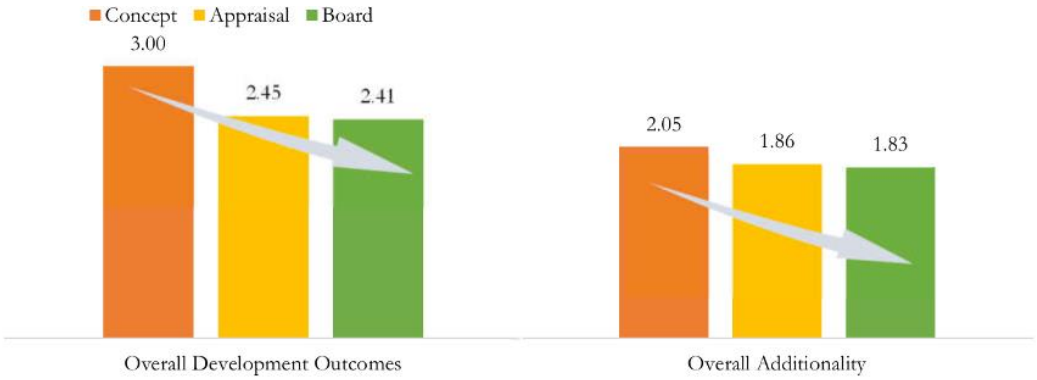
The use and usefulness of ADOA notes for the Operations Committee lie in helping provide a basis, through the rating and scoring system, with which to improve the project's quality at entry as it moves from concept note to appraisal stage. They also inform decisions on dropping projects at or before appraisal. Furthermore, the value of the ADOA team to the PAT provides it with discipline and assurance despite the inevitable tensions that sometimes arise.

> *"The Process of ADOA, through participating from the PCN to PAR stages, provides a useful support and challenge function to the PAT and so provides more credibility to the assessment" (SSI).*

At the early stage of the project cycle ADOA is one of the main pillars for decision- making. If a project's rating falls below a certain threshold, whether on additionality or development outcomes, the project is likely to be dropped. Its main use lies in providing an extra safety net ensuring that the bank is indeed funding projects according to its mandate and strives to mainstream expected value-added into its decision-making process.

As Figure 14 implies, the average overall development outcomes score goes from 3.00 at concept to 2.41 at Board stages (the lower the number, the better the rating). Similarly, the average overall additionality score moves from 2.05 at concept stage to 1.83 at Board stage. These improvements in ratings can be partially attributed to ADOA contributions through additional information on the project that is made available at the appraisal stage through the ADOA's due diligence tools, and the advisory role played by ADOA team.

Figure 14 - Average Rating Scores of Approved NSOs across stages



Source: ADOA 2019 Annual Report

There is no formal synthesis of ADOA's results portfolio level to aid steering decisions towards priority geographies and/or sectors. However, the Team produces an annual report that includes a useful tabulated summary of the specific NSOs assessed during the year. It breaks down the NSOs design quality as they move across the PCN and PAR stages though to approval. It is not clear whether this report is presented to OPSCOM or the Board.

The 2018 IDEV evaluation of quality at entry found a lack of incentives for NSOs to optimally assess and learn lessons from the Bank's contribution to private sector development. Once an acceptable rating of potential development impact has been obtained, there is little incentive to further enrich the development argument and thus better articulate a project's contribution to more strategic private sector development impacts. Evidence from a few ADOA notes attest to this: a rating of marginally satisfactory development outcomes is sufficient for board approval.

The ADOA team provides both internal and external capacity building to the NSO eco-system of regional and non-regional institutions. It delivered ten training sessions to build the capacity of the Bank staff and partner institutions. These included training of over 100 new investments officers from different department of the Bank, and six partner financial institutions. The team also produced knowledge products in the area of trade finance, risk participation agreement, and long-term financing in Africa.

The ADOA note is revised and reissued over the processing cycle as new information becomes available. Three to five ADOA notes are issued for each PSO that are presented to the Operations Committee at appraisal stage and to the Board for final approval. The final ADOA note includes a one-page summary description of ratings across the development outcomes and additionality in a standard template along with a one sentence description of the projects' strengths, weaknesses and likelihood assessment of outcomes. The remaining 4-5 pages sets out evidence, albeit with a limited analysis. Based on example ADOA notes and its contributions to the PAR, the analysis feels feint, is too assertive and plays down the rich data ADOA has collected. This limits its potential value.

> *"the ADOA notes undersell the value the team brings to the OPSCOM and the board" (SSI).*

The responsibility for fulfilling the purpose of ADOA lies with the Economic Governance and Knowledge Management Complex. The principal function of this complex is to devise and carry out research and analysis on priority social and economic development issues for the benefit of the Bank's member countries. The responsibility for ADOA does not appear to sit comfortably with this mandate and there is evidence of how this can distract if constrain its main purpose (SSIs).

Last year fifty-nine ADOA assessments are carried out by a team of 8 people. The level of effort in drafting and clearing ADOA notes at concept note, appraisal and approval stages is considerable: it can take up to 25 days/project from concept note up to and including submission of the final ADOA note to the board (ADOA Training Manual). It is noteworthy that the ADOA team's capacity has been constrained since late 2017: there has been a high turnover of staff (SSI). The reasons for this are not clear yet may have had some effect in reducing the value of ADOA (SSI).

3.1.4.5 Integration

ADOA's mandate stops at the approval stage, thus it is not integrated into subsequent stages of the project cycle.

The Private Sector Support Department (PINS) takes responsibility for taking this forward and trains its people to track progress against the indicators. However, no one is assigned to specifically do this, unlike other officers on supervision missions who monitor financial performance and adherence to legal agreements. As a result, PINS often loses sight of important information such as numbers of jobs being created and an analysis of this.

The issue of relative neglect of results mentioned earlier repeats itself during implementation and at exit.

> *"Supervision and completion reporting is excessively focused on administrative and fiduciary issues, with lesser attention to development results" (Evaluation of Quality Assurance across the Project Cycle of the African Development Bank Group (2012–2017) Synthesis Report, 2018).*

There is limited evidence on how ADOA play a part in broader stakeholder engagement. Project Appraisal reports and evaluations commissioned through the Bank's Independent Evaluation Function (BDEV)re publicly disclosed, but the ADOA framework and the notes produced are not.

As noted by BDEV's 2018 evaluation, Project Appraisal Reports (PARs) do not capture essential lessons learnt from past operations, either in the sector, the country or in similar operations of other partners as much as one would hope or expect. The most common source of lessons learned are previous projects or phases, while Independent Evaluations and Bank ESW are cited far less frequently (24% and 12% respectively for a sub-sample of 25 sovereign operations reviewed).

3.1.5 DEG

3.1.5.1 Introduction to the Tool

DEG uses the Development Effectiveness Rating (DERa) tool to measure and manage the development contributions of its investments. DERa was developed in 2016 replacing the former Corporate-Policy Project Rating (GPR) tool that had been in use at DEG since 2002. DERa has been in full use in DEG since the beginning of 2017.

DERa is an integral function in DEG's decision-making and impact management processes. It allows DEG to measure the anticipated development effects of all its investments across all sectors of operation at origination, as well as to monitor the achievement of these effects throughout the project lifecycle and their contributions to the Sustainable Development Goals (SDGs). The output of DERa is a development effectiveness rating score.

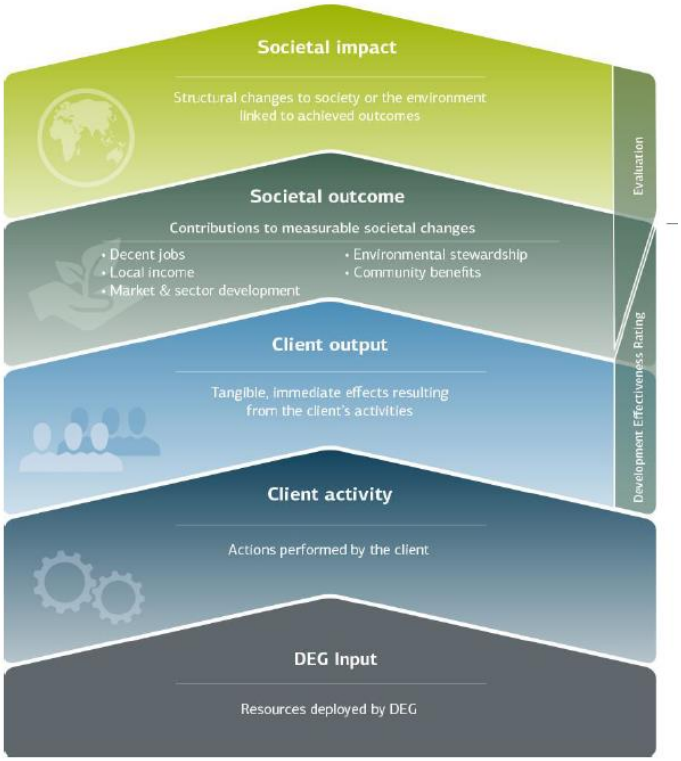
DERa is used by investment officers during the project appraisal and annual monitoring assessments, with support and oversight from the Corporate Strategy and Development Policy team within DEG that developed the tool. It relies on some outputs from other tools and

assessments separately conducted at DEG such as due diligence and environmental, social and governance assessments.

3.1.5.2 Adequacy of Scope

DEG's mandate is to promote sustainable growth and improved living conditions of the local population by investing in the private sector in developing and emerging economies. This mandate is underpinned by an overarching Theory of Change that specifies the outcomes and impacts that DEG wishes to achieve through its portfolio of investments (Figure 15).

Figure 15 - DEG's Theory of Change



Source: https://www.deginvest.de/International-financing/DEG/%C3%9Cber-uns/Was-wir-bewirken/Wir-messen-Wirksamkeit/

DERa uses five outcome categories to assess the development contributions of its investments. These are: decent jobs, local income, market and sector development, environmental stewardship, and community benefits. The majority of the outcomes measure direct effects of the interventions, with the exception of indirect job potential, which measures indirect beneficiary effects, and promotion of innovation and competitiveness in the market, which measures systemic market effects. Each of the outcomes (and their underlying indicators) are linked to the SDGs, allowing DEG to measure investments' contribution to the SDGs (Figure 16).

Figure 16 - DERa outcomes linked to SDGs

Source: DEG presentation on DERa, 2018

While the tool incorporates an assessment of indirect job potential and market development, these are inferred by direct project proxy outcomes (such as sector of the project, yes/no question on promoting innovation). However, the outcomes measured reflect economic, social and environmental dimensions and these are linked to SDGs. DERa does not differentiate between types of affected people/stakeholders at ex-ante – with the exception of the income level of the country (at macro level only).

Risk assessment is not incorporated into DERa. DEG separately conduct risk assessments (ESG) and include some 'do no harm' indicators in its ex-ante assessment, but they don't measure the likelihood of impact being achieved. Similarly, DEG does not measure additionality in DERa. It reports that it measures it separately, but the researchers did not have access to specific information.

3.1.5.3 Degree of Robustness

DERa measurement is standardised across projects, with some small differences for investments in funds or financial institutions. There is only a generic Theory of Change (TOC) at the institutional level (as in Figure 11). Development impact theses or TOCs are not defined at project or sector level and the same set of outcomes are measured for projects in all sectors.

The assessment takes into account the current level of all indicators (baseline) and a forecast for the indicators in 5 years' time. This allows a comparison of the increase in outcomes over the 5 year period (but this is not done with attribution in mind – rather where the clients seem themselves in 5 years, which is partly due to investment but also due to other trends).

A range of quantitative and qualitative indicators are used, and many are adopted from harmonised frameworks (HIPSO/IRIS+). The data requirements of DERa are not extensive and they are gathered as much as possible from existing sources such as financial and ESG assessments done by other teams, from international organisation data repositories, and from the client and judgement of investment officers. The final output are two single scores (one for baseline and one forecasted) that can be compared across projects as well as aggregated at portfolio and different levels (country, sector, etc).

There is a quality assurance mechanism in place: the DERa system checks for missing or inconsistent data and the impact team reviews all DERa assessments before they go to the investment committee for approval. An independent financial auditor reviewed the scores in 2018.

3.1.5.4 Use

DERa is an integral part of the decision-making. At investment level, it is seen by operation and impact teams as a key criterion for approval, even though there are no hard thresholds at project level. It allows DEG to adopt a portfolio management approach setting targets at portfolio level. The portfolio target is based on actual current DERa scores, which are the simple average of all current DERa scores for all projects in the portfolio. When a new project enters the investment pipeline, it is its baseline DERa score that will have an effect on the portfolio target for that year (given that the portfolio target is based on current DERa scores in a given year) and not its forecasted score.

Given its scoring mechanism, DERa provides a direct steer to invest in certain geographies and sectors. DERa rewards projects that are in lower income countries and in specific sectors, which enable private sector development by providing a direct uplift in the score if the project is in a lower income country or one of these enabling sectors.

The decision to approving investments is done by an Investment Committee. Depending on the size of the project, the management team or Board could be involved in approving, but the latter generally approves few projects.

In addition to its use in deciding whether to approve an investment, the DERa assessment is also used to identify where best to support clients during project implementation in order to enhance their contribution to development and their likelihood of achieving the intended outcomes. Support could include providing guidance and advisory services through DEG's Business Support Services, for instance on labour standards, supply chains, or training programmes.

DERa is very practical and efficient. It is fully integrated within DEG’s IT systems and allows for some data to be drawn automatically from other databases within the organisation. For investment officers who are familiar with the tool, it could take less than a few hours to input all the data and generate the project score once the client reports have been received. The Corporate Strategy and Development Policy team that is responsible for impact reporting and tool maintenance and data quality checks is a small team. Additionally, DERa has a low reporting burden on clients as almost all of its indicators already exist in other forms of reporting that the client has to do (such as tax payments and number of employees).

There is an internal DERa manual that describes the assessment process, as well as how each indicator is defined and rules for determining the baseline and targets. The research team did not review a copy of this manual to assess the level of its detail and quality in providing a consistent understanding of how to undertake the ex-ante assessment.

3.1.5.5 Integration

DERa is integrated into the monitoring cycle. All indicators defined ex-ante are tracked annually and the DERa score is recalculated and compared to the ex-ante baseline and forecasted scores. On the other side, it is less integrated with evaluation activities. Few evaluations are conducted per year and these are often thematic. Furthermore, there is no self-assessment at the end of the project (only a calculation of the final DERa score). No clear indication was provided on how it fits into overall RMS and if there are scorecards at corporate level for incentive management.

When it comes to stakeholder engagement and accountability on DEG’s management of development impact, there was not enough evidence. It is not clear if the ex-ante and annual

DERa scores are shared with the Board for all projects. The DERa scores for individual projects are not publicly disclosed, nor is there a narrative about the contribution of projects to development impact. DEG publish an annual development report that produces a summary of their development achievements each year; however, only portfolio level summaries are provided and for a selection of the indicators.

3.1.6 EBRD

3.1.6.1 Introduction to the Tool

EBRD's operations are governed by three key principles: sound banking, transition impact (TI) and additionality. Transition impact is centred on the market economy and is understood as the contribution to the process of transition towards a well-functioning market economy. New projects are assessed against each principle at appraisal stage, and this forms the basis for deciding whether to approve a project. The assessment is conducted independently for each principle using separate tools. The Economics, Policy and Governance (EPG) team, in collaboration with the Banking team[18], is responsible for assessing and managing the transition impact and additionality of projects, while the Banking team is responsible for assessing and managing the financial aspects of projects.

In 2016, EBRD introduced significant changes to its TI methodology and overall results architecture. This included, among others, a revised transition concept that underpins the Bank's operations followed by an updated transition assessment framework at country, corporate and project level that operationalises the revised transition concept. The updated assessment framework has been subject to constant refinements over time and is still in the pilot phase at the time of the undertaking of this study, with some elements expected to be further improved.

At the project level, EBRD adopts an end-to-end system to assess projects at origination and monitor them throughout their implementation period and at exit. The Transition Objectives Measurement System (TOMS), which was newly developed by the EPG team, is used to assess the expected transition impact of projects ex-ante. This is an automated and streamlined system that scores projects based on their expected contribution to promoting transition in the country of operation. TOMS also relies on the results from other assessments conducted at country-level including the Assessment of Transition Qualities (ATQs) and the Country Strategies (CS) that respectively define for each of the Bank's countries of operation the existing transition challenges as well as strategic priorities. The output of the TOMS is a transition impact rating score, *referred to as the Expected Transition Impact (ETI) score*, along with a TI narrative and a set of indicators to be monitored during project implementation. For monitoring projects during implementation, EBRD uses the Transition Impact Monitoring System (TIMS) which is closely aligned to TOMS.

The appraisal of each project takes place over three stages: Concept Review, Structure Review and Final Review, with the second stage optional if there are material changes after the first. The ex-ante assessment of potential transition impact starts at Concept Review using TOMS, and this is further enhanced and completed by the end of the Final Review stage. During the appraisal stage, due diligence and Environmental and Social risk assessments are also separately undertaken. At the end of the Final Review stage, the output from the project appraisal, including TI and additionality assessments, are summarised in a project document that is submitted to an Operations Committee for review and approval. Once approved, the project is submitted to the President for signature and then sent to the Board of Directors for final approval. After the Board approves a project, it transitions into the monitoring.

[18] The banking team are the equivalent of operational teams or investment officers – that is, the teams responsible for project identification, design and management.

This assessment framework, described in further detail below, applies to all of EBRD's investments and frameworks (which are a group of associated investments under an umbrella) in all sectors, but does not apply to their projects that only offer technical assistance, policy dialogue or business advisory services. Additionally, about half of projects that are assessed through TOMS are complemented by an in-depth manual TI assessment undertaken by EPG. These are typically projects that are deemed to be complex or have a high potential TI.

3.1.6.2 Adequacy of Scope

Unlike the other DFIs in this study, EBRD's TOMS is focused exclusively on transition impact - that is systemic changes in the market, which is in line with its mandate. As stated in Article I of the Agreement Establishing the Bank,

> *"the purpose of the Bank shall be to foster the transition towards open market-oriented economies and to promote private and entrepreneurial initiative" (EBRD, 1991).*

Transition impact is centred on the market economy and is understood as the contribution to the process of transition towards a well-functioning market economy. A well-functioning market economy in turn is defined as having the following six desirable qualities: competitive, well-governed, resilient, integrated, inclusive and green. TOMS measures a project's potential to contribute to transition impact along these six transition qualities. For each project, only two transition qualities can be selected in TOMS to be assessed: a primary quality and a secondary quality. These are the qualities that are deemed by the project team to be the primary and secondary drivers behind the transition impact of the project.

In addition to their contribution to transition, projects are also scored based on i) whether they target a transition quality that has a high gap in the country of operation (i.e. less developed markets), ii) their alignment with Country Strategy objectives, iii) their alignment with a selection of flagship projects in a given country , and iv) their use of strategic financing instruments (equity, local currency).

The potential for transition is the main part of the TOMS assessment and accounts for the largest share of the ETI score, with more weight given to the primary transition quality and less weight to the secondary quality.

The assessment of impact in TOMS is not based on the specification of a forecasted value for a specific set of anticipated outputs, outcomes or impacts such as job creation or poverty reduction. Rather, TOMS scores projects on their ambitions of impact. This is based on a recognition that the channels through which individual projects can achieve transition impacts cannot easily be measured through outcome measurement at the project level.

Thus, the IMP's dimension of 'HOW MUCH' – that is, the reach and depth of impact – is not assessed in a quantitative way but rather in the form of a narrative of what the project will achieve. However, the final and a key stage of the TOMS assessment involves defining a transition results framework whereby all claims of impact that the ETI score is based on have to be substantiated with indicators and associated targets for delivery. This involves selecting a set of output and outcome level indicators for each impact claim and setting a target that represents the scale of their expected change as a result of the project. These indicators are then defined as the monitoring indicators that are tracked through TIMS.

As for the duration of impact, all impact claims are assessed within the lifetime of the project. The Bank does not ask for commitments on impact claims that stretch beyond the closing of the project. There are however some outputs and outcomes that are expected to be achieved before the end of the project and the timeline for delivery would be set when defining the monitoring indicators.

The ex-ante TI assessment takes into account the probability of impact being achieved. The ETI score is risk-adjusted and reflects the full potential of expected impact discounted by the probability of delivery. However, at the moment this is not based on an actual assessment of the project-specific risks. Rather, for all projects, a blanket approach is applied whereby the ETI score is adjusted assuming high risks for project delivery. The ETI scores are therefore conservative. This is completed by a description of risks in free text format that is done by the Banking team. If there is clear information about a higher or lower risk, the ETI score may be modified manually. This was acknowledged in the interview with EPG members as being a challenge area for EBRD and a work in progress. Under the old TI methodology, ex-ante assessment was done manually by the EPG team who assessed both the potential TI and risks of a project. However, under the new system, there is a perceived trade-off between the need to have an automated TI assessment system that can be easily used by the Banking team, and an objective assessment of the risks to delivery (as Bankers cannot be asked to self-assess the risks). The EPG team is currently looking to enhance their risk assessment and are piloting methods in which this can be incorporated into TOMS.

As mentioned in the introduction to this section, additionality is separately assessed for each project and not incorporated into TOMS. Each project is evaluated along two dimensions of additionality: financial and/or non-financial. The former includes providing financing that is not available on reasonable terms and conditions in the local market, while the latter includes, among other things, the non-financial inputs that the Bank brings to an investment such as risk mitigation, support to adopt higher ESG standards, support in project design, capacity development and policy dialogue in country. Some of these non-financial aspects are typically included as conditionalities in the contract that the client has to deliver on (e.g. agreeing to adopt a specific ESG standard).

3.1.6.3 Degree of Robustness

TOMS allows for measurement that is partially project specific while at the same time grounded in systematic frameworks that allow comparability across projects. TOMS is based on underlying sector and transition quality level TOCs. Each project defines a narrative and is assessed on up to two transition qualities. Additionally, about half of projects also undergo a manual assessment by economists where scores could be adjusted based on the project specificities.

The assessment is based on answering multiple choice questions about the project characteristics, scope and objectives. These are overwhelmingly qualitative questions and rely mostly on questions about what the project is intending to do (so drawn from the investment team's understanding of the project). The questions are sector-transition quality specific. Based on the answers, TOMS automatically computes a score for each transition quality by averaging the individual scores of all answer-options within each quality that could be scored. The scores are then adjusted based on whether the project targets a transition quality that has a high gap in the country of operation and its alignment with country priorities and use of strategic financing instruments. The data for determining the gaps and country priorities relies on extensive assessments done in country. The indicators used to substantiate impact claims and monitoring are selected from a repository of 130+ indicators, some of which draw on harmonised frameworks (HIPSO/IRIS+). The selection of the indicators is done automatically by the system to avoid subjective selection and this is based on the answers to the multiple-choice questions, where each question-answer combination is associated with one or more pre-defined indicators. The final output is a single score that can be compared across projects as well as aggregated at portfolio and different levels (country, sector, etc).

There is currently no systematic framework for assessing additionality, but EBRD is piloting a tool for it at the moment after a review of the Bank's methodology carried out by the Independent Evaluation Department (EvD) in 2018 concluded that

> *"the operationalisation of additionality at project level is viewed by many as the least robust out of the three key operating principles of the Bank, and important aspects of rigour have been reduced." ... "The presentation of financial additionality (terms) commonly refers to the market conditions in fairly generic terms, providing little analysis/data on broader market conditions or relevant benchmarks" while "the presentation of non-financial additionality is often seen as overlapping with transition impact" (EBRD, 2018b).*

There is an internal quality assurance mechanism in place: at the moment economists work closely with investment teams to do the assessment but going forward they will reduce their heavy involvement and conduct spot checks and regular validations.

3.1.6.4 Use

TOMS is an integral part of the decision-making process at EBRD. At the investment level, it is seen by Management and Board as a key criterion for approval, even though there are no hard thresholds at project level. The tool on the other hand allows EBRD to adopt a portfolio management approach setting targets at portfolio level. While there are no hard thresholds on the TOMS score, a review of the distribution of TOMS scores in previous years shows that almost all projects approved have at least a 'Good' score. Management and Board are involved in the decision-making process and TOMS provides them with useful information on the project. The Board sees all projects but can delegate some projects within a framework that are under 25m euros and uncontentious. TOMS allows for steering investments in areas where transition gaps are widest (but not in a specific pre-defined list of countries or sectors).

There is internal guidance on how to use TOMS and economists work closely with the investment teams. On resource intensity it is mixed: while completing the assessment in TOMS might not be very resource intensive, about half of the projects undergo an additional manual assessment by the economist team and the tool does rely on country diagnostics and other assessments that are intensive.

3.1.6.5 Integration

TOMS is integrated into the monitoring cycle. All indicators defined ex-ante are tracked annually and the TOMS score is recalculated and compared to the ex-ante score. During monitoring, an assessment of the percentage of targets achieved is also done in order to see if projects are on track for delivery or not. All projects undergo a self-assessment at exit, and these are reviewed by the independent evaluation group and some are validated.

The accountability seems to be strong as the Board gets sight of the ex-ante scores as well as the annual performance score. The ex-ante scores and development impact narrative are also published online for the public. It is not clear if this is done consistently for all projects or if there are certain exceptions. The annual performance scores, however, are not published online.

TOMS is integrated into overall results management system. The results of TOMS and TIMS feed into country level and corporate level scorecards that have targets and these set incentives for staff to manage impact.

3.2 Comparator Group I

3.2.1 AsDB

As AsDB is undertaking an internal review of its appraisal system since late 2019 and the findings and the development of a tool comparable to that used by other DFIs are under discussion at management level, AsDB was not assessed.

3.2.2 AIIB

There is no ex-ante tool at AIIB. However, the AIIB's corporate strategy is planned for autumn 2020, the agenda for which will include discussing the design features of such a tool. As with the AsDB, it is hoped the results of this study will usefully inform this process.

3.2.3 EIB

No tool, equivalent to others we studied, could be found for the EIB. Rather, this section provides a descriptive summary of the guidance EIB provides for the appraisal of all Financial Instruments deployed under the European Structural and Investment Fund (ESIF).[19] The EU Member States who receive funding under the ESIF have a national body known as the Managing Authority (MA) which oversees the use of the available resources. MAs use ESIF allocations and place them in Financial Instruments through a Fund or a financial intermediary from which eligible projects can be financed.

In cooperation with the European Commission, the EIB Group developed this tool to help MAs conduct ex-ante assessments for implementing Financial Instruments. These Financial Instruments aim to transform EU resources under the European Structural and Investment Funds (ESIF) into financial products such as loans, guarantee, equity and other risk-bearing mechanisms. These are then used to support economically viable projects, which promote EU policy objectives. ESIF Financial Instruments can be developed for allocations from five EU funding sources:

- European Regional Development Fund (ERDF)
- Cohesion Fund (CF)
- European Agricultural Fund for Rural Development (EAFRD) and
- European Maritime and Fisheries Fund (EMFF)
- European Social Fund (ESF)

3.2.3.1 Introduction to the Tool

The European Investment Bank is the lending arm of the European Union. Its mandate is to support the economy, create jobs, promote equality and improve lives for EU citizens and for people in developing countries. The EIB Group has two parts: the European Investment Bank; and the European Investment Fund that specialises in finance for small businesses. Both parts lend to the public and private sectors and do not lend more than half of the cost of a project as the aim is for the loans to crowd in financing from private investors and other public financial institutions.

[19] https://ec.europa.eu/regional_policy/en/information/publications/guides/2014/ex-ante-assessment-methodology-for-financial-instruments-in-the-2014-2020-programming-period-general-methodology-covering-all-thematic-objectives-volume-i

The EIB group manages a range of Financial Instruments designed to contribute to reducing disparities between levels of development in the European regions and strengthen the economic, social and territorial cohesion of the EU.

> *"Union measures of financial support provided on a complementary basis from the budget in order to address one or more specific policy objectives of the Union. Such instruments may take the form of equity or quasi-equity investments, loans or guarantees, or other risk sharing instruments, and may, where appropriate, be combined with grants." (Article 2 (p) Regulation (EU, Euratom) no 966/2012 of 25 October 2012 on the financial rules applicable to the general budget of the Union).*

European Structural and Investment Fund (ESIF) Policies play a decisive role in reaching the objectives set up in the Europe 2020 strategy for a smart, sustainable and inclusive growth, while promoting harmonious development of the Union and reducing regional disparities.

The European Investment Bank designs Financial Instruments that support investments, which are expected to be financially viable but do not receive sufficient funding from market sources.

The objective of the ex-ante assessment is to provide evidence of the adequacy of the envisaged Financial Instrument against an identified market failure or suboptimal investment situation and to ensure the Financial Instrument will contribute to the achievement of the Programme and the ESIF objectives. In this way, it serves as a validation tool to check whether the decisions to deliver certain objectives laid down in the Programmes through an Financial Instrument are adequate. More specifically the assessment:

- Aims to ensure that ESI Funds allocated to Financial Instruments are fully aligned with the objectives of ESI Funds and Programmes and are used in accordance with the principles of sound financial management;
- Allows Managing Authorities (MA) to tackle high-priority market gaps; and
- Defines the specific priorities for the allocation of public resources.

3.2.3.2 Adequacy of Scope

There is no definition of development impact. The Financial Instruments' anticipated impact is explained by the extent to which they are consistent with the EU's 2020 strategic priorities for smart, sustainable and inclusive growth, aligned with ESIF's 11 Thematic Objectives and contribute to one or more of the relevant 38 investment priorities.[20],[21]

1. **Smart growth**: developing an economy based on knowledge and innovation - 75 % of the population aged 20-64 should be employed and 3% of the EU's GDP should be invested in R&D.

2. **Sustainable growth**: promoting a more resource efficient, greener and more competitive economy - The "20/20/20" climate/energy targets should be met (including an increase to 30% of emissions reduction if the conditions are right).

3. **Inclusive growth**: fostering a high-employment economy delivering social and territorial cohesion - The share of early school leavers should be under 10% and at least 40% of the younger generation should have a tertiary degree and 20 million less people should be at risk of poverty.

[20] See https://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf
[21] See https://keep.eu/faq/thematic-objectives-what-are-they-what-is-their-use/

There are seven elements that define the scope of the ex-ante assessment for Financial Instruments as defined by the CPR's requirements. These are broken down among two building blocks: A) Market Assessment covering elements 1-4 below; and B) Implementation and Delivery covering elements 5-7. The building blocks are intended to facilitate the development of robust ex-ante assessments.

Building Block 1: Market Assessment

1. Analysis of market failures and suboptimal investment situations through diagnosing a viability or financing gap to estimate the level and scope of public investment needs.
2. Assessment of the quantitative and qualitative value added of the financial instruments ensuring minimal market distortion and adherence to state aid rules (i.e. additionality).
3. Estimate of the additional public and private resources leveraged by the Financial Instrument, including co-financing.
4. Identification of lessons learnt from similar instruments and previous ex-ante assessments and integrated into the Financial Instrument's design.

Building Block 2: Implementation and Delivery

5. Proposed investment strategy, including an assessment of its possible combination with grant support, options for implementation arrangements and target groups that defines the thematic and geographical coverage, the relevance of the financial product to market needs and targeting of final recipients.
6. Expected quantitative results and outputs of the Financial Instrument that define its contributions to programme priorities.
7. Provisions allowing the ex-ante assessment to be reviewed.

An important feature of the tool is that it can be performed in stages. Should the results of first building block - the market assessment - lead to the conclusion that setting up an Financial Instrument is not justified, the Managing Authority (MA) could consider a different way to achieve the Programme objectives with other instruments. If the market assessment demonstrates the validity and the justification for establishing an Financial Instrument, the next step of the ex-ante assessment is to further develop the main characteristics of the Financial Instrument - the remaining elements 5-7 to facilitate its implementation. This includes mitigating possible risks (e.g. poor set-up, unsuccessful implementation and weak investment strategies in terms of financial products and volumes).

In this way, the ex-ante assessment is to be conceived as an iterative process rather than as a strictly linear one. This means that MAs will most likely go back and forth in their elaboration and will have to ensure the coherence of the whole assessment as described in elements 1-7 before finalisation.

3.2.2.3 Degree of Robustness

The expected results of the Financial Instrument are informed by previous parts of the ex-ante assessment: the market assessment, the expected value added and the investment strategy. This result orientation of all Financial Instruments is based on three pillars:

1. A clear articulation of the objectives of Programmes with a strong intervention logic (the result orientation of Programmes).

2. The definition of ex-ante conditionalities to ensure that the necessary prerequisites are in place for the effective and efficient use of Union support.

3. The establishment of clear and measurable milestones and targets to ensure progress is made as planned (performance framework).

What follows are findings relating to specific dimensions of the ex-ante assessment:

- The analysis of market failure (#1) and assessment of value addition (#2) carried out in the first building block; and

- The expected quantitative results and outputs of the Financial Instrument that define its contributions to programme priorities (#6) in the second building block.

These three areas are chosen for they are the most relevant to the sub-criteria under robustness and are comparable to evidence generated by other public funded DFIs selected for this study.

The analysis of Market Failure

The rationale for EU intervention, through Financial Instruments, is to support investments that are expected to be financially viable but are unable to raise sufficient funding on the market. This may be due to insufficient availability of funding (e.g. high risk of the sector or low profitability expectations) or due to the high costs associated with the available funding sources.

The results of the market failure and suboptimal investment analyses are a prerequisite for the identification of a need for support. In this sense, the ex-ante assessment has to provide an explicit statement on the identified investment gap that cannot be closed by market forces alone.

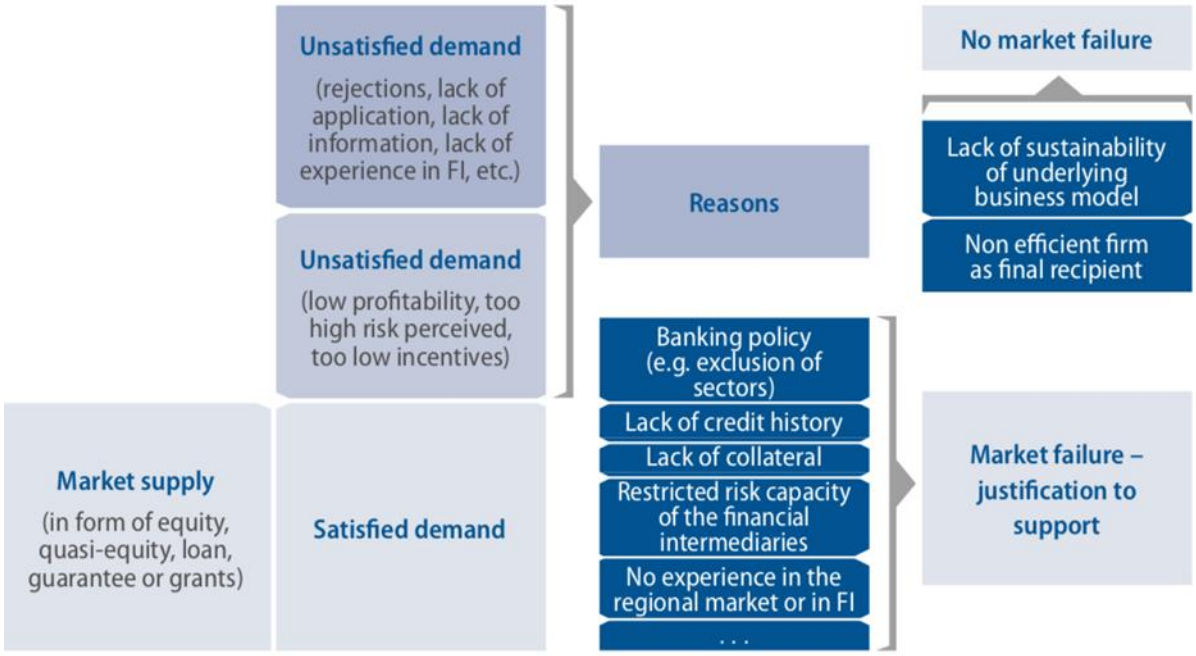
The analysis for the existence and, to the extent possible, the quantification of the market failure or the suboptimal investment situation allows determining the size of the investment gap to be filled by the Financial Instrument. This can result from the following:

1. A viability gap – in the case where the business plan of a project or of a group of projects demonstrates returns below market level. The viability gap is a cross-cutting issue, which tends to be independent from the financial structuring of the project. This can occur in sectors where project finance is the most common financial structure (e.g. energy, transport, urban development) but also where equity investment prevails (e.g. investment in SMEs and start-ups).

2. A financing gap – in the case where a certain sector or the economy as a whole shows evidence of unmet financing demand. The financing gap occurs especially for SME and mid-cap finance and in crises situations. Looking closer into the financing gap, it may be a gap for a certain financial product group like an equity gap for risk finance or a general lack of access to finance.

Following this analysis, the results of demand and supply analysis are collated that generate the quantification of the existing market failure and the investment gap to be covered by the envisaged Financial Instrument as shown in Figure 17 below.

Figure 17 - Calculation of unsatisfied demand and estimation of the level of market failure



The assessment of value addition for the Financial Instrument

After having identified the presence of market failure or suboptimal investment situations that justify public intervention and quantified the amount of support needed, the ex-ante assessment has to justify the value added of the envisaged Financial Instrument.

As a first step, the ex-ante assessment should analyse the quantitative dimension of the value added by the envisaged Financial Instrument. This analysis has to examine:

- The leverage of the EU (i.e. ESIF) contribution of additional contributions to the investment at all levels down to the final recipient. The higher the leverage achieved by the Financial Instrument the higher its value added.
- The intensity of subsidy of the Financial Instrument, which may be quantified in addition to the qualitative consideration (see below) of non-distorting the competition. The quantification helps to rank different options. The lower the intensity for a given project or group of projects the higher the value added.
- The revolving effect allowing the recycling of funds.
- Additional contributions coming from the final recipients, since these are excluded from the calculation of leverage.

The analysis of the value added implies comparing the envisaged Financial Instrument with other Financial Instruments, with grants or with other possible support mechanisms. Leverage represents one component of the quantitative value added and it assesses primarily the non-EU financial contributions by third parties during the first investment process.

After the quantitative dimension has been addressed, the ex-ante assessment identifies the qualitative value added of the envisaged Financial Instrument. Examples of qualitative categories of the value added include:

- Providing a financial product, which exactly matches the market gap without distorting the competition.

- Developing a new financial product type through the form of the envisaged Financial Instrument that has not been provided previously (e.g. microcredit).
- Supporting the building of or strengthening of the capacity of a sector, e.g. a nascent urban development fund sector.
- Giving preference to a Financial Instrument, which provides liquidity in the form of pre-financing of investment.
- Giving preference to a revolving long-term support scheme. This could be desirable for objectives such as seed support for SMEs, because the future generation of SMEs should also have the opportunity to be supported.
- Overcoming a specific market failure (e.g. lending capacity of the financial sector, which gives preference to a specific group of support schemes).
- Attract additional sources of expertise and know-how in delivering support to final recipients.

The expected quantitative results and outputs

While no logframe or ToC could be found, it is the responsibility of the Managing Authority (MA) to define two types of exclusively quantitative indicators.

1. Output indicators: the MA uses the set of common indicators already predetermined in the fund-specific Regulations or complementary documents provided by the Commission. These indicators cover the different forms of support to beneficiaries (including the reach of technical support and direct job creation) through Financial Instruments and include measures of operational efficiency or the performance (e.g. management costs and expected credit loss).
2. Results indicators: special attention to the definition of clear and measurable result indicators is critical and they must be clearly interpretable, statistically validated, truly responsive and directly linked to the specific objectives of the investment priority or focus area to which the Financial Instrument contributes. For that, the implementation of the Financial Instrument should affect the value of the selected result indicator under the corresponding investment priority or focus area. Result indicators across the different thematic guidance fiches are provided for the MA by the Commission.

Figure 18 - Example of indicators for a loan fund to SMEs

Type of FI	Loan funds to SME
Source of financing	ERDF
Funds budget	10M EUR
Specific objective of corresponding investment priority/ focus area	Supporting investments of SME with a focus to manufacturing

Result indicator of the corresponding investment priority/ focus area

indicator	unit	baseline	target	source
gross fixed capital formation in manufacturing	M EUR / national currency	800 (2011)	1.000 (2022)	National / regional accounts

Output indicators (as stated in the Annex of ERDF Regulation)

indicator	unit	baseline	target	source
enterprises receiving loans	number	-	100	monitoring
new enterprises supported	number	-	50	monitoring
Private investment matching public support to SMEs	M EUR / national currency	-	40	monitoring
jobs created in assist-ed SMEs	number	-	600	monitoring

Performance indicators

indicator	unit	target	source
credit loss	Volume of defaulted loans / volume of total loans outstanding	4 %	monitoring
Management costs	% on volume of total loans outstanding	4 %	monitoring
Leverage	Private investment matching public support to SMEs / public support to SMEs	5	monitoring

3.2.3.4 Use

No information available.

3.2.3.5 Integration

Monitoring and reporting requirements from the implementing body (e.g. a dedicated entity or entrusted financial intermediary) to the MA should be clearly defined in the funding agreement. A dedicated monitoring process should be defined at Financial Instrument level. The key elements of the monitoring process are illustrated below.

Figure 19 - Key elements of the monitoring process

The main purpose of the monitoring arrangement is about the steering of the Financial Instrument. The MA may pilot the Financial Instrument to some extent through conditions in the funding agreement about targeted results, leverage, reutilisation of resources and the responses of the Financial Instrument, when things develop differently, and deviations occur.

The MA could, therefore, decide to set up a monitoring and reporting system that provides them with information on the performance of the Financial Instruments in shorter intervals, e.g. with quarterly monitoring reports. A closer monitoring would allow the MA to identify possible constraints and issues that affect Financial Instrument implementation and to facilitate its management.

3.3 Comparator Group II

This section sets out discrete examples of good practice in assessing investments' anticipated impact as emerging from impact investors in the private sector. It has been informed by extensive desk research and two private sector case studies led by OPM. The analysis draws on three key sources:

- OPM's previous work on assessing impact due diligence systems in the private sector (OPM 2019)
- OPM's case studies on impact due diligence practices at Leapfrog and Actis
- The work of the Impact Due Diligence Initiative led by Pacific Community Venture (PCV), in partnership with the Impact Management Project (IMP) and with research support from the Global Impact Investing Network (GIIN) (PCV 2019).

In 2019, PCV reviewed due diligence practices of nearly 50 organisations and identified seven emerging best practices:

1. Assessing Impact Using the Impact Management Project's Five Dimensions, which are a widely accepted set of norms among Impact Investors.

2. Bridging the Divide between Environmental, Social, and Governance (ESG) and Impact Assessments to ensure all impacts an enterprise generate matter to people and planet including those related to business operations (i.e. ESG considerations) as well as products or services.

3. Aligning with the Sustainable Development Goals to help ensure anticipated impacts align with the global development agenda, as well as enable effective communication about expected impact across a diverse portfolio.

4. Elevating the Perspectives of Key Stakeholder by incorporating the perspectives of those who are impacted by investees helps investors amplify stakeholder voices, develop feedback loops between investors and investees, and assess both investor and investee contribution.

5. Evaluating a Commitment to Impact and Learning to help Investors ability to improve, adapt, and learn with a clear impact thesis informed key stakeholders' needs and linking financial incentives to impact performance.

6. Adopting a Portfolio-Wide Approach to help develop a consistent impact due diligence approach that enables direct comparisons of different types of investments across a portfolio.

7. Prioritising Accessibility to help ensure due diligence approaches can be easily adopted, use consistent language and are not burdensome to facilitate adoption.

In addition to these seven areas, our research points out other five emerging practices, which may complement the list above:

8. Integrating financial and social measures into one collective view of firm value.

9. Supplementing the impact score with a clear narrative by way of a Theory of Change/impact thesis, which provides evidence to support the rationale and explain intended development impacts of the investment.

10. Using due diligence tools to move beyond the debate over trade-offs between financial and development impact and inform decisions, which pursue development impact, while allowing for a broad range of viable investment profiles.

11. Integrating impact and financial management to enable investors to optimise investment performance across all dimensions and to communicate all dimensions of their investments' performance clearly and transparently.

12. Designing a simple and integrated system that supports the complete investment cycle to enable investors to clearly monitor investment performance against anticipated impacts and targets, as well as substantiate and validate their impact cases.

13. Committing to greater transparency through an independent assessment of internal due diligence that enable investors to demonstrate adherence to impact investing principles to enhance external accountability.

The remainder of this section breaks down good practices among the four leading criteria in our assessment framework – scope, robustness, use and integration – and their relevant sub-criteria. For each lead criterion, we locate relevant areas of emerging best practice and provide examples of applications by leading impact investors. The implications of these findings for target DFIs are discussed in the conclusion section.

This section presents evidence from the following institutions: Root Capital, Actis, HIP investor, LeapFrog, Acumen, Bridges Ventures and Omidyar Network. Table 2 provides a brief description of these entities. **Annex IV** presents the findings from Actis and LeapFrog case studies.

Table 2 - Description of entities

Institution	Mission statement
Root Capital	Root Capital is a non-profit social investment fund financing agricultural enterprise that support smallholder farmers across Africa, Latin America, and Southeast Asia.
Actis	Actis is a global platform offering a multi-asset strategy through the asset classes of private equity, energy, infrastructure, and real estate. Actis is a leading investor in growth markets across Africa, Asia and Latin America.
HIP investor	HIP investor is an investment advisor using a unique methodology to track, rate and rank investments' quantifiable impact on society.
LeapFrog	LeapFrog Investments is a profit-with-purpose investor. By backing high-growth, innovative, scalable businesses in Africa and Asia, the company seeks to fulfil the global unmet demand of billions of low-income, emerging-market consumers for critical services.
Acumen	Acumen is venture capital fund investing "Patient Capital", capital that bridges the gap between the efficiency and scale of market-based approaches and the social impact of pure philanthropy, in entrepreneurs bringing sustainable solutions to big problems of poverty.
Bridges Ventures	Bridges is a private fund manager that invests in solutions to pressing social and environmental challenges.
Omidyar Network	Omidyar Network is a philanthropic investment firm that aims to catalyse social impact on a large scale. The network works in multiple geographies, funding both commercial businesses and nonprofit organisations, focusing on investing in five sectors: education, emerging technology, governance and citizen engagement, financial inclusion, and property rights.
Sources: Root capital https://rootcapital.org/; Actis https://www.act.is/; HIP investor https://hipinvestor.com/; LeapFrog https://leapfroginvest.com/; Acumen https://acumen.org/about/;	

Institution	Mission statement
Bridges Ventures https://www.bridgesfundmanagement.com/; Omidyar Network https://www.omidyar.com/	

3.3.1 Adequacy of Scope

What is the scope of the development impact measured by the ex-ante tool?

Table 3 - Comparator Group II Adequacy of Scope

Good practice	Sub criteria	Application
Assessing impact using the 5 IMP dimensions	1.1 Impact definition - Who and What 1.3 Quantity of impact – How Much 1.4 Contribution 1.6 Risks and assumption underpinning the realisation of impact	• Root Capital • Actis
Alignment with SDGs	1.1 Impact definition 1.2 Alignment with the mandate of the institution	• HIP investor • Actis
Integrating financial and social measures into one collective view of firm value	1.5 Trade-offs between different dimensions of impact	• Leapfrog
Notes: We have not identified any emerging good practises related to sub criterion 1.7 - Horizon of realisation of impact and intertemporal distribution of impact. Actis and LeapFrog determine this by financial and direct impacts, not the indirect development impact return being realised recognising the latter will emerge after exiting the investment.		

The IMP's five dimensions of impact help investors to define and understand their development impact. By using the IMP's five dimensions investors will ensure that they are defining development impact consistently and comprehensively and robustly assessing expected impacts on people and the planet (PCV 2019).

Actis integrated its existing ESG framework with a bespoke sliding scale explicitly aligned to the IMP's five dimensions. For each investment, Actis identifies up to five material impacts aligned to specific SDGs (WHAT); for each impact Actis specifies who are the stakeholders that experience the positive social and environmental outcomes, in terms of the number of people benefitting and how well served they were already (WHO) [22]; the HOW MUCH dimension is measured by determining whether the investment is likely to achieve both deep and enduring positive change or something more short-term. CONTRIBUTION is assessed qualitatively, and it is understood as What Actis does differently and brings to the investee beyond providing finance. Actis assesses the RISK of the investment in failing to achieve its intended impact and scores this as Low, Moderate or High. While the result does not affect the impact score, it does help with investment decision-making by revealing the specific risks and how to mitigate these.

In a similar vein, Root Capital measures two types of impact, the Enterprise Impact and the Investment Impact, aligning these types to the IMP's 5 dimensions. The enterprise impact is the impact the enterprise has on its customers, suppliers and the environment and refer to the

[22] Actis is planning to do more work on better understanding the specific client base of their investee companies and the extent to which they are under-served. See Annex VI for more detail on how access assess the WHO dimension of Impact.

positive or negative social and environmental changes associated with the enterprise impact (the WHAT); their significance in terms of scale and depth (the HOW MUCH); and who experiences the changes and how underserved they are in relation to the outcome (the WHO). The Investment impact is the impact of the loan on the enterprise (the CONTRIBUTION dimension of impact), that is understood as the extent to which the investment adds value to a counterfactual scenario where Root Capital does not invest.

Incorporating the SDGs into impact due diligence can help investors to frame their impact as direct and indirect contribution to achieving the SDGs, aligning their operations with the global development agenda. This is particularly relevant for those organisations whose mandates overlap with the global goals (PCV 2019). [23]

Actis interprets impact as defined by select SDG goals relevant to particular investments and identifies up to five material impacts aligned to specific SDGs for each of its investment. HIP investor maps its impact metrics to the SDGs and looks for alignment between prospective investments' anticipated impacts and the SDGs. Each investment is assessed based on its performance against the specific impact metrics and its alignment with the SDG's indicators (full alignment, partial alignment and zero alignment).

Integrating financial and social measures into one collective view of firm value would help reducing the tension between development impact and commercial performance.

All LeapFrog funds have defined dual targets: top-quartile returns (profit) and emerging consumers reached with essential products or services (purpose). These are distilled to the level of each investee company, providing them with a clear measure of success.

FIIRM incorporates measurement of financial and operational Key Performance Indicators as well as governance indices, which are benchmarked to global best practice standards defined by IFC's impact principles. This enables LeapFrog's investment teams and portfolio company CEOs and CIOs to measure and drive performance towards both profitability and impact objectives.

Leapfrog sets impact measurements (KPIs) that integrate financial and social performance, and not trade one for the other. By doing so, both dimensions are seen as intrinsic contributors to building successful business models, entailing no necessary trade-off between social impact and financial return. These KPIs align financial and social performance and are integrated into existing management reporting structures. Just like operational KPIs, profit with purpose KPIs need to be tailored by industry and sometimes at the company level.

There is a strong business case for creating social value: For instance, in the insurance industry, products to a vast new emerging consumer segment represents a significant commercial opportunity (Profit) while enabling low income consumers to better mitigate risk and make better investment decisions for their future (Purpose). Leading business indicators go hand in hand with social outputs and outcomes as illustrated by, for example:

- Claims processing - the efficiency of the company in how quickly the company can get cash in the hands of the insured when they need it most; and
- Policy renewal – a leading indicator of profitability where customer retention is critical to small premium policies and an indicator of the value of the product to the low-income consumer

[23] It is important to mention that the SDGs offer little for aligning the objectives among financial service providers. There are no objectives and so targets relating to improving access to savings accounts, loans, insurance and other financial services. https://www.cgap.org/blog/financial-inclusion-has-big-role-play-reaching-sdgs

In addition, IRIS metrics provide a good means to identify indicators of social outputs for a range of businesses across many sectors. It is essential these metrics are aligned with the business' commercial objective. This linkage is at the very heart of integrated reporting, a critical tool for profit with purpose decisions.

3.3.2 Degree of Robustness

What evidence does the tool rely on and how is this gathered, analysed and synthesised?

Table 4 – Comparator Group II Degree of Robustness

Good practice	Sub criteria	Application
Elevating perspective of key stakeholder & Evaluating a commitment to impact and learning	2.1 Clarity of Theory of Change and result framework 2.3 Access to data from investees and stakeholders 2.8 Adequate consultation with the investee	• Acumen • Leapfrog
Adopting a portfolio wide approach	2.2 Range of indicators used 2.5 Use of harmonised indicators 2.6 Approach to measure impact across projects, sectors and geographies	• Actis • LeapFrog • Root Capital • Bridges Ventures
Notes: We have not identified any emerging good practises related to sub criterion 2.4 – Methodology to assess additionality and 2.7 – Quality assurance of the ex-ante impact assessment. Proving additionality is unique to DFIs, given that their mandate is to add value while not distorting/crowding out private impact investors. There is evidence of quality assurance practices from Leapfrog and Actis case studies. Quality assurance is provided by internal committees or independent third parties, similarly to how the DFIs' tools are checked by their internal panels and independent evaluation offices.		

Effective engagement of key stakeholders, including the investees and those who are impacted by the project investors are funding, can improve impact measurement and management in four ways. Consultation practices may help to mitigate and understand risks; understand and improve the system of incentives of the investees to ensure alignment to the impact goal; spot opportunities for innovation and improvement of products and services; assess performance of the project in terms of expected impact on ultimate beneficiaries and identify drivers behind outcomes, to be able to better inform assessment of contribution (PCV 2019).

Acumen systematically consults stakeholders beyond the investees, to validate their Theory of Change and understand the anticipated impact of the product or service and potential impact risks prior to investment. It does this by researching the customer base of its potential investee companies through a "quick and dirty application of its Lean Data approach: who are they, what choices they have in relation to the products and services the company sells and what is their poverty profile? By collecting such information through its Lean Data approach, Acumen generates in-depth characteristics of ultimate beneficiaries', including their poverty level, to better understand the enterprises' degree of poverty outreach.

LeapFrog's approach to impact is built upon a clear Theory of Change: by investing capital and expertise (inputs) in innovative companies, LeapFrog aims to equip emerging consumers with essential tools (outputs) that enable better risk mitigation, enhancement of financial and health well-being (outcomes), and that ultimately empower the customer to take entrepreneurial leaps out of poverty as a result of different life choices (impact). As part of the due diligence process, Leapfrog supplement data collected on the investee's internal performance by Consumer Insights. The diligence for impact risks and opportunities is

conducted using a standardised Impact and ESG Due Diligence toolkit, in collaboration with a consumer centricity diligence. The Consumer Insights team gleans feedback from customers in-store, by telephone and online, as well as extensive emerging consumer research data sets. It provides insights into customer profiles, needs, behaviour, and impact experienced. In addition, field interviews capture the experience of low-income consumers who are the target beneficiaries, and this generates learnings about consumers' diverse needs and preferences. This is similar to Acumen's use of the Lean Data approach at ex-ante and ex-post.

Alongside the engagement of key stakeholders, due diligence practices should assess the investee's commitment to impact, its capacity to learn and adapt and ability to achieve impact. Investors should assess the investee's impact objectives, incentive structure and organisational responsiveness, as well as impact model and robustness of the Impact Measurement and Management (IMM) to track and evaluate impact. This would facilitate alignment on impact goals and measurement and management of impact. It would also mitigate impact risks and identify opportunities to refine investors and investees' strategies to increase impact (PCV 2019).

Both Acumen and LeapFrog emphasise the importance of assessing the investee's knowledge and understanding of its target beneficiaries and its target impacts, including any negative outcomes related to their activities, when deploying capital. Acumen uses the Lean Data Approach to systematically look at these dimensions at due diligence. LeapFrog works closely with its partner companies to deepen their understanding of consumer needs. Through human-centered design processes, LeapFrog helps them deliver relevant, affordable and quality products. In addition, in the event of impact underperformance or other negative effects, LeapFrog enacts Environmental and Social Action Plans (ESAPs) or interventions mandated by LeapFrog's investment governing forums as applicable, which could include the need to update the impact targets in light of the performance of the investee.

PCV recommends that adopting a portfolio wide approach based on sector-agnostic frameworks, generalisable questions and indicators would help to make due diligence process more easily implementable and outputs of the assessment more comparable. A consistent approach would enable different sectors to adopt a common language and enhance comparability. Notwithstanding, frameworks need to allow for a certain degree of flexibility and specificity for different contexts to ensure a certain level of rigor in the impact assessment (PCV 2019).

Findings from our review reveal that investors commonly used a consistent framework to assess development impact ex-ante (Actis, LeapFrog, Vital Capital, Root Capital, Bridges Ventures). Some DFIs adopt this approach already. For example, Proparco assesses four main dimensions for each investment: [24]

- Expected impact on development: measured as the estimated development outputs and outcomes of the investment;
- Profitability: measured as the commercial viability of the investment and established by the financial department at Proparco;
- Level of risk: measured as a rating of the risks that the investment could face and established by the Risk department at Proparco; and
- Fit with Proparco's strategy: measured as the adequacy of the investment's fit with Proparco's investment and development strategies.

[24] Proparco is a subsidiary of Agence Française de Développement (AFD) focused on private sector development. https://www.proparco.fr/fr

Proparco's assessment of development impact always relies on a common set of four output/ outcome categories and four cross-cutting effects that are measured for all projects. At the same time, the tool allowed to select a set of outputs/outcomes that are specific to the type of client to ensure the tool appropriately captures potential impact areas of the investment.

Standardised impact metrics have been developed to help investors measure their impact in a consistent way. For instance, in 2019 the GIIN launched IRIS+, a generally accepted system for impact measurement, which identifies Core Metrics Sets of performance indicators by impact theme or category. The system is aligned to the SDGs, the IMP's five dimensions, and more than 50 other frameworks and conventions. IRIS Catalog of Metrics and the recent IRIS + Core Metrics Sets are two of the most used set of indicators adopted by impact investors (GIIN 2019).

3.3.3 Use

How is the tool used by project teams and decision-makers to steer investment decisions and what resources does it rely on?

Table 5 - Comparator Group II Use

Good practice	Sub criteria	Application
Supplementing impact score with clear supportive narrative	3.1 How is data synthesised and interpreted 3.3 Use of output across sectors 3.4 Use of output across geographies	• Actis • CDC
Using due diligence tools to inform investment across a range of possible returns	3.2 Use of the tool for go/no go decisions	• Omidyar Network • Root Capital
Prioritising accessibility	3.5 Guidance on using the tool 3.7 Capacity of M&E function 3.8 Level of effort needed to use the tool	• Root capital
Notes: We have not identified any emerging good practises related to sub criterion 3.6 - How the output from the tool is presented to decision-makers.		

Impact scores might not be the answer to everything. While indices are useful to classify or rank investments based on a comparative assessment of their impact potential, sometimes they do not allow for easy interpretation of the actual expected impacts. **Some elements of impact might be better explained through a narrative.** The narrative of impact would complement the impact score and provide evidence to support the rationale and intended development impacts of the investment. This would improve the basis to inform decisions and might help to steer investments in certain priority sectors or countries for the institution.

Actis's assessment of anticipated impact always includes a narrative capturing elements, which are hard to quantify and synthesise in a score. Some DFIs use a similar approach. For instance, CDC Group plc, a development finance institution owned by the UK government complement its impact score by a Development Impact (DI) thesis and DI case.[25] The DI thesis is a concise and clear statement that articulates the rationale for the investment and what its intended achievements are. The DI case is the supporting argument for the thesis that provides evidence to support the rationale and intended development impacts of the investment. The DI case aspires to be credible, compelling and concise. To establish the DI thesis and case, the investment team gathers evidence from several sources, including a

[25] https://www.cdcgroup.com/en/

review of: the fit/relevance of the investment with the team's/sector strategy; the recent relevant literature from other DFIs and academia; the outcomes of past comparable investments; and, in discussions with the client, conducting due diligence that includes ESG assessments.

In order to act on their impact mandate, investors should consider using due diligence tools to inform a broad range of viable investment profiles, adjusting their return expectations to their expectations of the investment development impact. For this reason, investors need to make sure that their due diligence tools provide a clear view of impact expectations when assessing specific opportunities (Bannick et al 2016).

Private sector investors, as well as DFIs, are developing models, which evaluate investment options based on the interplay between development impact and financial returns. These models vary widely. For example, Acumen pursues development impact by assessing the extent to which Acumen's investees reach under-served beneficiaries (i.e. poverty outreach) and use the assessment as a key determinant of a go/no go decision to invest, alongside financial returns. The Omidyar Network uses a framework to help making decisions in favour of investments that maximise development impact while maintaining financial sustainability. This philanthropic investment firm created a "returns continuum" framework, which lays out the necessary investment conditions given an expected development impact. Investments are assigned to a category on the return's continuum depending on the combination of expected financial returns and expected market impact. The lower are the financial returns, the more compelling the expected market impact needs to be. The "return continuum framework" allows the firm to consider sub commercial investments and grants alongside commercial investments, depending on the expected market impact (Bannick et al 2016). Similarly, Root Capital uses 'The Efficient Impact Frontier' to set goals at the portfolio level for how much they can improve the impact and financial performance of their portfolio as a whole. This is done by charting an efficient impact frontier, where a portfolio that lies on that frontier provides the highest level of impact expected relative to the expected return of investments. See section on integration for further details on the integration between impact and financial management.

In developing due diligence approaches, investors should prioritise accessibility. Due diligence practices and processes should be easy to adopt and not burdensome on users, and more readily incorporated into the existing investment processes. Incentives and methods need to be in place for easy engagement of internal and external parties on impact due diligence given their expectations and responsibilities (PCV 2019).

Root Capital's 'Expected Impact Rating is a good example of an accessible approach to due diligence. Root mainly rely on information already collected during financial due diligence, reducing data collection efforts. Loan officers are encouraged to build a relationship with the borrowers to facilitate communication and exchange of information. Borrowers complete the assessment themselves, with loan officers supporting the borrowers and entering only some information.

3.3.4 Integration

How is the tool integrated into the project lifecycle and more broadly into the oversight and accountability mechanisms of the institutions?

Table 6 - Comparator Group II Integration

Good practice	Sub criteria	Application
Bridging divide between ESG and impact assessment &	4.4 Integration of impact assessment and other aspects of due diligence within RMS	• Bridges Ventures • LeapFrog

Good practice	Sub criteria	Application
Enhancing integration of impact and financial management	4.2 Contribution of the tool to management oversight and accountability	
Designing an integrated system that support the complete investment cycle	4.1 Ex-ante assessment results followed through in monitoring and evaluating investment implementation	• Acumen • Actis • LeapFrog
Committing to a greater due diligence process transparency	4.3 Use of the tool for stakeholder engagement and external accountability	• Actis • LeapFrog
Notes: We have not identified any emerging good practises related to sub criterion 4.5 – Mechanisms to ensure learning and consequences in cases where planned objectives have not been met. The LeapFrog's case study reveals interesting findings, although it is not clear the extent to which similar practices are used by other investors. In the event of impact underperformance or other negative effects, LeapFrog enacts Environmental and Social Action Plans (ESAPs) or interventions mandated by LeapFrog's investment governing forums as applicable, which could include the need to update the impact targets in light of the performance.		

While often thought of as separate areas of practice, **investors should incorporate an examination of ESG factors within their assessments of expected impact** to systematically consider their investees' business practices and evaluate all impacts, including potential negative ones (PCV 2019). Additionally, **integration of impact and financial management would enable investors to optimise investment performance across all dimensions** and to communicate all dimensions of their investments' performance clearly and transparently (Impact Frontier Collaboration 2020). [26]

Bridges Ventures follows a two-step approach to assess anticipated impact: the Bridges Impact Radar to assess the development impact of an investment and then using the Impact Scorecard to define key performance indicators (KPIs) and targets that substantiate the impact claims and can be tracked to monitor and manage impact. The Bridges Impact Radar assesses impact of an investment along the following four key criteria. 1) Social or Societal Outcomes, their scale, depth and systemic change, 2) Additionality or whether the target outcomes would have occurred without the investment; 3) Environmental, Social and Governance factors; and 4) Alignment that assesses the relationship between an investment's ability to generate impact and its ability to deliver competitive financial returns.

LeapFrog's framework was designed to integrate not only ESG and impact considerations, but also the financial aspect throughout the entire investment design process. Investment officers use a distinctive measurement framework, which encompasses Financial, Impact, Innovation and Risk Management factors (FIIRM). Officers do not work in separate teams. On the contrary, each investment opportunity is evaluated from the start using FIIRM on key financial and impact considerations, including ESG and sustainability. The results enable clear identification of the "profit-with-purpose opportunity" and highlight areas of focus for further diligence. All aspects of the FIIRM framework are applied with due diligence to each investment. Due diligence is supplemented by data collected by the Consumer Insights team on consumers' unmet demands, "pain points," perceived future risks, and drivers of satisfaction. The investment committee integrates FIIRM results and Customer Insights to holistically evaluate the performance of potential investments. The due diligence results from FIIRM help crystalise company-level impact targets and action plans and enable alignment with LeapFrog's principles for responsible investment.

Another area of good practice emerging from the private sector relates to the way results from the anticipated impact assessment are followed through across the implementation /

[26] https://ssir.org/articles/entry/how_investors_can_integrate_social_impact_with_financial_performance_to_improve_both#bio-footer

monitoring and evaluation phases of the project cycle. **Integrated systems that support the complete investment cycle can help investors to clearly monitor investment performance against anticipated impacts and targets, as well as substantiate and validate their impact cases.**

Actis monitors the Impact Score throughout an investment to see precisely how it is performing from an impact perspective. The current score is compared against the score at the time of initial investment ('baseline') to understand how much impact has been added. This is called the impact multiple; the greater the increase in impact, the bigger the multiple.

Similar to the impact score mentioned earlier, assessments of the impact multiple are also presented across investments.

Similar to Actis, LeapFrog and Acumen monitor the expected case for impact throughout implementation. Tracking of impact is done through quantitative and qualitative indicators, which include monitoring of customers' responses, enabling a better understanding of the impact on the ultimate beneficiaries.

An interesting finding emerges from the Actis's case study with regard to the way ex-ante tools can play a part in external accountability mechanisms. While most of due diligence tools are proprietary and assessments are usually not publicly available, Actis's due diligence tool is open source and its description and examples of its outputs are publicly disclosed and presented on Actis's website. Committing to a greater process transparency would enable investors to clearly show their investment rationale, increasing trust in their investment decisions or as argued by PCV 2019 "reducing concern about impact washing, or unsubstantiated claims of impact". LeapFrog also does the same with an independent "audit" that assesses its adherence to IFC's impact investing principles.

4 Comparative assessment

This comparative assessment is based upon our detailed assessment of the target DFIs and the analysis of the comparator investors, which include two DFIs (CDC and Proparco) and five private impact investors. We found sufficient diversity within and between the target DFI and comparator group to identify significant differences in performance of *ex-ante* impact measurement tools and approaches. This has been helpful in informing our conclusions in the next section.

This assessment is structured around the four key assessment criteria that form the basis of this investigation: adequacy of scope; robustness; ease of use and integration. It concludes with a brief description, drawing on good practice, of what we judge to be features of a "best in class" tool.

4.1 Scope

Development impacts are defined and measured differently in the DFI tools. DFIs all adopt a partial view of impact and, in some cases, an extremely narrow approach is used, which precludes any indirect or systemic impact. IDB Invest aside, there is a scarcity of rigorous results chains in most DFI investments. This contrasts with most private impact investors analysed, where impact is comprehensively defined, and each investment has a Theory of Change detailing how the investment will achieve the impacts envisaged.

All DFI definitions of impact are partial: None of the DFIs impact measurement tools capture 'development impact' across all five dimensions of impact defined in the sector-wide guidance from the Impact Measurement Project. MIGA has only recently started to measure its impact systematically in a pilot exercise, whereas the AfDB tool has been in place since 2007. This represents a contrast with our comparator group, where impact measurement has been central to the operations of several of the private investors from inception. Two of the private impact investors – Actis and Root Capital – assess impact across all five dimensions of the IMP framework. This suggests that these privately-financed investors take a wider approach to development impact than some of the publicly-financed DFIs – an incongruous finding when the rationale for public funding of DFIs is based upon them stimulating development impact when the market fails but perhaps understandable in light of heterogenous investor interest, as well as the lack of a need to harmonize impact measurement.

Most DFIs have a compelling strategic vision: Most DFIs have a strong strategic focus, where the overarching aim and cumulative impact of numerous investments is explicit and clear. Others have an opaquer goal and, as a result, it is not immediately obvious that the portfolio of investments resulting amounts to much more than a collection of projects rather than incremental steps towards a greater, overarching goal. The EBRD defines its impact as supporting countries to transition towards well-functioning markets. DEG has a similarly bold vision of impact as structural change in society and the environment and contribution to SDGs. IFC and MIGA have a vision aligned to the WBG twin goals of reducing inequality and eliminating poverty, with IFC also contributing to SDGs and market creation.

IDB Invest understand impacts as benefits generated on the economy and society, with particular focus on the development priorities identified by the SDGs, as well as IDB Invest's institutional priority areas (i.e. climate change, gender equality and diversity, MSMEs, and serving the region's smaller economies and small and island countries). Only the AfDB has a more modest 'vision' of impact, which is effectively the aggregation of individual project direct impacts aligned with the second tier of its Results Measurement Framework that also reflect priority areas (climate and gender).

Interestingly, the HIP Investor and Actis in our comparator group have also aligned their investments with the SDGs. The fact that some progressive private investors are structuring their impacts around the framework of the SDGs gives a clear indication of the power of the SDGs to convene support from a diverse range of stakeholders.

The link between the strategic vision and the project operations of the DFI is not always clear: Some DFIs, notably IFC, come closest to having a results framework linking the two levels of strategy and operations. The problem is defined as gaps in access to infrastructure or services in their sector plans and constraints to market creation and individual projects are evaluated in terms of their ability to contribute resolving these issues. IDB Invest adopts a more project-based approach. The DFI requires projects to present a project logic articulating the development challenges and causes and how the project will contribute to the development objective, and related SDGs and IDB Invest's priority areas. The robustness of the project logic is assessed prior to investment approval through an evaluability assessment. Some DFIs have a much less clearly defined linkage between the impact of individual projects and their strategic vision. For instance, the AfDB appears to consider each investment on its own merits – without an assessment of its contribution to a broader strategic goal.

Most private investors in our comparator group, with the exception of Actis, use a Theory of Change as the narrative to explain the logical pathway between an individual investment and the development outcomes sought. **No DFI tools conciliate transparently between financial viability and development impact:** In some DFIs, such as AfDB, only financially viable projects are assessed for development impact – so the implicit prioritisation for financial viability is 100% because non-viability is effectively an exclusion criterion. EBRD will drop any project if not deemed financially viable (sound banking) yet the board and investment committee are more likely to accept projects with lower return when the impact is expected to be high. Other DFIs, such as DEG, appear to avoid precluding projects with low financial returns by assessing the sector portfolio rather than each project. IFC adopts a portfolio approach with no hard thresholds on either development impact nor financial scores and is developing a tool that visualises the two scores. IDB Invest also uses a portfolio approach, but it assesses each project, as well as how each project impacts on financial metrics at the portfolio level. This approach is widely adopted amongst private investors.

Financial and non-financial dimensions of impact need to be explicit and comparable. Decisions on one often have consequences on the other, including the investment's objectives and duration. Yet most of the tools have their own teams with their own frameworks, all of which exist in varying degrees of isolation from their financial counterparts. Some projects are only considered eligible for preparation when they have already demonstrated that the financial case has been met. There is some evidence of DFIs giving projects from prioritised sectors or countries with high fragility additional points. However, this process is not fully transparent, and we found little evidence of the explicit treatment of the trade-off between commercial and development returns. There are some examples of 'good practice' from the private sector, notably Acumen, Root Capital and Leapfrog. The CDC development impact case and thesis makes explicit the trade-off between the fragility of the development context and the achievement of financial and development returns. The 'returns continuum framework' used by the Omidyar Network is a particularly transparent approach, where lower financial returns can only be justified by increasing social returns.

Making a robust case for additionality is fundamental to the rationale for DFIs being able to demonstrate positive net impact and, ultimately, justify their use of public funds. If DFIs displace appropriate commercial funding for projects, it is not only a waste of public money, but it also threatens to distort existing capital markets and impede development. Of the tools reviewed, IDB Invest is closest to assessing additionality with sufficient seriousness, followed by IFC and the EBRD. The other DFIs assessment of additionality is relatively weak and. Although most DFIs aspire to adopt the MDB harmonised framework for additionality, practice varies. Neither MIGA nor DEG provide an assessment and, while EBRD and IFC do, it is not

included in their project rating. IDB Invest gives serious consideration to financial additionality, although evaluations suggest that the result tends to be clouded by non-financial additionality. In the case of AfDB tool, the definition of additionality is stretched even wider (to include political risk mitigation), which can yield a high score for a project with low financial additionality. Surprisingly, several of the private financiers included in their impact measurement frameworks a serious treatment of the counterfactual scenario – what would happen if the investment was not made by this specific investor.

Some tools assess impacts at end-beneficiary and system level - such as IDB Invest and IFC. Others such as AfDB and to a certain extent DEG, lack this strategic focus and take a partial view of direct effects. However, all DFI tools, even those, which include end-beneficiary impacts, are relatively 'top down'. There is little evidence of the customer centricity of private sector investors, which engage more actively with investees and their customers.

DFIs treat and measure the term development impact differently among their indicators: IFC and IDB Invest, in addition to including impacts on market systems, follow the impact of their investee projects to the end beneficiary of the investment. EBRD does not, rather it focuses on the broader market. DEG focuses on measuring job creation and local income creation with limited evidence on market development and there are some indicators on environmental sustainability and community benefits. The AfDB focuses neither on the market nor the end beneficiary, but rather on the very partial measures of direct job creation and government revenue investees generate.

There are clear differences in the extent of effort to engage both with the investee and the investee's customers amongst the DFIs. However, the contrast between the DFIs as a group and the private investors in the comparator group is striking. Leapfrog, for instance, regard customer centricity as an issue of due diligence and have a team dedicated to collecting customer insights – to inform their assessment of the bankability of the investee as well as the likely impacts resulting from an investment. Acumen also assesses their potential investees' knowledge and understanding of their customers using Lean Data approaches. This focus on customer centricity is partly motivated by private investor's concern that the investee will be able to finance their investment but also on the impact this investment will have. The relative, although variable, lack of attention paid by DFI to engaging with 'end beneficiaries' contrasts with the focus of private investors on the final consumer.

Some DFIs recognise their systemic impact and others do not: IDB Invest, IFC and MIGA measure the impact of projects not only on end beneficiaries but also on the market system in which the DFI is operating. Root Capital distinguishes between the 'enterprise impact' of its investments (the impact on customers, suppliers and the environment) and the 'investment impact'. However, the latter is not an assessment of the systemic impact of the investment but rather an assessment of the value added by the investment compared with a counterfactual where Root Capital does not invest. In other words, this is an assessment of additionality rather than systemic impact. The much larger size of the DFIs compared with the comparator group – for instance the capitalisation of EBRD with total equity of almost €18 billion compared with €1 billion capitalisation of LeapFrog - could explain the latter group not focusing on systemic impacts.

The nature and treatment of risk of failing to achieve projected outcomes differs between DFIs: IFC and MIGA provide a detailed risk adjusted score to inform the investment decisions, and IFC covers several dimensions. IDB Invest extends this analysis to include sensitivity analysis. EBRD assumes all its investments are high risk. AfDB undertakes a lighter likelihood assessment of development outcomes compared to IFC and IDB Invest, while DEG does not consider them beyond risks of breaching safeguards through an ESG assessment.

4.2 Robustness

Closely related to the issue of scope, is the **importance of establishing a robust case for the investment.** This should involve a clear diagnosis of the problem to be resolved, the intervention that will resolve the problem and the consequences of this with an emphasis on systemic effects and impacts on the end beneficiaries. Although some DFIs do measure impact at end beneficiary and systemic level, most DFIs, with the exception of IDB Invest and IFC, have unclear Theories of Change for specific investments and some do not even consider most of their impacts at either beneficiary or system levels. Establishing baseline data is crucial for evaluating impact.

All DFIs make use of harmonised development indicators: Most DFI tools develop and predict values for indicators derived from harmonised sets, notably HIPSO and IRIS+. In most cases, these are expressed quantitatively and typically broken down by gender. Only EBRD, AIMM and DERa include an assessment of qualitative characteristics. Most DFIs only assess indicators quantitatively and there is limited evidence of indicators being used sensitively to assess development impacts on end beneficiaries.

All tools, excepting TOMS, collate, score and rate the adequacy of information collected by different members of the appraisal team within DFIs. Financial information from investment and credit officers is collated with data from those responsible for environmental and social safeguards and legal officers all of whom are typical representatives on project appraisal teams. IDB Invest do this in a more integrated way, with different assessments undertaken by the same team, including impact officers as well as risk and investment officers.

Experience from the impact investors suggests that the nature and extent of consultation with investees and their clients is critical. Non-financial support from DFIs in the form of technical assistance or policy dialogue can have a significant impact on the likelihood of investment success and, without consultation, DFIs may not even be aware of the need for this non-financial support. Consultation is also critical for the investor to understand their likely development impact and even whether the investment will work commercially. While most tools carry out varying types of due diligence with DFI investees, most of these focus on risk reduction rather than investee support or understanding development impact, in particular the investee's outreach and their customers. The implicit role, which most DFIs appear to hold for themselves is as the provider of investment capital rather than expertise and policy influence.**Amongst DFIs, investee consultation is generally lacking:** All tools invest resources to consult with investees. However, this communication appears to be motivated by the need to reduce credit, operational or reputational risk. This communication is limited to validating data generated through the social, financial and environmental assessments together with management and operational capacity assessments. What is missing is the notion of customer centricity, which is critical to many of the private investors. This requires detailed consultation not only to understand the non-financial needs of the investee and the likely development impact of the investment among their clients but also extensive analysis of the perceptions and behaviour of the customers themselves. This informs discussions about opportunities for innovation in products and services. IDB Invest discusses the project logic with the client but the degree to which the lender's objectives and targets are aligned with those in the business plans and the extent to which these reflect a commitment to development impact is not clear.

DFIs usually do not establish valid counterfactuals: The apparent absence of rigorous consideration of the counterfactual scenario is a serious weakness in DFI impact assessment methodology. As a critical input to an assessment of additionality, without a clear counterfactual analysis the DFI cannot robustly estimate its own net development impact. Where such assessment is explicitly conducted, such as with DELTA and ADOA, it is not

always clear how this assessment is made. This makes it difficult to judge whether their impact *ex-ante* refers to the gross or net impacts of investments.

The use and treatment of Theories of Change in assessing the investments' evaluability vary considerably: DELTA assesses the evaluability of specific investments based on a Theory of Change. The analytical framework for AIMM is based on sector level theories of change but each project defines its own development impact thesis. Sector level theories of change inform design features for EBRD projects and ADOA's indicators are presented separately to those found in project results frameworks. DERa has a corporate level Theory of Change only. Several private sector investors in our comparator group make extensive use of investment specific theories of change. These are an important tool to make explicit the linkages between an investment and the development impacts. They also made explicit the assumptions – and therefore risks – on which these linkages are based.

The degree of sector specificity with indicators differs: Some tools measure all projects against the same set of outcomes (DEG and the AfDB's two deep core indicators). Others, including IDB Invest's DELTA and ADOA's sector indicators of AfDB, allow for some flexibility and tailoring. IFC measures a different set of outcomes, which are tailored to each sector and, within this framework, project promoters can select the outcomes that are most relevant to the investment; EBRD's indicators are tailored to each sector and component of a market.

Contrasting quality assurance processes: IDB Invest and IFC have robust quality assurance processes in place, both internally - with peer reviewers and expert panels - and also externally - provided by an independent third party – similar to Leapfrog. Other DFIs, notably DEG and AfDB, have a more 'light touch' QA process, with AfDB's being more procedural. DFI quality assurance processes are generally characterised by 'top down' assessments by the financier of the investee and, on occasion, the end beneficiary. Amongst some of the private sector investors in our comparator group this 'top down' assessment is triangulated with a 'bottom up' assessment directly from the customers of the investee.

4.3 Use

The application of tools to investments is universal: Although DFIs introduced *ex-ante* impact measurement tools at different times over the past 14 years, the use of tools is now ubiquitous. All tools are used to assess every investment and in the case of IFC even advisory projects, and their outputs are used in the decision-making process, including appraisal and investment approval, for all organisations.

Tool results have a direct influence on the stop/go investment decision: All tools estimate individual investment level results that score their different dimensions of outcomes/modules and provide scores for each that are summarised into an overall rating.

No tools adopt clear thresholds or decision rules for approval at board level: Yet most apply minimum thresholds based on their rating system for senior management/ investment committees to filter investment at earlier preparation and appraisal stages. At IDB Invest, the minimum DELTA project score required for an investment to advance depends on the expected risk-adjusted return on capital. The returns continuum framework of the Omidyar Network, which explicitly requires a higher development return from projects with a lower financial return appears to be mirrored in IFC's Portfolio approach.

Some tools are highly quantitative and require all issues of impact to be distilled to a number. Others, particularly those used by the private sector, recognise that this approach risks being reductionist; they have therefore created space for a narrative on impact to accompany the quantitative analysis.

Most tools can steer projects towards specific sectors or geographies: DERa provides an uplift specifically for projects in certain enabling sectors, and projects in low income countries and DELTA, AIMM and EBRD's tool provide an uplift for projects that address an existing gap (which is correlated with countries / sectors that are less economically developed) or an institution's priority. ADOA uses standard weighted scores for projects depending upon the sector of the investment – but these are partial in that they only focus on direct effects.

User manuals for the tools and the training of responsible teams are most comprehensive for those using AIMM, DELTA and ADOA. There has been a clear effort by impact measurement teams to disseminate awareness of tools amongst the investment staff of DFIs. A different approach is adopted by DEG where the tool has been designed to be so straightforward (the assessment can be completed in one day by a non-technical member of staff) that it is implemented by the project team itself.

Several DFIs (IFC, IDB Invest, DEG and EBRD) adopt a portfolio approach to investment appraisal: Individual investments are assessed but the stop/go investment decision is made taking into consideration their contributions to portfolio level targets. The portfolio approach to appraisals is characteristic of several of the private impact investors.

ADOA, IMPACT, TOMS, AIMM and DELTA influence pre-board approval investment decisions among senior management committees: As projects move from the preparatory to appraisal stages, the tools improve the quality of their design (to the extent that criteria adequately assess 'quality') and, for others, prompt rejection from the pipeline.

The tools' outputs are presented differently at investment appraisal stage: For example: the results of DELTA, TOMS and AIMM are integrated into the investment proposal to the Board; those of ADOA are presented separately to the Board in the form of a final note alongside the appraisal report and credit risk memorandum.

DERa's and ADOA's influence on board decisions, whether to invest, is somewhat circumscribed: DEG's Board is not a resident board. It is not involved in decision-making for individual projects and most projects decisions are made by Investment officers and Dept Heads. The impact of ADOA at Board level appears to be limited by a format, which does not reflect the richness of the analysis underlying it. For DELTA, TOMS and AIMM, the presentation of the overall portfolio has a greater and more strategic bearing on the investment decision.

The positioning, capacities and associated level of effort in impact assessment vary significantly: 21 staff work in the IDB Invest's development impact team and each investment assessment takes about 5 days; 8 staff work in the AfDB's dedicated ADOA team, independent from the appraisal team, and each project takes up to 25 days to assess. AIMM employs the most people in its development impact team and work appears very intensive. DEG has a small team of approximately 5 people responsible for DERa and impact measurement, taking only a day to assess each project.

4.4 Integration

The only way to prove the robustness of the ex-ante tools ex-anteis to find out how the projects assessed, perform during and after implementation. Some DFI impact tools have no role after the investment decision, which prevents the dissemination of learning between the design and implementation phases of a project. This also undermines the ability of staff to assess and improve the impacts forecast by the tool. IFC just announced systematic project-cycle assessment of impact, allowing for systematic assessment of ex-ante reliability.

Most tools are integrated into the project life cycle somewhat: Most DFIs monitor the impacts estimated by tools into the implementation phase of investments through their routine monitoring system through, for example, supervision missions and at investment completion. However, use of impact measurement tools to actively encourage learning and the management of impact appear to be less frequent amongst DFIs than impact investors

Most tools track indicators. Impact monitoring involves a process of re-generating scores defined at approval, for example, annual re-assessments of the project score by DELTA and re-computation of DERa score based on an annual re-assessment of indicators. To what extent this is accompanied by a narrative analysis is uncertain. This study did not look at the adequacy of monitoring systems.

DELTA and AIMM generate reports on project and portfolio performance for both senior management and the Board. For DELTA it is clearer how lessons learned are assessed and fed back for improvement to ex-ante impact assessment, than for AIMM. In the case of DEG and EBRD, based on the evidence found, it is not clearly communicated to the general public how lessons are learned from tools.

ADOA's mandate stops at approval stage and so, too, currently does IMPACT: More importantly, ADOA's outputs are dis-connected from those that are subsequently monitored during implementation and evaluated at exit.

The extent to which results and lessons are tracked and identified through corporate results management systems varies: While the crowding-in of private capital is monitored, it is unclear how the same is true for the quality of DFI's non-financial added value. This contrasts strongly with the application of tools by Acumen, Actis, Bridges and LeapFrog, where the ex-ante impact measurement tools are fully integrated into the project cycle and are regarded as important mechanisms for oversight and accountability. In these examples the ex-ante tools are not just used as a framework for post-approval monitoring, but as a tool to manage impact during implementation. This process allows investors to substantiate and validate their impact cases. The Actis due diligence tool is open-source and examples of its outputs are publicly disclosed and presented on the investor's website.

In contrast, DFIs appear relatively more opaque: *ex-ante*most of the institutions do not fully disclose either the tools or the results they generate externally. Similarly, ADOA, DELTA and IMPACT, do not disclose methodologies used for the tools externally. While some information on appraisal reports is made publicly available, neither the tools *per se* nor the results generated by them for individual investments are subject to a high degree of transparency.

4.5 Features of a "best in class" tool

Informed by the above analysis of practice among DFIs and private sector impact investors, the following define core features of what makes for a best in class tool:

1. A clear definition of development impact, making explicit the change to which the investment's outcomes contribute.

2. A compelling answer to the question: what do DFIs bring (financially and non-financially) to the investees' operations that commercial lenders cannot or do not bring?

3. Making clear how the financial and development returns are conciliated, in parallel, not in sequence.

4. A clear Theory of Change for each investment, breaking down the results (outputs, outcomes and impact) that makes explicit underlying assumptions.

5. The establishment of a valid counterfactual that complements the case for additionality (the inputs to, and outputs of, the investee) to attribute the indirect outcomes to the DFI's investment.

6. Independent, third party assessments as to the quality and usefulness of the tool and the lessons learnt, the results of which are made public. Making publicly available the description of the tool and its results.

7. An approach to investee consultation that clearly aligns the objectives of the lender with those of the borrower, understands the non-financial needs of the investee and analyses the perceptions and behaviour of the investee's clients – the ultimate beneficiaries - and specifies who they are.

8. Using a risk adjusted score to ensure that uncertainties of the anticipated outcomes are reflected in the final recommendation to decision-makers.

9. A Portfolio approach to help decision-makers steer investments to countries and sectors prioritised in the institutions' strategic plans and that makes clear their contributions to portfolio-level targets.

10. Ex-ante impact data, on which the investment is approved - e.g., predicted values of indicators, the actual non-financial value bought by the DFI and assumptions – to be followed through into project implementation, taking into consideration existing responsibilities and reporting arrangement among those responsible for implementation and independent evaluation.

5 Conclusions

5.1 Target DFI Conclusions

5.1.1 IFC' AIMM

The scope of AIMM is adequate. Similar to other DFIs, AIMM assesses development outcomes, not impacts. However, the tool is unusually holistic. It assesses direct and indirect outcomes along economic, social and environmental dimensions in terms of reach and depth for each project, as well as assessing the contribution to market creation (systemic effects). The tool is also strategic in the sense that the project assessment takes account of the existing gap in markets (i.e. extent to which markets are underserved) and can also differentiate between different categories of end stakeholder (i.e. gender, income, youth). The final AIMM score is risk-adjusted for each project based on likelihood of achieving impact and takes account of operational, sector, country and political risk factors. Due diligence and ESG assessments, whilst external to AIMM, can influence project scores by identifying potential negative effects of the project. Additionality is not part of AIMM but is measured based upon financial and non-financial aspects and is a key factor in the investment decision.

AIMM allows for measurement that is partially project specific while at the same time grounded in systematic frameworks that allow comparability across projects. The analytical backbone of the assessment relies upon sector frameworks that define theories of change for each sector. Each project then has a development impact thesis supporting it, which identifies the main expected effects of the project. Not all projects measure the same set of outcomes and the choice of outcomes and indicators to select rests with the rating teams. While these outcomes and indicators are selected from pre-defined lists in the sector frameworks, it is not clear if there are any incentives to cherry pick certain outcomes and indicators that are more likely to be achievable. The assessment methodology and data used in the tool is guided by detailed sector frameworks, which define for each sector the list of potential outcomes and indicators, the data sources, benchmarks for assessing the gaps and intensity of change, and how the rating is to be completed. The frameworks also include detailed market typologies that guide the assessment of contribution to market creation, which is quite ambitious in comparison to the other DFIs studied. Evidence is provided by economic models, the judgement of sector specialists and in-country staff and extensive diagnostics. Some of the indicators used in AIMM are adopted from harmonised frameworks with the aim to minimise proliferation of metrics. The final output is a single, risk-adjusted score that can be compared within and across sectors and countries and aggregated at portfolio level. There is a good quality assurance process in place whereby assessments made by investment officers are reviewed and validated by sector specialists who work closely together during project appraisal. In addition, half the projects are reviewed by the AIMM Panel, which is at arm's length from the operations team. When it comes to additionality, the assessment relies upon a qualitative evidence-based assessment and a guidance note has been produced, which aims to promote consistency in additionality assessment across different teams within IFC.

AIMM is integral to the investment decision, with management and Board review. AIMM provides the information in a format to aid decision-making. Individual projects are assessed using a portfolio management approach, with targets set at portfolio level. A certain level of financial returns is not assumed as a pre-condition, rather the portfolio approach helps to make informed decisions on a case by case basis, allowing to invest in projects with high impact potential but more uncertain financial returns and vice versa. Even though there are no hard thresholds on AIMM score, distribution of scores from past years show that implicitly almost all projects accepted have at least a 'Satisfactory' score, which corroborates the importance

of the AIMM score in decision-making. The tool assesses projects to steer investments towards areas where investment gaps are widest, which are typically in lower income and FCS countries. However, implementing the tool is resource-intensive, requiring significant resources and expertise.

AIMM is well integrated into the monitoring cycle with ex-ante indicators monitored annually. Given AIMM is new, it is not yet fully integrated with the monitoring system (DOTS), with indicators related to market creation mostly missing in DOTS; however, IFC is working to update DOTS to have a fully integrated system. AIMM scores are recalculated annually and compared against the ex-ante score, which provides the investment teams with real time data to identify potential issues and take corrective measures. When it comes to ex-post assessments, AIMM is less integrated. Although impact teams conduct a self-assessment of projects and the Independent Evaluation Group validates these, this is not done for all projects. Furthermore, IFC conducts few ex-post evaluations and these are mainly thematic, so there is limited independent assessment of project outcomes. AIMM is well integrated into overall RMS with results from the tool feeding into corporate level scorecards. Internal accountability is strong, but AIMM scores are shared with clients inconsistently and not at all with the public.

5.1.2 IDB Invest' DELTA

Despite having some weaknesses, IDB Invest's approach to manage for impact stands out for three reasons.

DELTA uses a hybrid model combining aspects of the impact rating framework and impact monetisation framework. The use of a combined approach is not unique, with several impact investors relying on features from multiple archetypes (IFC 2019). However, the way DELTA combines these two frameworks with an analysis of alignment to ESG requirements, project financial sustainability, and an assessment of additionality, represents an interesting example, which has no comparator across the other DFIs reviewed by this study. Through the economic analysis, the tool makes an explicit attempt to frame impact in the most used language of value, i.e. monetary terms, and capture a comprehensive range of positive and negative, direct and indirect effects of an investment. At the same time, DELTA acknowledges that there are inherent challenges in this approach, with some dimensions of impact that are difficult to capture and monetise. For this reason, the tool complements the economic analysis with a stakeholder analysis, which explores the types of direct and systemic effects generated by the investment, and who ultimately benefits from these effects. The stakeholder analysis adequately relies on the project logic and results framework to identify the expected effects of the project and ultimate beneficiaries. The project results frameworks use harmonised indicators, which are selected by sector and type of investment/transaction, with the aim to minimise proliferation of metrics. Beside the comprehensive assessment of the investment's benefits, the final impact score incorporates an examination of ESG factors, an analysis of the investee's business sustainability and an assessment of the additionality that IDB Invest brings to the project. The impact score practically synthetises several dimensions of the investment appraisal, bridging divides between different appraisal areas. The score provides a holistic, fact-based judgment of the impact potential of an investment. Impactful projects are identified as those investments that generate economic and social benefits to the underserved, that would have not been available without IDB Invest's contribution and are implemented by profitable and sustainable businesses with ESG practices aligned to IDB Invest's requirements.

Despite the exhaustive scope of the DELTA assessment, the interpretation of the score may be sometimes problematic. Given the way the impact score is computed, different dimensions can drive a similar score. This limits comparability of the score across projects and complicates the interpretation of the development impact measured by the tool. In addition, while the robustness of the assessment is ensured by a good internal and external quality assurance

process, as well as by the use of best practices and clear assessment guidelines, stakeholders' consultation practices appear to be somehow limited. There is little information on the extent to which DELTA consults the investees and on the involvement of ultimate beneficiaries in the assessment of the investment's impact. Similarly, while the assessment of impact is explained to be based on a comparison with a counterfactual scenario, it is not clear how this operationalised.

The way DELTA is used for decision-making is usefully rooted in a portfolio approach, which assesses each project for its impact potential and financial returns (FCR), as well as its contribution to the overall portfolio. The Board approves new investments based on the interplay of the impact score and the expected risk-adjusted return on capital, with higher impact potential required for investments with relatively low financial returns. At the same time, investment decisions aim to move toward a portfolio of projects, which represent an efficient impact frontier, optimising impact and financial performance simultaneously. This approach appears to be pioneering, with its merit laying in accounting for the differences in the underlying risk among the investments and so helping assess how risks and returns trade-off. A certain level of financial returns is not assumed as a pre-condition, rather the portfolio approach helps to make informed decisions on a case by case basis, allowing to invest in projects with high impact potential but more uncertain financial returns and vice versa. The Investment Proposal used for approval includes the overall score, the development outcome score and additionality score, as well as scores of each subcomponent. This provides a consistent and comprehensive basis for decision-making.

The DELTA is integrated in the project life cycle, with measurement of development outcome and additionality reassessed and validated throughout implementation. DELTA metrics are collided in the Development Effectiveness Analytics platform and used to provide management with real time data to identify potential issues and take corrective measures. This enables IDB Invest to apply a consistent approach in the way the performance of the investment is assessed throughout the project cycle and validate assumptions and the basis for approving the investment given contextual factors affecting implementation. The impact score is also computed for completed projects. This is particularly relevant and could be leveraged to create a body of knowledge available to the teams to better inform future impact assessment.

5.1.3 MIGA's IMPACT

MIGA developed IMPACT in 2018 and has only concluded piloting its tool. Given its recent development and implementation, there is limited evidence to systematically assess the performance of the tool across the four criteria of our research framework. In addition, the team was not able to conduct interviews with MIGA's staff to help fill outstanding information gaps and generate further insights on the tool. While the documentation available presents an overview of the tool and the scope of impact assessed by IMPACT, detailed information on the methodology used to assess impact, on how IMPACT is used for decision-making and the extent to which the tool is integrated into MIGA's overall Results Management System is generally lacking. This represents a clear limitation to our review in terms of breadth and accuracy.

The tool mimics IFC's ex-ante assessment framework, although the documentation available does not provide sufficient information to assess similarities and differences in an extensive way. Similar to AIMM, the tool is designed to assess MIGA's development impact in support of the SDGs and the World Bank Corporate Goals to end extreme poverty and boost shared prosperity in a sustainable manner. However, it remains unclear how the tool operationalises this definition, with the assessment of impact ex-ante only focusing on the project's financial, economic and environmental performance (i.e. project outcome) and the project's contribution to encourage the flow of foreign investment in the country (i.e. beyond-project outcome). The

key difference between AIMM and IMPACT lies in the way the institutions define the beyond-project outcome. MIGA's beyond-project level outcome focuses on the DFI's ability to encourage additional flow of investments, whereas IFC views beyond-project level outcome as its contribution to market development. No further information is available.

Similar to AIMM, the tool includes an assessment of risks, with the final impact score discounted to take into account relevant risks at sector, country, policy and political economy level. As in IFC's system, additionality is not incorporated in MIGA's assessment of impact ex-ante. However, in contrast to IFC, it is not clear whether and how the project teams assess additionality separately and whether additionality is one of the components upon which a decision to invest is made. Based on the documentation available, MIGA provides additionality to private markets in terms of risk classes, coverage amounts, tenors and member states where few or no private insurers will provide coverage.

Our assessment of robustness of IMPACT is constrained by data availability. Disclosed information on the methodology used by MIGA to assess impact ex-ante, dates back to 2013. The tool seems to use inputs from a stakeholder analysis to inform the assessment of impact. This analysis is a common approach among private sector-oriented development finance institutions. Robustness of the analysis is enhanced by the use of general and sector specific indicators aligned to HIPSO. Nonetheless, it is not clear how IMPACT processes the findings from the stakeholder analysis across its dimensions and related sub-dimensions and there is no information available to assess whether MIGA's assessment process aligns to IFC's one. No information is available on the arrangements in place to quality assure the evidence generated by IMPACT, nor the extent to which investees and borrowers are involved in the ex-ante assessment process.

While the evidence is still preliminary, the tool seems to improve the project's quality at entry, as it moves from origination to the approval phase. IMPACT also helps to inform decisions on deprioritising projects with low expected IMPACT ratings at or before approval by the Board. Nonetheless, the way the outputs of IMPACT are summarised and presented to the Board provides little explanation on how the project impact is expected to realise, especially with regard to the project effects on the economy and society overall. This might undermine the effectiveness of the tool to steer decision-making towards high-impact projects.

Insights from a semi-structured interview indicate that MIGA is moving toward a more integrated system although it is too early to evaluate any changes in accountability mechanisms and MIGA's overall Results Management System. At the time of writing this report, IMPACT is used at ex-ante only, and it is not clear whether IMPACT indicators are followed through by MIGA's system of indicators used in execution (DEIS) and how the IMPACT assessment relates to the ex-post evaluation presented by MIGA's Ex-post Project Evaluation Report.

5.1.4 AfDB's ADOA

ADOA was the first DFI impact measurement tool to be implemented in 2008.

Given its purpose and mandate, the organisational positioning of the tool, the functional relationships the ADOA team have with other departments and its accountability to the Board define a questionable fit. ADOA sits uneasily with a Department whose principal function is to devise and carry out research and analysis on priority social and economic development issues for the benefit of the Bank's member countries and its programming. ADOA is removed from the mainstream of AfDBs operations and provides a challenge function similar to that of the Operations Committee and Quality Assurance (SNOQ). It is separate to the Independent Development Evaluation Office (IDEV) who also report to the Board.

The two questions that define the tool's purpose are excellent a) what do DFIs bring to the operation that commercial lenders cannot or do not bring?; and b) what are the expected development outcomes? These inform ADOA's seven categories of development outcomes, its three dimensions of additionality and provide a consistent and clearly defined basis for ADOA's independent assessment for projects in country and at regional level. However, ADOA does not adequately answer either question. AfDB's unconventional approach to assessing additionality potentially dilutes the significance of financial additionality. By including factors, which have little to do with the first question, such as political risk mitigation - projects with low financial additionality and improved development outcomes can receive a high 'additionality' score. ADOA does not investigate or inform any trade-off between financial returns and development impact because projects, which do not meet the financial viability threshold are excluded before being considered by the tool. The tool relies on exclusively numeric and limited measures of development outcomes. In the absence of a definition of development impact, this leaves unclear their significance in relation to Private Sector Development Impacts. Indicators of 'development' focus upon revenue generation for government and the creation of direct jobs among investees. This favours government-led formal sector projects and does not consider – and therefore tends to preclude - the indirect effects that arise from the type of broad-based economic development project that benefit low-income groups. There is a distinct lack of clarity regarding to which degree investee products and services are relevant to and adequately target the underserved.

The indicators for Development Outcomes are harmonised, prioritised through a weighting system and broken down by sectors. These are notable strengths along with the internal quality assurance relating to its procedures, the notes it generates for OPsComm and the Board and the team's participation in investee due diligence missions. There remain challenges, however. There is no assessment of the non-sovereign operations' evaluability based on their results framework or Theory of Change. ADOA's guiding principles of the counterfactual and attribution are a confusing interpretation of terms. The indicators for Development Outcomes and how measures of non-financial value and risks contribute to them are all dis-connected from results frameworks found in the Project Appraisal Reports. Related to this, development outcomes align inadequately with the Investee's own objectives (which are also not reflected in the financing agreement). Finally, there is limited evidence to avoid concluding ADOA's assessment of financial additionality is too lenient and forgiving.

ADOA notes' rating and scoring system helps the Operations Committee improve a project's quality at entry, as it moves from concept note to appraisal stage. They also inform decisions on dropping projects at or before approval by the Board. Furthermore, the ADOA team provides the Project Appraisal Team (PAT) with discipline and assurance. However, the lack of apparent incentives among the PAT, driven by an approvals culture after the financial case has been cleared, raises issues as to how well the bank understands and communicates itself as a development bank.

The format of ADOA notes' presentation to the Board does not reflect the quality of underlying research work. This diminishes the weight of the tool's findings in comparison with the Project Appraisal Report and Credit Risk Memorandum that are considered at the same time by senior AfDB decision-makers. There is no portfolio-level synthesis of individual project assessments, thus limiting the strategic usefulness of ADOA in guiding non-sovereign operations, in particular across priority countries on the continent. ADOA's integration into the non-sovereign operations ecosystem is limited, as ADOA starts after the programming stage of the project cycle and the approval of the Project Evaluation Note at preparation stage. Given the resources the Bank affords to non-sovereign operations' preparation and appraisal, there are questions about ADOA's efficiency given its purpose - a team of eight people spending up to 25 days per project define significant opportunity costs to the Economic Governance and Knowledge Management Complex.

ADOA's mandate stops at the approval stage, thus the tool becomes redundant as soon as the investment has been approved. The consequences are numerous and not negligible: a) it leaves the bank vulnerable regarding the different decision rules for investment approval versus performance standards that define their success during implementation and at completion, including the actual performance of the bank in adding non-financial value to the investee; b) it leaves results "on the ground" given the Private Sector Department's limited current capacity to monitor and analyse them and the supervision missions' focus on financial and administrative performance ; c) it leaves unclear who decides which indicators to track, as much as who tracks them and how; and d) in the absence of any systematic feedback loops, it limits the degree to which ADOA can learn and so improve its performance as much as take heed of lessons learnt from completed projects they advised the board to approve.

5.1.5 DEG's DERa

The scope of DERa is to an extent limited. As with most DFIs, DERa estimates development outcomes and not impacts. While the outcomes measured reflect economic, social and environmental dimensions and are linked to the SDGs, they are mainly restricted to the reach and depth of direct project outcomes. There is a crude assessment of indirect job creation and market development, which are inferred by direct project proxy outcomes (such as project sector or a binary 'yes' or 'no' on promoting innovation). Furthermore, the tool takes no account of the characteristics of end-stakeholders, beyond the average income prevailing in the project country. Beyond measuring some 'do no harm' indicators, risk assessments for projects are considered separately to DERa and the final DERa score is not risk-adjusted. Additionality is not assessed in DERa, and while it was indicated that DEG does assess additionality, the team was unable to access relevant evidence. Therefore, it remains unclear how DEG measures additionality and whether it is a core part of the decision-making.

The only Theory of Change is a generic DEG-level assessment to cover projects across all sectors and countries. This results in an assessment which is very high level. This is an aid to standardisation and comparability between projects, but at the cost of providing less useful information on the effects of any specific project. Furthermore, all projects measure the same set of outcomes, with slight changes for investments in funds and financial institutions. A baseline assessment is made with respect to all the outcomes and indicators together with a forecast in 5 years' time. The range of quantitative and qualitative indicators are mostly adopted from the HIPSO/IRIS+ harmonised frameworks with the aim to minimise proliferation of metrics. Data requirements for DERa are light and based on secondary sources as well as data sourced directly from the client as part of other due diligence activities. The final output is two scores (baseline and forecast) that can be compared across sectors and countries and aggregated at portfolio level. For quality assurance, the impact team reviews all DERa assessments before projects are submitted to the Investment Committee for approval.

DERa is an integral part of decision-making, however the board is rarely involved, and management are not always involved in investment level decisions. Individual projects are assessed using a portfolio management approach, with targets set at portfolio level. A certain level of financial returns is not assumed as a pre-condition, rather the portfolio approach helps to make informed decisions on a case by case basis, allowing to invest in projects with high impact potential but more uncertain financial returns and vice versa. No account is taken of size of gaps in guiding where to make project investments. On the other hand, based on how the portfolio target is defined, there are indications that there might be incentives to select projects that already have a high baseline DERa score (i.e. already perform well across outcomes prior to the investment). DERa encourages projects in lower income countries and in certain sectors that enable private sector development. The tool is efficient and quick and easy to use.

DERa is well integrated into the monitoring cycle and all ex-ante indicators are tracked annually and compared with the original scores. This provides the investment teams with real time data to identify potential issues and take corrective measures. However, it is less integrated with ex-post activities. DEG conducts few evaluations – and these are mainly thematic – so there is limited independent assessment of project outcomes and additionally project teams do not conduct self-assessments at exit as is done by other DFIs. It is not clear how DERa scores influence performance management. The internal accountability of DERa is blunted by the limited involvement of the Board in investment decisions and monitoring data for individual projects. Additionally, no project level development impact data is provided to the public.

5.1.6 EBRD's TOMS

TOMS' restricted dimensions of impact and its treatment of risks limit its scope. Unlike the other DFIs in this study, the focus of the EBRD tool is primarily with transition impact on the economy, meaning the systemic changes in the market resulting from the investment. This is an important consequence of investments, and one that is often overlooked by DFIs. However, the focus on economic transformation eclipses an understanding of the effect of investments on people. TOMS does not quantify the anticipated direct or indirect outcomes of their investments in terms of reach (i.e. number of people benefiting) or depth (i.e. scale of benefit) - an important omission in assessing development impact. While all impact claims made in TOMS are substantiated by indicators and targets for delivery, these indicators are only used to monitor projects once in implementation stage and they do not affect the expected transition impact score that is calculated ex-ante. When it comes to risk assessment, the tool applies the same high-risk rating to all projects, and therefore makes no attempt to identify the particular risks facing specific investments. Although additionality is not part of the tool, EBRD does assess additionality, both in financial and non-financial aspects and it is a key factor in the investment decision. However, it is not clear how EBRD measures additionality. In our understanding, the assessment of additionality has not been consistent to date and EBRD is currently piloting a questionnaire with the aim of standardising this assessment.

TOMS allows for measurement that is partially project specific while at the same time grounded in systematic frameworks that allow comparability across projects. TOMS is based upon sector level theories of change and each project has a narrative and selects up to two transition qualities upon which the ex-ante assessment is based. The indicators used to substantiate impact claims are selected from a repository of 130+ quantitative and qualitative indicators, which draw upon the HIPSO/IRIS+ harmonised frameworks and these are selected automatically by the system based on the impact claims made in TOMS (that is, there is no cherry-picking). The final output is a single score, which can be compared with other projects in the sector or country and can be aggregated at portfolio level. Quality assurance is provided by economists working closely with investment teams and about half of project scores are assessed manually following a preliminary assessment in TOMS.

TOMS is an integral part of decision-making by management and the Board, although smaller <€25m, uncontentious projects that are part of frameworks may be delegated by the Board. Individual projects are assessed using a portfolio management approach, with targets set for transition impact at portfolio level. A certain level of financial returns is not assumed as a pre-condition, rather the portfolio approach helps to make informed decisions on a case by case basis, allowing to invest in projects with high impact potential but more uncertain financial returns and vice versa. TOMS scores projects based upon whether they target a transition quality that has been identified as a gap in the country and the extent to which they align with country priorities – so the portfolio reflects the strategic intent of the EBRD.

The tool is well integrated into the monitoring cycle and all ex-ante indicators that substantiate impact claims are tracked annually during implementation – and compared with the original

score. This provides the investment teams with real time data to identify potential issues and take corrective measures. When it comes to ex-post assessments, TOMS is less integrated. Although investment teams conduct a self-assessment of all projects at exit, EBRD conducts few ex-post evaluations and these are mainly thematic, so there is limited independent assessment of project outcomes. Accountability is promoted internally by the Board monitoring annual performance scores and an independent evaluation group reviewing the self-assessment of all projects at exit and validating a small sample of them. These scores feed into the country level and corporate scorecards. When it comes to external accountability, unlike other DFIs, EBRD publishes the ex-ante TOMS scores of its projects on its website; however, these were not available for all projects and the annual performance scores are not shared with the public.

6 Recommendations

6.1 Summary

Five themes emerge that cut across all target DFIs: transparency, additionality, measurement of development impact, portfolio approach and customer centricity.

Transparency and accountability – as publicly-owned development institutions, DFIs should be at least as accountable as the 'best in class' in the rest of the development sector to the public who own and finance them. This implies publishing the impact measurement tool and methodology for public scrutiny. Producing information in the public domain on each investment (impact score, monitoring results and evaluations) is also recommended.

Additionality – at the very least all DFIs should make more explicit their assessment of financial additionality. In addition, given the centrality of financial additionality to the rationale for using public funds to finance DFIs, the ambition should be to improve the quality of additionality assessment to be able to demonstrate convincingly that DFIs are not displacing other investors and explain the effectiveness of the non-financial value they bring to the investee.

Definition and Measurement of development impact – all DFIs should clearly define development impact and have a Theory of Change for each investment to allow the project to be evaluable and then measure the contributions to (the action theory) and the consequences of these to the direct and indirect effects (the change theory) on end-beneficiaries as well as on the broader market system, making explicit the assumptions.

A portfolio approach – all DFIs should use a framework[27] to help make decisions in favour of investments that maximise development impact while maintaining financial sustainability. Doing so will allow them to consider investments with different financial returns depending on the expected development impact.

Customer centricity – there are compelling commercial as well as developmental reasons for investors understanding the end-customers of the products and services, which they are supporting. There is considerable scope for all the DFIs reviewed to significantly improve their customer centricity.

6.2 IFC

1. **Strengthen and afford priority to the results of the stakeholder analysis.** AIMM needs to make clearer who are the ultimate beneficiaries and the degree to which they are underserved. The sector frameworks include indicators related to beneficiaries (such as gender, income level). However, given that teams select outcomes/indicators most relevant for their project, will all projects assess the profile of beneficiaries?

2. **Improve its external accountability**. AIMM scores and a description of the anticipated development impact for each project should be published. We would also recommend that IFC publish the annual scores and those at exit.

[27] Omidyar Network's "Returns Continuum" and Root Capital's "Efficient Impact Frontier" are best in class examples among our comparator group.

3. **Incorporate actual results of predicted market outcomes.** IFC are currently working on updating their monitoring system. At the moment predictions of market outcomes ex-ante are not yet fully incorporated into the monitoring system.

4. **Improve the consistency of type of indicators selected.** Project teams purposively select outcomes / indicators that are seen as most relevant to the project, yet indicators vary, and quite significantly – from numbers of farmers reached to more ambitious changes such as changes in wage levels.

5. **Clarify the implications on the AIMM score of indicator selection**. The scoring mechanism works whereby a high score can be achieved on overall project outcomes if at least one of the indicators is rated very highly. Are there incentives to include other indicators in the assessment, which will not affect the score?

6.3 IDB Invest

1. **The DELTA should spell out more clearly what the tool defines as "development impact"**, for a better understanding and easier interpretation of the project score. The DELTA scoring system assesses an investment along multiple dimensions and provides a synthetic measure of impact, which reflects how the investment performs against each dimension (financial and non-financial additionality, financial sustainability, economic and social benefits, alignment to ESG requirements, and type of benefits and type of beneficiaries reached). However, a similar score can be driven by different dimensions. For example, a project score of 5 can result from high additionality (7) and low development outcome (4), or the other way around (DEO 2018). This implies the project score might be hard to interpret and compare across investments. In this regard, a plain definition of what is measured by the project score in terms of what kind of changes are expected as a result of the investment and how these relate to development impact would be beneficial.

2. **The DELTA should make clearer the degree to which products and services delivered by investees reach those underserved** - the ultimate beneficiaries – and level/degree of engagement with them in defining and clarifying impacts. The DELTA already explores who are the beneficiaries of the investment, with particular attention to the poor and most vulnerable. This could be made clearer and strengthened to systematically consult stakeholders beyond the investees. This would provide better insights on expected impact, as well as help identifying drivers behind outcomes. Collecting stakeholders' perspectives could also help testing assumptions underpinning the realisation of impact and better inform assessment of contribution

3. **As part of the consultation practices, the tool should make clearer the degree to which development impact is integrated into the business plans among investees**, as a demonstration of their commitment. Especially with regard to poverty outreach among underserved areas/populations, the relevance of the products and services and their anticipated responses among the ultimate beneficiaries. The tool already assesses the robustness of the project logic and causal chain between project inputs, outputs and outcomes. However, it could be made clearer the extent to which the tool takes into account any potential constraints to the investees' capacity to achieve the expected impact, including investees' capacity to learn and adapt and incentive structure.

4. **The portfolio approach, whereby higher impact potential is needed for projects with relatively low financial returns is "best in class" and this word should be spread.** The approach does not rule out investments with lower financial returns, rather it allows for optimisation of impact and financial performance simultaneously

5. **IDB should focus more on measuring development impact ex-post and link ex-post measures to expected impact metrics computed ex-ante and throughout execution.** This would help validate the ex-ante assessment but also generate a body of evidence, which could better inform future impact assessment. While DELTA is already used both ex-ante and in execution, the extent to which DELTA is used ex-post could be made clearer and strengthened. Also, the institution should work on improving incentives to leverage the impact knowledge available to the teams and take heed of lessons derived from previous investments.

6.4 MIGA

MIGA would benefit from greater transparency and clarity on the approach to measure development impact. This would help understanding how the institution adapted IFC's AIMM to its roles and mandates and what areas could be strengthened. Based on the limited information available, we recommend:

1. **Strengthening how the project outcomes reflect MIGA's contributions to the WBG's twin goals and relevant SDGs**, for example, through systematically designing a project Theory of Change, which clarifies the relationship between project inputs, results and objectives, as well as assumptions underpinning the achievement of the objectives.

2. **Considering including other dimensions of impact** beyond leveraging in additional investment flows or clarifying why this can't be done if MIGA's business model really precludes this.

3. **Reviewing more closely the value IMPACT has for decision-making** and its positioning regarding the project cycle – when IMPACT is used and by whom.

4. **More clearly explaining the outputs IMPACT generates**, with emphasis on analysing how and in what ways: a) they inform whose decisions in preparing, appraising and approving projects; and b) projects are anticipated to have an effect at societal and system level.

5. **Clarifying how MIGA takes into account additionality** and how the institution assesses what value its support (financial and non-financial) brings to its investee companies.

6. **Make clearer how the basis for approving projects, i.e. IMPACT's baseline and anticipated values of indicators , is followed through in monitoring its implementation and evaluating at exit.** In particular, how do IMPACT indicators relate to MIGA's system of indicators used to monitor development effectiveness throughout execution (DEIS)?

6.5 AfDB

For ADOA, we have framed our recommendations into three categories: Organisation and Systems; Value; and Methodology.

On Organisation and Systems, the following applies:

1. **AfDB needs to re-consider the institutional positioning of ADOA**. Given its strategic mandate, the responsibility for fulfilling the purpose of ADOA is segregated from operations and does not appear relevant to and aggravates the capacity of the Economic Governance and Knowledge Management Complex. AfDB should review

this arrangement with a view to either moving it to the Independent Development Evaluation Office (BDEV) or, perhaps more appropriately from an operations perspective, the Operations Committee and Quality Assurance (SNOQ). The SNOQ has the mandate of developing improvements and overseeing their implementation, and non-sovereign operation quality at entry is very much part of the current quality agenda.

2. **AfDB should explore ways on how ADOA can be integrated into the monitoring of project portfolios**. That ADOA becomes redundant after approval is well known. Spotting opportunities to resolve this needs to answer two questions: a) why, unlike all other DFIs, is this the case and persistently so?; and b) how, why and by whom can what ADOA does and produces be best followed through, particularly in relation to existing M&E processes and outputs?

3. **AfDB needs to develop and mainstream incentives to improve the quality of non-sovereign operation project design**. While we understand the AfDB is in the process of resolving problems concerning the approvals culture, of note, through developing KPIs for project quality, by themselves these KPIs will not resolve the issue. It will be important to diagnose the underlying reasons for the prevailing approvals culture. The results of this analysis could usefully inform ways of encouraging, not just measuring, how well task managers improve the quality of non-sovereign operation design and, so too, supervision missions during implementation.

On Value, the following is recommended:

4. **Non-sovereign operations' Development Outcomes need to more clearly explain their significance for the Bank's contribution to Development Impact.** While ADOA makes passing reference to relevant High 5 priorities in Level 2 of the AfDB's Results Measurement Framework - the Bank's contribution to Development Impact – it is not clear what this means.

5. **AfDB needs to better define what value ADOA's process and outputs bring**. ADOA is appreciated and does bring some value, despite limits to its Notes, to the PAT team, informs decisions by the OPsCoMM and, to a lesser extent, the Board. That said, the efficiency of the process and the resources allocated seem cumbersome and disproportionately high respectively. To remedy this, we recommend the efficiency of the process, the resources required, and the robustness/quality of its outputs be reviewed in relation to other efforts and outputs at the preparation and appraisal stages.

On Methodology, we recommend the following:

7. **The development outcomes need to prioritise micro level effects**. ADOA's seven development outcomes cover a balance of categories. Undue emphasis, however, is placed on jobs created by investees and revenues they raise for government. AfDB should prioritise reference to the investee's clients, the ultimate beneficiaries, and clarify who they are and to what extent they are underserved (as specified in the Level 2 indicators in the Bank's Results Measurement Framework). Neither the positioning of the Bank - through on-lending to wholesalers - nor the sectoral nature of non-sovereign operations adequately explain why this should not be done. Equivalent emphasis should also be placed on private sector development and demonstration effects, so prioritising broad-based economic development that benefits low-income groups, as well as crowding in private sector investment.

8. **Strengthen and focus the assessment of additionality**. The basis on which additionality is assessed should prioritise and make a more robust assessment of

financial additionality, make clearer the benefits of the non-financial additionality investees realise through improving development outcomes and drop the count of political risk mitigation. This last dimension is not part of the MDB harmonised approach. Identifying political risk and ways of mitigating it should be part of the credit risk team's mandate.

6.6 DEG

1. **Incorporate a measure to assess the risk of development impact.** We recommend incorporating a measure of risk or likelihood of impact claims being achieved. We would recommend using this to calculate a risk-adjusted score for development impact similar to IFC and MIGA

2. **Strengthen and afford priority to the results of the stakeholder analysis.** DERa needs to make clearer who are the ultimate beneficiaries and the degree to which they are underserved. DERa provides incentives to invest in lower income countries – but the stakeholder analysis is limited at macro level and not at individual or community level.

3. **Review the incentives for project selection associated with setting the target based on actual and not forecasted scores**. DERa computes a baseline and forecasted score for each project at appraisal. This may imply that there might be incentives at times to select projects that already have a high DERa score at baseline, rather than those that create the biggest change.

4. **To take into account the size of the existing market gaps.** DERa doesn't currently do this except at country level. Differentiating specific market gaps would help **steer** projects in underserved markets.

5. **Improve its external accountability**. The DERa scores and a description of the anticipated development impact of each project should be published. We also recommend publishing the scores annually and at exit.

6. **Strengthen learning from projects by carrying out a self-assessment at exit**. DEG monitors projects annually and re-computes the DERa score. DEG also does this at exit. However, unlike EBRD or IFC, DEG does not conduct a self-assessment of projects at exit. The scope of this exercise would look at effects beyond those solely measured by DERa (which is limited given it measures a set of common indicators across all projects).

7. **Consider designing project specific theories of change**. DERa measures a common set of outcomes across most projects that define a limited basis for assessing their development impact ex-ante and ex-post (see above). Setting out how the project inputs will lead to changes in outcomes and contribute to the ultimate impact, as well as risks and assumptions underpinning them, will improve the robustness of project design at appraisal. It will also, as explained in the previous recommendation, provide more opportunities for learning, including during project implementation.

6.7 EBRD

1. **Strengthen and afford priority to the results of the stakeholder analysis**. EBRD's assessment focuses more on the market and not on direct project outcomes at micro level. There is a need to strike a better balance by defining who are the project beneficiaries and the degree to which they are underserved and to consider giving higher scores to projects that are serving those most in need.

2. **Conduct an impact risk analysis for each of their projects.** While the TOMS ex-ante score is risk-adjusted, this is not based on an actual assessment of the risk/likelihood of impact materialising for each project. All projects are rated as high risk. Project specific risk assessments will help EBRD identify where they can support projects to mitigate risks and so improve prospects for impacts.

3. **Establish a common framework for assessing additionality for all projects**. While defining some indicators. EBRD does not have a common framework. The current initiative in piloting an additionality questionnaire, coupled with a review of the harmonised approach to additionality adopted by MDBs, could usefully reveal the design features of such a framework including the identification of indicators. Tracking these during implementation would substantiate the claims they make on additionality ex-ante.

4. **Publish the scores for all projects**. EBRD do publish the TOMS score online, but not for all projects including their actual performance scores during implementation.

Annexes

Annex I - Bibilography

Literature Review

1. A Maffioli (2018) Development Effectiveness Learning, Tracking, and Assessment Tool – DELTA. Audiovisual presentation Development Effectiveness, IDB In July 13, 2018 - Washington D.C.
2. C Broccolini Lotti G Maffioli A Presbitero A F and Stucchi R (2019) Mobilization effects of multilateral development banks." International Monetary Fund, Working Paper WP/19/28.
3. DW te Velde (2011) The role of development finance institutions in tackling global challenges *ODI working paper*
4. Eurodad (2015) Monitoring and evaluation at Development Finance Institutions *A Eurodad briefing paper*
5. Global Impact Investing Network 2020 The State of Impact Measurement and Management Practice, Second Edition GIIN publication https://thegiin.org/research/publication/imm-survey-second-edition
6. IFC CREATING IMPACT (2019) The Promise of Impact Investing http://documents.worldbank.org/curated/en/146131573015554306/pdf/Main-Report.pdf
7. *I Massa, M Mendez-Parra, D W te Velde (2016) The macroeconomic effects of development finance institutions in sub-Saharan Africa ODI report https://www.odi.org/publications/10664-macroeconomic-effects-development-finance-institutions-sub-saharan-africa*
8. I So & A Staskevicius (2015) A Measuring the "impact" in impact investing, Harvard Business School https://www.hbs.edu/socialenterprise/Documents/MeasuringImpact.pdf
9. Kenny C and Moss T (2020) What to Do When You Can't Prove DFI Additionality CGD note
10. M Bannick P Goldman M Kubzansky and Y Saltuk (2017) Across the returns continuum Omidyar Network https://ssir.org/articles/entry/across_the_returns_continuum# Stanford Social Innovation Review
11. M Heinrich (2014) Demonstrating additionality in private sector development initiatives. A Practical Exploration of Good Practice for Challenge Funds and other Cost-Sharing Mechanisms DCED
12. Multilateral Development Bank: African Development Bank, Asian Development Bank, Asian Infrastructure Investment Bank, European Bank for Reconstruction and Development, European Investment Bank, Inter-American Development Bank Group, International Financial Corporation and Islamic Corporation for the Development of the Private Sector (2018) Multilateral development banks' harmonized framework for additionality in private sector operations
13. N Van de Sijpe & P Carter & R Calel, 2019. "The Elusive Quest for Additionality," Working Papers 2019022, The University of Sheffield, Department of Economics.

14. OECD Secretariat (2020) Impact Standards for Financing Sustainable Development (IS-FSD) Background Note to inform the expert discussion

15. OECD. (2015). An Alternative/Expanded Institutional Approach for ODA Eligibility of Private Sector Instruments. OECD Publishing. See: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DCD/DAC/STAT/RD%282015%293/RD5&docLanguage=En

16. Oxford Policy Management Limited (OPM) (March 2020) Analysis of International Financial Institutions' (IFI) and Development Finance Institutions' (DFI) private sector operations' impact measurement and project assessment tools. GIZ Concept Note

17. P Brest and K Born (Fall 2013) When Can Impact Investing Create Real Impact? Stanford Social Innovation Review

18. Pacific Community Ventures (2019) The Impact Due Diligence Guide Practical guidance for investors seeking to systematically assess investments' anticipated impact http://www.pacificcommunityventures.org/wp-content/uploads/sites/6/FINAL_PCV_ImpactDueDiligenceGuide_web.pdf

19. S Attridge, R Calleja, M Gouett and A Lemma (2019) The impact of development finance institutions: rapid evidence assessment. London: Department for International Development.

20. S Bilal et al (2019) Harmonisation of impact measurement among development finance institutions and beyond: opportunities for European Leadership *Chapter in S Attridge, DW te Welde and SP Andreasen eds Impact of Development Finance Institutions in Sustainable Development: An essay series ODI-EDFI*

21. World Business Council for Sustainable Development and IFC's Measuring Impact Framework Methodology, 2008

22. Y Arvanitis, M Stampini & D Vencatachellum (2015) Balancing development returns and credit risks: project appraisal in a multilateral development bank, Impact Assessment and Project Appraisal, 33:3, 195-206 https://doi.org/10.1080/14615517.2015.1041837

Additional literature analysed to inform our research

1. Barder, O. (August 2012) 'What is Development?', Center for Global Development Blog. Available at: https://www.cgdev.org/blog/what-development
2. McCreless, M. (Winter 2017) 'Toward the Efficient Impact Frontier', Stanford Social Innovation Review.
3. Organisation for Economic Co-operation and Development (2013) 'Development Results: An Overview of Results Measurement and Management. Briefing Note'.
4. Programme Coordination Unit of the Impact Programme (December 2018) 'Impact Measurement and Management in Impact Investing: review and reflection for DFID staff'.
5. World Business Council for Sustainable Development and International Finance Corporation (July 2008) 'Measuring Impact Framework Methodology'.
6. S Spratt (2008) Development finance: debates, dogmas and new directions Routledge

7. OHCHR (2019) Benchmarking study of the development finance institutions' safeguards and due diligence frameworks against the UN guiding principles on business and human rights
8. JP Morgan (2011) Insight into the Impact Investment Market Social Finance Research Report
9. S Spratt and LR Collins (2012) Development Finance Institutions and Infrastructure: A systematic review of evidence for development additionality Private Infrastructure Development Group
10. J Xu et al (2019) Mapping development finance institutions worldwide: definitions, rationales and varieties Institute of New Structural Economics Peking University
11. S Sprat, P O'Flynn (2018) DFIs and Development impact; an evaluation of Swedfund EBA report 2018:01 Expert Group for Aid Studies, Sweden
12. N Shafik (2011) The future of development finance Center for Global Development Working Paper 250
13. T Adams, M Ripley and A Speyer (undated) At the heart of impact measurement: listening to customers Deep Dive DFID The Impact Programme
14. C Santiso (2004) Development finance, governance and conditionality: politics matter International Public Management Journal 7 (1)
15. C Garmendia and A Olszewski (2014) Impact investing in development finance Impact investing Policy Collaborative
16. Oxford Policy Management (2019) Review of tools used for measuring development impact ex-ante among DFIs and Impact Investors
17. OVE (2017) Comparative Study of Equity Investing in Development Finance Institutions (DFIs) https://publications.iadb.org/publications/english/document/Comparative-Study-of-Equity-Investing-in-Development-Finance-Institutions.pdf ; https://ecgnet.org/about-ecg
18. Savoy C M, Carter P and Lemma A (2016) Development finance institutions come of age: policy engagement, impact, and new directions ODI CSIS Report https://www.odi.org/publications/10591-development-finance-institutions-come-age-policy-engagement-impact-and-new-directions
19. GIIN (2018) Annual Impact Investor Survey https://thegiin.org/research/publication/annualsurvey2018
20. MA Jouanjean and DW te Velde (2013) The role of development finance institutions in promoting jobs and structural transformation a quantitative assessment ODI Working Paper 377
21. I Massa (2011) Impact of multilateral development finance institutions on economic growth ODI Briefing Paper
22. S Spratt (2018) Finnfund impact report
23. F Donou-Adonsou and K Sylwester (2016) Financial development and poverty reduction in developing countries: new evidence from banks and microfinance institutions Review of Development Finance
24. A Dreher, A Fuchs, B Parks, AM Strange, MJ Tierney (2017) Aid, China and Growth: evidence from a new global development finance dataset AIDDATA working paper 46

25. N Marwa (2014) Micro, Small and medium enterprises external financing challenges: The role of formal financial institutions and development finance intervention in Tanzania International Journal of Trade, Economics and Finance 5 (3)
26. N Lee (2020) Eight Principles for DFI Crisis response CDG Note https://www.cgdev.org/sites/default/files/eight-principles-dfi-crisis-response.pdf
27. OVE 2016 Corporate-Evaluation-Contingent-Lending-Instruments https://www.ecgnet.org/document/corporate-evaluation-contingent-lending-instruments
28. Bridges Ventures (2013) 'Bridges Impact Report: a Spotlight on our Methodology'.
29. Commonwealth Development Corporation 'The Development Impact Grid'. Available at: https://d3s6ftg26lsiet.cloudfront.net/wp-content/uploads/2018/07/06125405/Development-Impact-Grid.pdf
30. DEG (2018) 'DEG's Development Effectiveness Rating (DERa)', Presentation in Cologne, 2018.
31. European Bank for Reconstruction and Development (July 2018) 'Transition Impact Methodology Update. Memorandum'.
32. Global Innovation Fund (February 2019) 'Practical Impact: GIF's approach to impact measurement'.
33. International Finance Corporation (March 2018) 'AIMM Scoring Process & Implementation
and Update on New Additionality Framework', Presentation in March 2018.
34. Investment Fund for Developing Countries 'Rating Methodology for DIM'. Available at: https://www.ifu.dk/wp-content/uploads/Rating-methodology-for-DIM.pdf
35. Independent Evaluation Group (2013) Assessing the monitoring and evaluation systems of IFC and MIGA Biennial Report on Operations Evaluation
36. IFC/SecM2020-0025 (03/2020) IFC Quaterly Board Reports FY20 Quarter 2 (July 1,2019 – December 31,2019), IFC Operations Report to the Board (IOR) IFC Quaterly Budget and Human Resources Report (QBHR)
37. World Bank/IFC/MIGA (09/2017) Office Memorandum OM2017-0070 IFC's Anticipated Impact Measurement and Monitoring System: Pilot Framework and Implementation Update
38. World Bank/IFC/MIGA (10/2017) Office Memorandum OM2017-0076 Eds' Seminar – Friday, October 6,2017 The Anticipated Impact Measurement & Monitoring System
39. IFC (03/2018) Committee on Development Effectiveness, IFC AIMM System: Update for CODE2018-0011
40. IFC/SecM2018-0106 (08/2018) AIMM Scoring Methodology Changes
41. IFC/SecM2019-0127 (10/2019) AIMM Scoring Methodology Changes, Second Revision
42. Impact Management Project (2019) Impact monetisation huddle
43. Pacific Community Ventures (2019) Impact Due Diligence: Emerging Best Practices. A synthesis of due diligence practices employed by leading impact investors who systematically assess investments' anticipated impact anticipated impact
44. Impact Frontiers Collaboration (2020) How Investors Can Integrate Social Impact With Financial Performance to Improve Both https://ssir.org/articles/entry/how_investors_can_integrate_social_impact_with_financial_performance_to_improve_both

Annex II - List of Interviewees for Semi-Structured and Key Informant Interviews

Semi Structured Interviewees (Tier 1 DFIs)	Organisation	Position
Ousman Gajido	AfDB	ADOA Team Leader
Roland Michelitsch	AfDB	Evaluator General, Independent Evaluation Office
Khaled Samir, Independent		Evaluation Officer, Independent Evaluation Office
Mouna Ben Dhaou	AfDB	Portfolio Support and Operations, Private Sector Development Dept
Boris Honkpehedji	AfDB	Risk Mitigation Officer, Private Sector Development Dept
Natascha Weisert	AfDB	Senior Advisor, Board of Executive Directors
Nico Westphal	BMZ	Senior Policy Officer
Helmut Fischer	AsDB	Executive Director for Austria, Germany, Luxembourg, Turkey and the United Kingdom, Asian Development Bank
Gloria Paniagua	AsDB	Senior Results Management Specialist in the Private Sector Development Operations
Mark Kunzer	AsDB	Director, Operations Dept, Private Sector Development Department
Mike Barrow	AsDB	Director General and Head of Private Sector Department
Julian Frede	DEG	Senior Manager, Department for Corporate Strategy and Development Policy
Elisabetta Falcetti,	EBRD	Director for Sector Economics and Policy, Economics, Policy and Governance Department;
Pawel Krasny		Principal Economist, Economics, Policy and Governance Department
Astrid Harnisch	BMZ	Senior Policy Officer
Alessandro Maffioli	IADB	Chief, Development Effectiveness Division, IDB Invest
Rodolfo Mario Stucchi	IADB	Head, Development Effectiveness, IDB Invest
Philipp Hauger	WBG	Advisor to the German Executive Director's Office, World Bank
Linnea Kreibohm	BMZ	Senior Policy Officer
Issa Faye	IFC	*Director,* Development Impact Unit
Camilo Mondragon Velez	IFC	Principal Research Economist, modelling team, Development Impact Unit
Luyen Doan Tran		*Adviser*, Economics and Private Sector Development Vice Presidency Unit
Zekebweliwai Fuh Kah Geh	IFC	Adviser, Economics and Private Sector Development Vice Presidency Unit, IFC.
Key Informant Interviews (Impact Investors)	**Organisation**	**Position**
Mike McCreless	IMP	Head of Investor Collaboration
Leslie Labruto,	Acumen	Head of Global Energy
Shami Nissan	Actis	Head of Responsible Investment
Plum Lomax	New Philanthropy Capital	Principal, Impact Investing
David Pritchard	SVT Group	Advisor

Annex III – Private Sector Case Studies

Actis

The reasons for selecting Actis are:

Actis is a signatory to the UN-supported Principles for Responsible Investment (PRI) since 2009 and has reported publicly to the PRI since 2010. In 2015, PRI launched a Reporting and Assessment process to measure signatories ESG activities and reporting frameworks, for which Actis achieved an A rating.

Actis is one of a core group invited by the World Bank's International Finance Corporation (IFC) to provide input to developing Operating Principles for Impact Management, and as a contributor to the Impact Management Project (IMP), a global forum for organisations to build consensus on common standards for disclosure and management.

Links

https://www.act.is/responsible-investing/
https://vimeo.com/328685850
https://www.act.is/media/1841/vdv-report-2018-update.pdf

To Actis

Actis is a leading investor in growth markets across Africa, Asia and Latin America across private equity, energy, infrastructure and real estate asset classes. Born out of CDC and founded in 2004, it has raised US$15bn since inception and employs over 200 people, including a team of c.120 investment professionals, working across 16 offices globally. Their investors' capital is at work in c.70 companies around the world, employing over 116,500 people.

To the Tool

The Impact Score and Multiple has been operating since January 2019 and is applied to all investments. It generates an impact score for each investment allowing for impact returns to be compared in the same way as financial returns. It is also open source.

The Responsible Investment (RI) Team at Actis built its Impact Score Methodology for three reasons: a) investors began to ask questions about value creation and how Actis contributes to the SDGs; b) the DFI community began asking more questions as to how this can be done; and c) while the industry has generated many metrics, principles and frameworks, it lacks a measurement system (SSI).

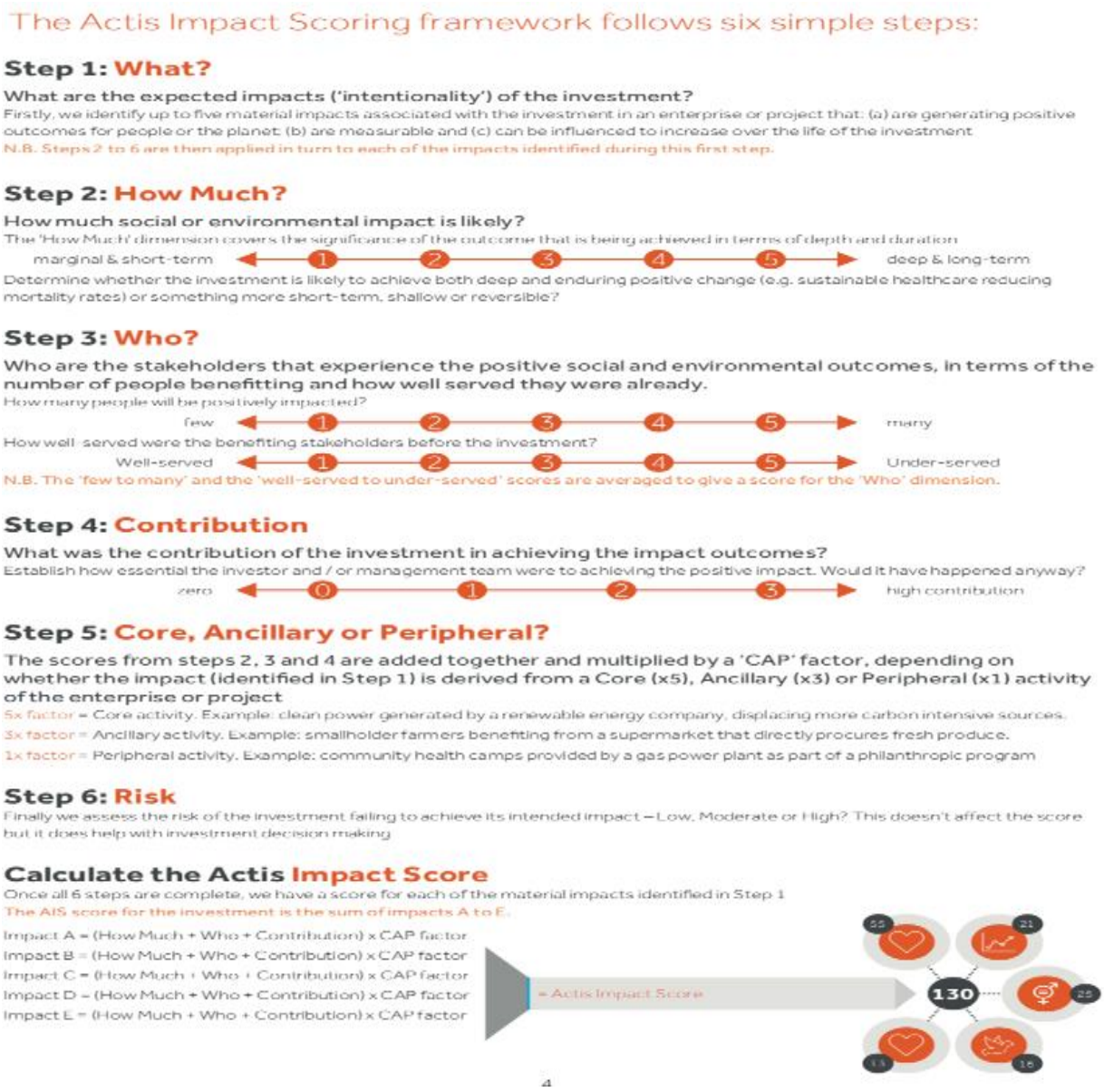
It is developed over three stages:

Figure 20 - Three stages of Actis

Actis integrated its existing ESG framework with a bespoke sliding scale aligned to IFC's impact investing principles, the SDGs and the IMP's five dimensions along a six-step process.

This six-step impact measurement methodology provides an impact score (at appraisal) and an impact multiple (through implementation) that allows the firm to report portfolio companies' overall positive impact along- side financial performance.

Figure 21 - The six-step methodology



Adequacy of Scope

There is no definition of development impact, rather Actis interprets impact as defined by select SDG goals relevant to particular investments.

The What

For each investment, the tool identifies up to five material impacts aligned to specific SDGs that: (a) generates positive outcomes for people or the planet; (b) are measurable; and (c) can be influenced to increase over the life of the investment. These dimensions are prioritised as it is impossible to measure and aggregate all impact, so Actis focuses on areas that are the most transformative and enduring. For example, CO2 avoidance, jobs created, educational improvements (enrolment) and healthcare outcomes (patient outreach).

The Who

For each of these the tool specifies Who are the stakeholders that experience the positive social and environmental outcomes, in terms of the number of people benefitting and how well served they were already. This is assessed on two counts:

1. How many people will be positively impacted? - 1= few, 5 = many

2. How well-served were the benefiting stakeholders before the investment? 1=well served; 5= underserved

Actis compares statistics from the investee, such as employment of women and access to services among its customer base, with national averages. Those on customers are quite general and, in response, Actis is planning to do more work on better understanding the specific client base of their investee companies and the extent to which they are under-served.

The How Much

This dimension covers the significance of the outcome that is being achieved in terms of depth and duration with 1 =5 marginal and short term, 5 = deep and long term. It determines whether the investment is likely to achieve both deep and enduring positive change (e.g. sustainable healthcare and reducing mortality rates) or something more short-term, shallow or reversible?

Targets are defined to help inform the scale, for example on jobs created 5000 would be given a score that helps compare like with like. A metric such as Jobs created is just a data point for a particular investee company. It's how to analyse and compare this metric across the portfolio that the rating system helps you.

The Contribution

What Actis does differently and brings to the investee beyond providing finance is part of its due diligence. The approach keeps this assessment simple by asking whether this investment, with the accompanying non-financial support, would never have happened. It is entirely subjective and based on discussions with the potential investee. This is very difficult to validate (SSI.) For example, if a firm won the rights to develop a solar power plant via a competitive auction, it's fair to say that someone else would have done the same thing, so the contribution would be scored as zero. However, the investment overall might still achieve a high score by introducing a deep and enduring service to a community.

Trade-offs between financial and impact

> *"We don't see a tension between intentionality and returns. Impact does not dilute commercial performance" Interview with Head of Responsible Investment, Actis (Private Equity International 2019)*

Actis assesses deals in the normal way as well as identifying positive impacts it wants to achieve, such as whether the investment will create jobs, improve access and affordability to an essential service or generate clean energy. A strong case can be made that impact can contribute directly to company value and therefore returns. Many of the financial or commercial performance metrics it uses also track impact. For instance, in the education sector metrics such as the number of enrolments, graduation rates, post- graduation employment levels and starting salaries and number of courses offered define financial and operational reporting tools. They also provide a window on how much impact the investment is contributing to student outcomes. A similar comment can be made on investments in healthcare: the number

of clinics opened, number of patients seen, improvements in clinical outcomes – these are all operational and financial metrics that blend with social outcomes (SSI).

The Risk

The tool assesses the risk of the investment in failing to achieve its intended impact and scores this as Low, Moderate or High. While the result does not affect the score, it does help with investment decision-making by revealing the specific risks and how to mitigate these.

The timeframe for investments is determined by the timing of the financial return at exit while most indirect impacts will emerge, or not, long after Actis exits the investment, leaving uncertain the resilience of direct impacts.

The Core, Ancillary or Peripheral impacts

The scores from steps 2, 3 and 4 are added together and multiplied by a 'CAP' factor, depending on whether the core impact is derived from a Core (x5), Ancillary (x3) or Peripheral (x1) activity of the enterprise or project.

For example, a project's core impact could be a renewable energy company displacing carbon emissions (Core), an entity building a supermarket for smallholder farmers that directly procures fresh produce (ancillary) and a gas power plant establishing community health camps as part of a philanthropic programme (ancillary).

Degree of Robustness

There is deliberately no use of a Theory of Change at either portfolio or investment level. Actis is very nervous about carrying out what it sees as an academic exercise to do a full analysis of impact within time periods for private sector.

> *"This is not something a Private Equity Firm does. it provides false comfort given the inherent uncertainty and may, if insisting on it, create tensions in the relationship between the Responsible Investment and sector teams." (SSI)*

The range of indicators, or impact metrics, is limited to quantitative measures informed by those in IRIS + and/or the relevant SDGs. Actis keeps it metrics simple and most, like IRIS+, are easy to use and measure. They are exclusively limited to direct impacts (jobs created, student enrolments, Co2 emissions and numbers of patients).

Figure 22 - Example impact metrics for an education investment

Impact Metric	2017	Outcome
Employees	4,700	96% ↑
Female Employees (FTE)	2,553	113% ↑
Total Students (Enrolment / year)	161,800	350% ↑
Female Students (Enrolment / year)	100,316	390% ↑
Scholarships (# / year)	20,366	350% ↑

The method used to measure these impacts is both conservative and simple. This means the wider or system wide benefits of an investment might not contribute to the overall score. These are rarely predictable. By way of an example, Actis invested in Credit Services Holdings, which

included a company addressing financial inclusion in Uganda by providing credit scores on potential first-time borrowers. In order to do this, the company first of all had to create an identity system in the country. Four years on, this biometric system has become the main form of ID for Ugandan citizens, helping with everything from affordable financial services to voting. Such widespread secondary and tertiary impacts, particularly from first movers in a country or sector, are especially hard to measure empirically and better communicated through a qualitative narrative.

The in-house value creation teamwork with the RI team and do due diligence together on the ground. Community management is essential, and the RI team is usually the first on the ground.

Quality assurance is provided through establishing an ESG sub-committee, consisting of at least two board members, company management and a RI team member. This establishes appropriate senior level scrutiny on ESG issues on a regular basis. This committee also typically oversees the company's long-term community investment strategy – an important mechanism for de-risking projects but also improving social impact.

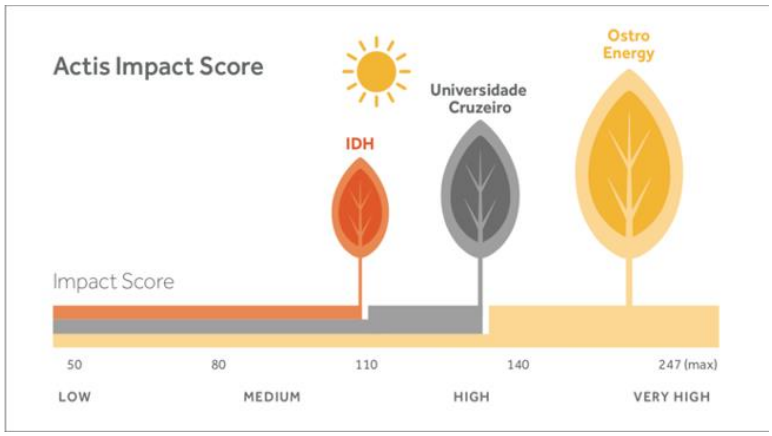
Consultation with the investee is through a due diligence process designed to highlight any gaps in ESG management so that, if an investment is made, Actis knows how to strengthen those areas of weakness. The team spends considerable time with management teams to ensure agreement on what needs to be done. This is key as it ensures there is an alignment of mindset, and it secures senior support for embedding responsible practices across the investee business.

The responsible investment team has an important vantage point across all the Actis companies and sectors. It is highly focused on bringing this experience to bear when working with our companies. To codify our knowledge, we have developed toolkits and frameworks to help 'fast track' ESG improvements so our companies do not have to reinvent the wheel.

Use

Evidence generated by the tool at ex-ante is aggregated into a single impact score for individual investments that can also be presented across different investments.

Figure 23 - Impact Scores across investments



The impact score is not the answer to everything. It is mis-leading to reduce some aspects of impact to a single number, so Actis always includes other elements that need to be expressed through a narrative. For example, and in relation to the above example in Uganda, Actis helped setup the country's first credit bureau. The score reflects predicted numbers of many

customers who will benefit from access to credit and how new credit ratings will lower the cost of borrowing. However, what a single number cannot capture is how the creation of the first credit bureau in the country catalysed a country-wide biometric identification system, which in turn has generated further benefits to Ugandan citizens and the state.

The decision on whether to invest is first and foremost a financial one. Once the commercial deal is agreed, Actis then identifies relevant SDGs and associated indicators.

Actis is a long way from having a consensus on what is an impact threshold. The approach is simple: If there is a commercially viable deal the aim is to then maximise its impact. Actis does not exclude deals based on impact assessment.

Originally, Actis assigned a higher impact score for countries with lower real GDP per capita. This was recently changed as determining a societal need at a country-level can obscure regional pockets of relative poverty and the need for impact investment. Actis assesses this societal need by considering whether stakeholders are well-served or under-served for a particular outcome prior to our investment. As pointed out earlier, the method in understanding the needs of investee clients is work in progress.

While the methodology has a systematic approach, it is refined by sector rather than ‘one size fits all’. There is a deliberate focus on a relatively small number of ESG issues, which are potentially material, and where improvements can be made. This could be improving health and safety standards, environmental performance, business integrity or tackling risks in supply chains. Because the RI team is part of the decision-making process, ESG value creation commitments are embedded into the company’s 200-day plans developed with management, and alongside deal teams, to deliver against these commitments.

Actis installs a head of ESG across the sector teams. Having a dedicated ESG professional in each Actis investments helps ensure that good management principles can endure even when Actis has exited the investment.

The Head of Responsible Investment goes to the investment committee, and the papers submitted to the IC includes a mandatory section on impact. This forces the IC to scrutinise and go through the discipline of to develop a plan, quantify metrics for each investment such as how many jobs created, co2 emissions and the value it creates for the community.

Integration

As with financial evaluations, the Impact Score is measured throughout an investment to see precisely how it is performing from an impact perspective. The current score is compared against the score at the time of initial investment (‘baseline’) to understand how much impact has been added. This is called the impact multiple; the greater the increase in impact, the bigger the multiple.

Similar to the impact score mentioned earlier, assessments of the impact multiple are also presented across investments.

Figure 24 - Comparing Impact Scores with Impact Multiples

Actis Impact Score	IDH	Cruzeiro	Ostro*
A. Latest Score	128	116	147
B. Baseline Score	99	61	26
C. Multiple (A / B)	1.3x	1.9x	5.6x

These results are presented to the board.

Actis has been a signatory to the UN-supported Principles for Responsible Investment (PRI) since 2009 and has reported publicly to the PRI since 2010. In 2015, PRI launched a Reporting and Assessment process to measure signatories' ESG activities and reporting frameworks, for which Actis achieved an A rating.

The tool is open source and its description and examples of its outputs are publicly disclosed and presented on Actis's website.

For all portfolio companies, there is an Impact Sub-committee that reports to the board and who follows through on the results and lessons learnt concerning impact at the community level.

Leapfrog

The reasons for choosing Leapfrog are:

- It has been a pioneer in measurement and management of impact since its founding in January 2007, 8 months prior to the adoption of the term “impact investing”. Throughout its history, the firm has made essential contributions to the metrics and industry bodies that now define the industry.
- LeapFrog was one of the external members of the consultation group that designed the IFC Impact Investing Principles and became a founding signatory to the Principles in April 2019.
- An independent impact verification was completed by Tideline, a leading impact investing consultancy, in September 2019. Tideline concluded: “LeapFrog’s IM system demonstrates an advanced level of alignment across all of IFC’s Impact Principles.”
- Globally, LeapFrog’s was the first independent impact audit ever completed and announced against these agreed principles.
- The audit applies across every LeapFrog fund, rather than to one or another fund in a larger suite of funds not dedicated to impact investing.

Links

https://leapfroginvest.com
https://www.ft.com/content/2236b95e-9998-11e8-88de-49c908b1f264
https://www.businesswire.com/news/home/20190919005589/en/Impact-Investing-LeapFrog-Announces-Audit-Impact-Principles

To Leapfrog

LeapFrog Investments is a profit-with-purpose investor. By backing high-growth, innovative, scalable businesses in Africa and Asia, the company seeks to fulfil the global unmet demand of billions of low-income, emerging-market consumers for critical services. LeapFrog launched its first fund 10 years ago with the goals of generating top-tier, private equity returns. The group currently manages over $1.2 billion in commitments across four funds, reaching 131.4 million emerging consumers with affordable healthcare and finance.

It invests in companies that provide financial tools to millions of low income or financially excluded people across Africa and Asia and in so doing generate top-tier financial and social returns. Its two funds are driven by two inter-related goals:

- To reach 25 million low income or financially excluded people living on less than $10/day with quality, relevant and affordable insurance products while achieving top-tier private equity returns

- To extend fund I’s reach to 50 million low income or financially excluded people with quality, relevant and affordable financial products while achieving top-tier private equity returns.

To the Tool

To drive profit-with-purpose performance, the LeapFrog team have developed a distinctive proprietary measurement framework, which encompasses Financial, Impact, Innovation and Risk Management factors (hereafter FIIRM). It has the following, distinct characteristics:

- Social measurement or KPIs focus on what business can control: Outputs and outcomes vs. Impact
- KPIs aim to integrate financial and social performance, and not trade one for the other
- KPIs are used with a view to drive management decision-making and social performance, not just measure it

The FIIRM system is complemented by an in-house Consumer Insights team, gleaning feedback from customers in-store, by telephone and online, as well as extensive emerging consumer research data sets. It provides insights into customer needs, behaviour, and impact experienced. Leapfrog's design of FIIRM is based on three lessons:

1. Clearly define impact. With a clear vision (on poverty and life choices), it is important to set clear goals around outputs and outcomes fully reflecting the business case for achieving this purpose.
2. KPIs that integrate profit and purpose. By integrating financial and social measures into one collective view of firm value both are seen as intrinsic contributors to building successful business models, and CEOs do not have to choose one at the expense of the other.
3. KPIs for the emerging insurance customer. There is a strong business case for creating social value: products to a vast new emerging consumer segment represents a significant commercial opportunity (Profit) while enabling low income consumers to better mitigate risk and make better investment decisions for their future (purpose).

LeapFrog Investments became the first impact investor globally to announce the results of an independent audit of its impact against the Operating Principles for Impact Management, new industry standards developed by IFC in consultation with leading impact investors and other stakeholders.

Together, FIIRM and Consumer Insights provide a rich dataset of financial and non-financial indicators.

Adequacy of Scope

How is development impact defined and assessed?

Leapfrog defines development impact as poverty reduction. However, FIIRM deliberately does not assess this. Assessing the impact at societal level that measures the impact of one societal intervention relative to a control group without that intervention is problematic. The issue with such an approach for Leapfrog is that measures of societal impact (e.g. reductions in poverty) are not used to drive business behaviour. A business has limited reach and therefore a limited control on impact. While Leapfrog may contribute toward an impact such as poverty reduction or a cleaner environment, a multitude of variables, including consumer choice, means that large scale social deliverables are beyond Leapfrog's exclusive control of any one firm. Social measures unrelated to the core business necessarily, therefore, command less executive attention at Leapfrog. It can distract management from their more natural profit for purpose objective. Leapfrog has found that in order to drive business behaviour, it is necessary to focus on what a company can realistically control and achieve - its outputs and outcomes. Typically, these measures include the values of products produced, their use among clients and how effectively they are used.

How does the definition of impact align with the mandate of the institution?

In this regard, the FIIRM model makes an important distinction between outputs, outcomes and impact. Leapfrog's ultimate social impact vision is to support financially excluded, low-income people in breaking their cycle of poverty. To achieve this, Leapfrog invests in businesses that provide quality, relevant and affordable insurance and other financial services products (outputs) to millions of low-income people, enabling them with critical financial tools to better manage risks (the outcomes). The outputs and outcomes mentioned above are reflected among Leapfrog's objectives and its two associated goals:

The FIIRM framework measures company and fund performance against profit with purpose goals that contribute to the firm's overall objective.

How does the tool define the quantity of impact?

FIIRM is designed to drive business around quantifying outputs and outcomes that are critical to ultimate impact but are directly measurable and aligned with that company's financial bottom line.

All LeapFrog funds have defined dual targets: top-quartile returns (profit) and emerging consumers reached with essential products or services (purpose). These are distilled to the level of each investee company, providing them with a clear measure of success.

How does it mediate trade-offs between different dimensions of impact?

The core tenet of the profit with purpose philosophy is that businesses generating measurable and positive social and environmental outputs are more likely to achieve competitive and sustainable returns. Leapfrog asserts there doesn't need to be a trade-off between social and financial returns.

Does the tool account for risks and assumptions?

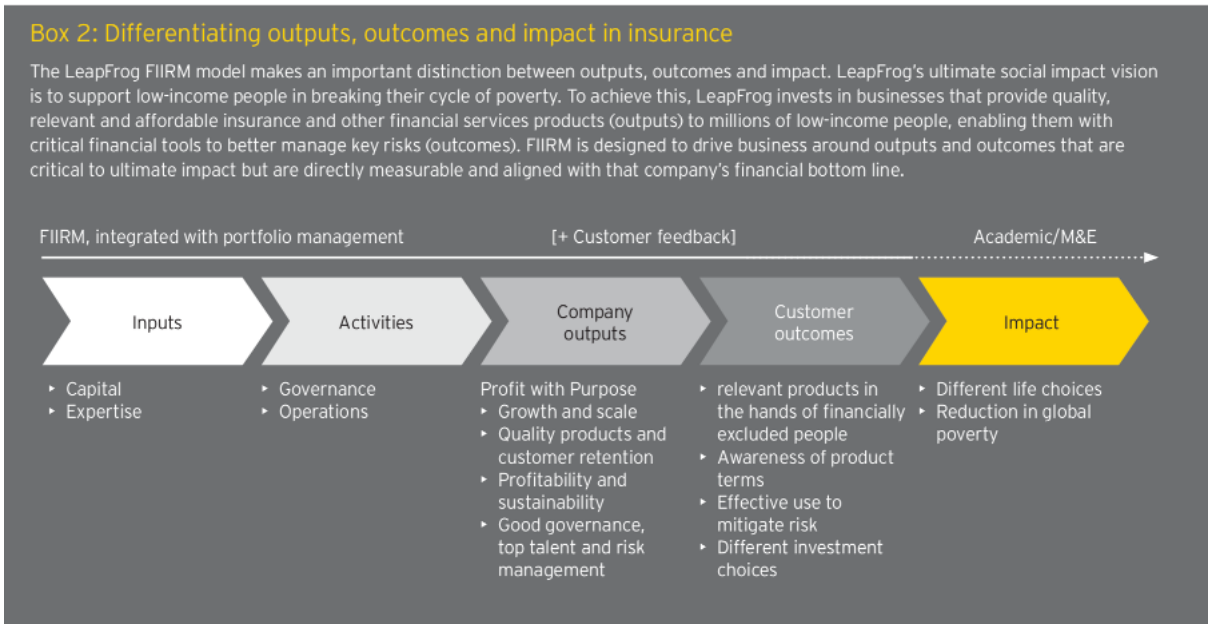
Risk Management is one dimension of FIIRM and data on ESG, Quality Governance and Enterprise Risk Management is collected through investee completing questionnaires, which are then converted to indices and benchmarked with the Dow Jones Sustainability Index.

Degree of Robustness

Is there a clear and logical Theory of Change?

LeapFrog's approach to impact is built upon a clear Theory of Change: by investing capital and expertise (inputs) in innovative companies, LeapFrog aims to equip emerging consumers with essential tools (outputs) that enable better risk mitigation, enhancement of financial and health well-being (outcomes), and that ultimately empower the customer to take entrepreneurial leaps out of poverty as a result of different life choices (impact).

Figure 25 - Leapfrog's Theory of Change



Source: Driving Integrated financial and social returns, EY (2014)

Informed by this over-arching framework, Leapfrog uses the following results chain to frame each business case, or impact thesis.

The Investment team, Impact Labs team, and Customer Centricity team collectively build the impact case. Impact is evaluated across four core areas: (1) scale of emerging consumers reached (current and expected), (2) quality of the product being provided, (3) affordability of the product relative to the emerging consumer income, and (4) the presence of good governance standards to protect the various stakeholders. In determining the emerging consumer opportunity, LeapFrog assesses the financial or health inclusion of low-income consumers in market and the expected provision of products and services by the company to the underserved consumer segment.

What is the range of indicators used?

Leapfrog's approach integrates financial and social measures into one collective view of firm value through KPIs that integrate financial (operational) and social performance. In this way, both are seen as intrinsic contributors to building successful business models, and CEOs do not have to choose one at the expense of the other. KPIs that align financial and social performance are integrated into existing management reporting structures.

Just like operational KPIs, profit with purpose KPIs need to be tailored by industry and sometimes at the company level. IRIS metrics provide a good means to identify indicators of social outputs for a range of businesses across many sectors. It is essential these metrics are aligned with the business' commercial objective. This linkage is at the very heart of integrated reporting, a critical tool for profit with purpose decisions.

All LeapFrog funds have defined dual targets: top-quartile returns (profit) and emerging consumers reached with essential products or services (purpose). These are distilled to the level of each investee company, providing them with a clear measure of success.

FIIRM incorporates measurement of financial and operational Key Performance Indicators as well as governance indices, which are benchmarked to global best practice standards. This enables LeapFrog's investment teams and portfolio company CEOs and CIOs to measure and drive performance towards both profitability and impact objectives.

Financial Performance
Company and Fund financial performance as measured by the investee's revenue, profitability and internal rate of return benchmarked with data from ILPA, IFRS and GAAP.

Impact and Innovation
Non-financial KPIs as measured by scale, quality, innovation, outcomes broken down by type of consumers and service benchmarked with IRIS taxonomy (GIIN).

Consumer Insights

Insights from emerging consumers as measured by Customer social outcomes, experience and loyalty benchmarked with World Bank's Findex.

FIIRM measurement were designed to be standardised across the portfolio and can be supplemented with company specific KPIs also critical to portfolio management.

Central to these are KPIs for the emerging insurance customer. There is a strong business case for creating social value: products to a vast new emerging consumer segment represents a significant commercial opportunity (Profit) while enabling low income consumers to better mitigate risk and make better investment decisions for their future (Purpose). In the same way, leading business indicators go hand in hand with social outputs and outcomes, for example:

- Claims processing - the efficiency of the company in how quickly the company can get cash in the hands of the insured when they need it most; and

- Policy renewal – a leading indicator of profitability where customer retention is critical to small premium policies and an indicator of the value of the product to the low-income consumer.

How is the evidence gathered?

Data collected on the investee's internal performance is supplemented by Consumer Insights. The diligence for impact risks and opportunities is conducted using a standardised Impact and ESG Due Diligence toolkit, in collaboration with a consumer centricity diligence.

Field interviews capture the experience of low-income consumers who are the target beneficiaries, and this generates learnings about consumers' diverse needs and preferences. At the heart of LeapFrog's focus on the emerging consumer opportunity is a deep understanding of the needs and preferences of these consumers in multiple markets. The insights are driven by six key measures that predict the strength of a relationship and future loyalty:

- Satisfaction with products and services
- Value for money
- Comparison with competitors
- Responsiveness
- Propensity to shift
- Trust in the brand

This is done in collaboration with LeapFrog's impact labs. At the heart of Labs' mission is the collaboration of profit with purpose. Labs are demonstrating that catalysing purpose-driven businesses meet need, attract better people, have stronger relations with their customers, are smarter and more profitable, and are thus attractive investment propositions.

Does the tool use harmonised indicators?

HIPSO is not used, and as mentioned above, indicators reflect a mix of harmonised indicators and benchmarks across FIIRMS' three dimensions and the complement, Consumer Insights.

What arrangements are in place for assurance?

FIIRM is quality assured by an independent third party and an assessment in December 2019 concluded that:

- **Areas of strength**: LeapFrog's IM system demonstrates an advanced level of alignment across all of IFC's Impact Principles. LeapFrog has integrated impact considerations in a consistent manner throughout its investment process, supported by clear decision-making protocols and standardized documentation. The process has been refined and strengthened over the course of several funds.

- **Areas for improvement:** There are no notable processes in need of improvement at this time. [28]

On consultation with the investee?

LeapFrog works closely with its partner companies to deepen their understanding of consumer needs. Through human-centered design processes, LeapFrog helps them deliver relevant, affordable and quality products.

FIIRM focuses on a given company's impact goals, defines them in measurable outputs and outcomes (purpose) and integrates purposed KPIs with operational KPIs to reflect the business case for achieving the purpose. FIIRM also includes policy level indices constructed with a formal policy level questionnaire:

- Good governance and policy
- Product quality (client protection)
- Enterprise risk management

Use

How is the tool used in decision-making?

LeapFrog's Investment Committee (IC) scrutinizes the expected impact and the synergy between profit and purpose. Post review, the IC formally signs off on the impact targets (base case, downside case, and upside case). Each Investment Committee paper has a separate section for impact and ESG. This provides the strategy of the underlying investment, the impact gap experienced by low income consumers, the material risks involved in delivering this strategy, and commitment of management to mitigate the risk, including evidence of a clear and viable action plan to increase the impact of the investment.

LeapFrog's framework was designed to integrate impact considerations throughout the investment design process broken down by three stages:

1. Screening—Each investment opportunity is evaluated from the start using FIIRM on key financial and impact considerations, including ESG and sustainability. The results enable clear identification of the profit-with-purpose opportunity and highlight areas of focus for further diligence.

[28] https://leapfroginvest.com/wp-content/uploads/2020/05/LeapFrog_Tideline-verifier-statement_Detailed-assessment.pdf

2. Due diligence—All aspects of the FIIRM framework are applied with due diligence to each investment. Impact considerations are examined, both top-down and bottom-up, by assessing alignment and performance of the company against the four tenets of scale, quality, affordability, and governance. At the same time, the Consumer Insights team collects data from a range of consumers on their unmet demands, "pain points," perceived future risks, and drivers of satisfaction.

3. Investment decision—The investment committee integrates FIIRM results and Customer Insights to holistically evaluate the performance of potential investments. The due diligence results from FIIRM help crystalize company-level impact targets and action plans and enable alignment with LeapFrog's principles for responsible investment.

Integration

To what extent is the basis for approving an investment followed through in M&E?

FIIRM forms the backbone of LeapFrog's impact measurement and management system for achieving integrated performance results. It provides integrated dashboards and highlights gaps between targets and actual performance. These dashboards are used by Leapfrog stakeholders, including deal teams, investors and management to monitor and drive performance.

FIIRM data and indicators are used on a monthly and quarterly basis by LeapFrog's investment governing forums and investment functions to manage the portfolio's performance on profit and purpose. The expected case for impact is built at the time of investment using FIIRM and Consumer insights (quantitative and qualitative) and progress is monitored quarterly.

Tracking of impact is done through quantitative and qualitative KPIs that measure the scale of people reached, quality of products being offered, affordability relative to low-income consumers, and institutionalization of good governance standards.

LeapFrog has predefined reporting guidelines that provide detailed guidance to investment teams and investee businesses. These guidance documents include protocols and templates for reporting performance, the method and responsibilities for data collection, and to whom the data will be reported, as well as state the regular, predetermined intervals at which performance data is reported and reviewed internally. The reporting timelines are agreed at the start of the year. LeapFrog's reporting includes monthly reviews and quarterly deep dives by LeapFrog's investment governing bodies, quarterly reports for LPs, and specific reporting requirements.

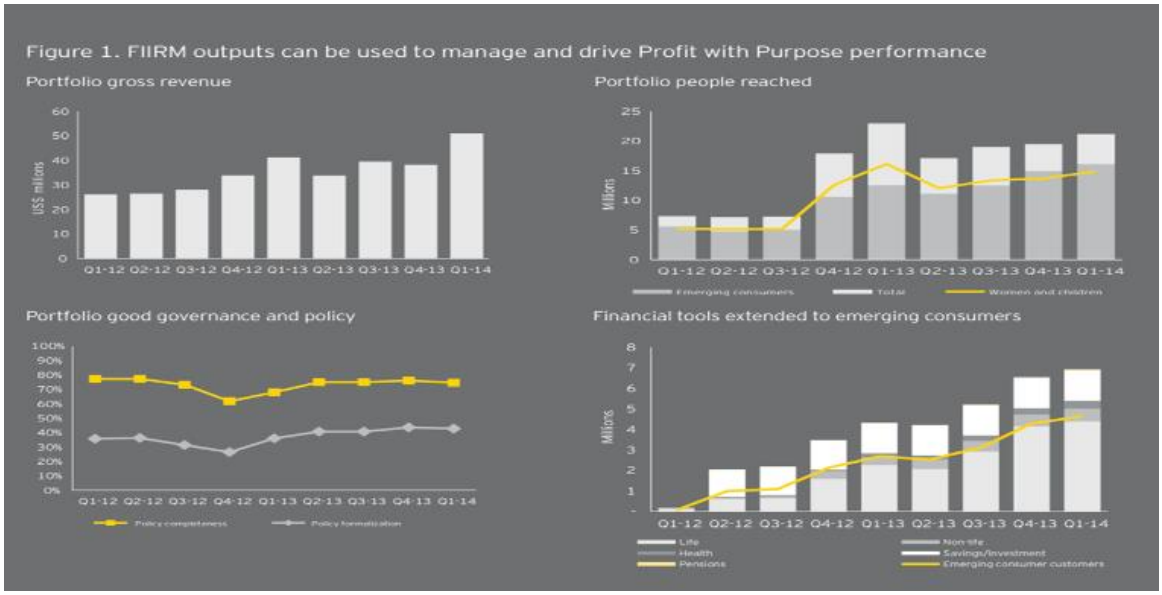
In the event of impact underperformance or other negative effects, LeapFrog enacts Environmental and Social Action Plans (ESAPs) or interventions mandated by LeapFrog's investment governing forums as applicable, which could include the need to update the impact targets in light of the performance.

FIIRM captures outcomes data for each product, which demonstrate the customer's ability to understand, afford and use the product. This is analysed along with direct outcomes feedback from the customers to establish whether the beneficiary is experiencing the expected outcomes and impact.

Investment management—FIIRM forms the backbone of ongoing portfolio review and management. All companies report FIIRM data quarterly, ensuring timely and integrated results. Targeted consumer feedback further supplements FIIRM results, charting the

trajectory for value creation and risk management across a range of financial and non-financial impacts.

Figure 26 - Dashboard of Results



At exit, the data and insights are used to evaluate impact, and financial and ESG performance against initial targets in order to prove decisively the value generated by LeapFrog's profit-with-purpose approach. A Responsible Exits Framework also helps ensure companies graduate to a suitable next owner, while protecting emerging consumers. Public disclosure on adherence to IFC's Operating Principles for Impact Management is available on its website.[29]

[29] https://leapfroginvest.com/wp-content/uploads/2020/05/LeapFrog-OPIM-Disclosure-Statement-May-11-2020.pdf