THE DEVELOPMENT ASSISTANCE COMMITTEE:
ENABLING EFFECTIVE DEVELOPMENT

BLENDED FINANCE EVALUATION: GOVERNANCE AND METHODOLOGICAL CHALLENGES

Ole Winckler Andersen, Irene Basile, Antonie de Kemp, Gunnar Gotz, Erik Lundsgaarde and Magdalena Orth

OECD DEVELOPMENT CO-OPERATION WORKING PAPER 51

Authorised for publication by Jorge Moreira da Silva, Director, Development Co-operation Directorate

OECD
BETTER POLICIES FOR BETTER LIVES

January 2019

Working Paper

Working Papers describe preliminary results or research in progress by the authors and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on the present Working Paper are welcomed and may be sent to dac.contact@oecd.org — the Development Co-operation Directorate, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

Abstract

There is an urgent need to better understand the role that the use of blended finance in development co-operation can play in achieving the Sustainable Development Goals (SDGs). By adopting the "Blended Finance Principles for Unlocking Commercial Finance for the SDGs" in 2017, members of the OECD's Development Assistance Committee committed to "monitor blended finance for transparency and results". The practical implications of monitoring and evaluating blended finance are currently being explored with a view to providing further policy guidance on the implementation of these Principles. This paper contributes to the ongoing consultation process by discussing governance and methodological challenges in blended finance evaluation and proposing a few options to deal with them.

In their review of the present practice of blended finance evaluation, as of end 2018, the authors identify some key issues that need to be addressed and put forth initial ideas in order to ensure that evaluations improve the knowledge base on blended operations. The paper starts with a discussion of some key management and organisational challenges that influence how blending operations are monitored and evaluated. It continues with an overview of main evaluation methodologies that could be used for blended finance evaluation and challenges associated with applying them, and it outlines the challenges of assessing additionality. The paper concludes with a summary of identified issues from a review of a sample of completed evaluations of blended finance, highlighting the methodological challenges that they reveal.

Acknowledgements

The OECD Development Co-operation Directorate would like to thank the external authors and their respective institutions for their commitment and expertise: Ole Winckler Andersen (Danish Institute for International Studies), Antonie de Kemp (Ministry of Foreign Affairs of the Netherlands, Policy and Operations Evaluation Department), Gunnar Gotz (DEval, German Institute for Development Evaluation), Erik Lundsgaarde (Danish Institute for International Studies), Magdalena Orth (DEval, German Institute for Development Evaluation).

Furthermore, the Secretariat would like to thank Government of Denmark for their leadership and support in pursuing this work. In particular, our gratitude goes to the Evaluation Department in the Ministry of Foreign Affairs of Denmark for hosting the OECD Workshop on "Addressing the evidence gap in blended finance" in Copenhagen, 22 October 2018.

The authors are also thankful to all members of the DAC Network on Development Evaluation and of the DAC Results Community who actively contributed to this research.

Finally, the report has benefited from helpful comments from colleagues of the OECD Development Co-operation Directorate and from the Association of bilateral European Development Finance Institutions (EDFI).

1. Background and introduction[1]

The OECD's Development Assistance Committee (DAC) adopted the 'Blended Finance Principles for Unlocking Commercial Finance for the SDGs' at its High-Level Meeting in October 2017. Principle 5 sets out the need to "Monitor blended finance for transparency and results" and highlights various challenges related to monitoring and evaluation of blended finance.[2] In January 2018, the OECD published 'Making Blended Finance Work for the Sustainable Development Goals' (OECD, 2018[1]), which provides more background and explanation for the 'Blended Finance Principles'. Neither the 'Principles' nor the background document addresses the practical implications of Principle 5, and a consultation process has been initiated to lead to the development of guidelines for monitoring and evaluation of blended finance operations. This paper contributes to this consultation process in discussing challenges in blended finance evaluation and by proposing a few options to deal with them.

Multilateral and bilateral development agencies are increasingly turning to blended finance as a means of mobilising private resources to address investment needs in connection with the Sustainable Development Goals.[3] This paper takes as its starting point the OECD definition of blended finance: 'The strategic use of development finance for the mobilisation of additional finance towards sustainable development in developing countries' (OECD, 2018[1]).[4] This broad definition can provide an umbrella for a number of different models of blended finance. Blending can combine financing from varied public and private sources through a mixture of distinct modalities.

Blended finance transactions can include the use of financial instruments to crowd in commercial investments as well as mechanisms to structure or intermediate instruments with the same purpose. According to the DAC Creditor Reporting System, official development finance is provided using five main groups of instruments:

- Grants: transfers in cash and in kind where no legal debt is incurred.
- Debt instruments: transfers in cash and in kind where legal debt is incurred (e.g. loans, bonds and other securities) or could be incurred when certain events occur (e.g. reimbursable grants).
- Equity: a share in the ownership of a company or a collective investment scheme.
- Mezzanine finance: hybrid instruments, such as subordinated loans and preferred equity that present risk profiles between senior loans and equity.

[1] This paper is a revised version of the paper ('Issues Paper on evaluation of blended finance') prepared for the OECD Workshop on "Addressing the evidence gap in blended finance" held in Copenhagen, 22 October 2018.

[2] The other principles are: (1) 'Anchor blended finance use to a development rationale'; (2) 'Design blended finance to increase the mobilisation of commercial finance'; (3) 'Tailor blended finance to local context'; (4) 'Focus on effective partnering for blended finance'.

[3] (OECD, 2018[1]) mentions that 17 members of the OECD's Development Assistance Committee (DAC) now engage in blending, that 167 facilities have been launched 2000-2016, and that $81 billion private finance has been mobilised by development finance 2012-2015.

[4] For an overview and discussion of various definitions, see (OECD, 2018[1]), chapter 3. Compare also with (DFI Working Group, 2018[5]).

- Guarantees/insurance: risk-sharing agreements under which the guarantor agrees to pay to the lender/investor part of or the entire amount due on a loan, equity or other instrument in the event of non-payment by the borrower or loss of value in case of investment.

These can be further structured into more complex mechanisms to mobilise private capital such as funds, syndication, securitisation or PPPs.

Blended finance pursues both development and commercial objectives, underlining its hybrid character, operating between public and private spheres. Similarly, blending may be justified as a response to different types of problems. For example, it may be proposed as a means of addressing market failures (perceived or not) and to improve the risk-return relationship of investment projects (OECD, 2018[1]). The variety of actors, forms of financing, and objectives associated with blending present a challenge in terms of generating an evidence base on its effectiveness, both in relation to the goals associated with blending itself and in identifying the added value of blending in comparison to other development instruments.

This paper reviews the present practice of evaluating blended finance,[5] identifies some key issues that need to be addressed and puts forth initial ideas in order to ensure that evaluations improve the knowledge base on blending finance operations.[6] Thus, the purpose of the paper is not to assess or re-assess the potential relevance and results of blended finance as an instrument.

The paper starts with a discussion of some key management and organisational challenges that influence how blending operations are monitored and evaluated. It continues with an overview of main evaluation methodologies that could be used for blended finance evaluation and challenges associated with applying them, and it outlines the challenges of evaluating the additionality of blended finance. This subject receives special attention due to the centrality of this concept in justifying the use of blending. The paper concludes with a summary of identified issues from a review of a sample of completed evaluations of blended finance, highlighting the methodological challenges that they reveal.

2. Management and organisational challenges[7]

Organisational diversity

One fundamental challenge in establishing a common knowledge base on the effectiveness of blended finance initiatives is the diversity of organisational setups and legal provisions that influence how public actors engage in blending and what influences blending practice. Blended finance involves entities with more diverse legal settings than other development cooperation modalities: public administration, public and commercial banks, pension

[5] Parallel work is ongoing on monitoring and results measurement of private sector engagement in the OECD/DAC Result Community (OECD, 2018[39]).

[6] As stated in (OECD, 2018[1]), "the evidence base on blended finance is still quite limited" and "monitoring and evaluation of blended finance funds and facilities are less developed than for other development co-operation activities".

[7] This section is partly based on (OECD, 2018[1]), chapter 4.

funds, local financial institutions, multinational corporations, micro small and medium enterprises, individual borrowers, etc.

Public institutions that participate in blending include national governments and their development arms. Most countries distinguish between public agencies in charge of development cooperation and private law entities with market-oriented operations.[8] These development finance institutions (DFIs) may also mobilise non-public funds in their own right, whether from foundations, multi-national corporations (MNCs), or other investors. The type of financial instruments each entity is allowed to deploy will depend on their contractual arrangements.

Because development agencies, development banks, DFIs and export credit agencies (ECAs) have different mandates, the choice to entrust the mobilisation of private capital in developing countries to an agency or a DFI has strategic consequences for the whole development cooperation approach. Development agencies and banks tend to have a broad, impact-oriented mandate, whereas bilateral DFIs often focus on supporting domestic companies and generating a return on investment (ROI) for the national government. Export credit agencies also have a strong interest in investing in majority-owned national companies abroad.

Both DFIs and multilateral development banks (MDBs) are, at their core, banks and incentivise their employees accordingly. Within their portfolios, public interest projects driven by policy objectives often coexist with opportunity-driven deals initiated by private sponsors. Account managers are under pressure to meet specific targets for disbursal of funds, generating competition between DFIs to distribute allocated funding and contributing to a lack of an adequate project pipeline. Similarly, DFIs are meant to make investments in environments where traditional investors balk at the investment risk. Despite this, account managers are evaluated as traditional investors (e.g. through deal flow and ROI). When account managers are measured on funds distributed and not on development impact achieved, they might be less motivated to conduct proper ex-ante evaluations of project proposals and to assess development outcomes.

In a blended finance setup, the incentives of the different actors need to be understood from the start, as the motives of a financial intermediary might be very different from those of donors. Such different intervention logics may be difficult to reconcile in the same evaluation framework. Different stakeholders also bring along different information needs, which evaluations should try to address. When documenting results, evaluators should thus not only consider upwards accountability to the taxpayers, but also multi-stakeholder accountability by engaging all partners involved.

CSOs lament the lack of transparency in monitoring development finance due to the concern that without adequate oversight, public and private sector funds may fall prey to corrupt or predatory practices predicated on the rationale of returns and volumes. While many DFIs have adopted provisions to ensure a competitive tendering process for example, as private law entities they are not subject to the more stringent rules of public procurement regulation. This increases the risk that companies may collude with DFI selection committees for favourable results. DFIs disclose varied levels of information on

[8] For example, the UK's Department for International Development (DfID) provides ODA funds both directly and through its bilateral development finance institution (DFI), the CDC Group, which differs from DfID in that it invests with a return on investment in mind. The same is true of the Agence Française de Développement (AFD) with Proparco, and the European Commission (EC) with the European Investment Bank (EIB).

procurement, tendering, contracting and the results of reported grievances but, overall, MDBs are still considered doing a better job at public information disclosure than their bilateral counterparts (Romero, M. J. and Van de Poel, 2014[2]).

Complex governance patterns affect monitoring and evaluation

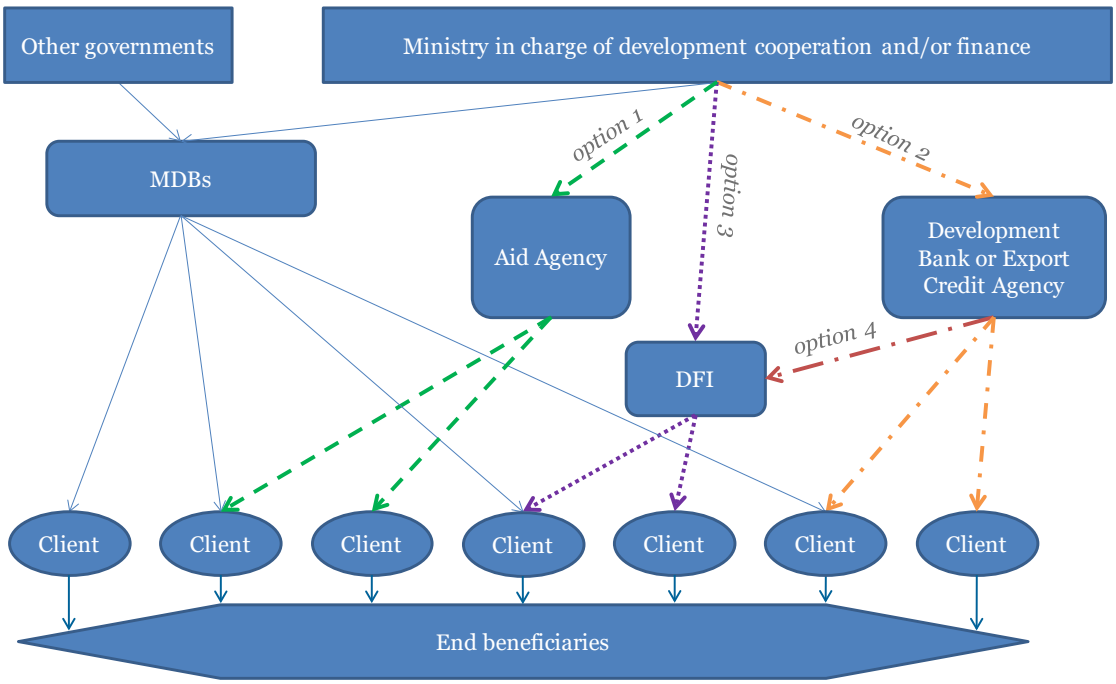
The delegation of development cooperation to specialised agencies carries implications for the availability and quality of information provided on blended finance activities. Several DAC members[9] implement development cooperation via a development bank or aid agency, with no separate DFI (options 1 and 2, as depicted in Figure 1). As the central and only operator, their monitoring and evaluation strategy is validated and closely overseen by the ministry in charge of development cooperation. Private sector mobilisation is thus steered as an integral part of their statutory duties, although separate monitoring and evaluation arrangements might be established ad hoc on specific blended mechanisms. In most DAC members, the DFI or export credit agency (ECA) is directly overseen by the government, but not necessarily by the ministry in charge of foreign affairs (option 3).[10] In these cases, monitoring and evaluation requirements may differ between the two entities, as they are spelled out in separate institutional arrangements. The obligations DFIs and ECAs have to face in terms of reporting and transparency are generally much less strict than for their public counterparts.

[9] Australia, Czech Republic, Japan, Korea, Luxembourg, Slovak Republic, and Slovenia.

[10] Examples include Canada (new-born FinDev Canada), Denmark (IFU), Finland (FINNFUND), the Netherlands (FMO), the UK (CDC Group), but also the European Union (EIB).

A further alternative is that DFIs may coexist with other implementing agencies (options 1+3 in graph 1). Examples include Austria (ADA and ÖEB), Belgium (ENABEL and BIO + SBI-BMI), Italy (AICS and CDP), Norway (NORAD and NORFUND), Portugal (IPAD and SOFID), Sweden (SIDA and SWEDFUND), Switzerland (SDC and SIFEM), Spain (AECID and COFIDES), and the US (USAID and OPIC).

Figure 1. Governance patterns



In a minority of cases (option 4 in Figure 1), the DFI is a subsidiary of a national development bank or export credit agency (France with AFD and Proparco, Germany with KfW and DEG). There, the steering of the monitoring and evaluation framework is fully entrusted to the main executive entity, which acts as an intermediary between the DFI and the government. In this latter case, where DFIs are far removed from policy making, ministries have little oversight or direct communication with the DFI. Thus, they might lose track of what information is available and how DFI activities actually contribute to the overall development cooperation strategy. The involvement of many intermediaries such as fund managers in the implementation of blended finance presents a further complication in promoting transparency.

A general conclusion is that as the delivery chain grows longer, it becomes more difficult for a government to exercise its steering and control functions over its private sector arm(s). Ministries will typically hold close working relationships with their development banks and/or aid agencies, but DFIs and ECAs follow a different logic. Due to their mandate and design, they both respond to the private sector mode of functioning and objectives. To ensure alignment across their institutional frameworks, DFIs typically agree evaluation arrangements with donors in blended finance delegation agreements. MDBs and larger DFIs have their own evaluation teams, while smaller DFIs rely on external experts; all being subject to their parent or shareholding institutions' independent evaluation requirements. Moreover, the DFI's sensitivity to public accountability will depend on their ownership, which can be either fully in the hands of the central government, delegated to another governmental entity, or partially taken over by private shareholders (as with most European DFIs).

Especially in the case of a three-tiered model, ministries have little or no information on the monitoring evidence produced by their DFIs or ECAs. Accountability is even lower when blending is entrusted to a private and purely commercial actor with limited reporting

requirements (often self-imposed, as in the case of impact investors). Thus, donor countries are today unable to draw an overall picture of their blended finance operations, let alone the outcomes or impact of all their development cooperation activities, including private sector mobilisation.

Financial information on blended operations is not systematically disclosed, on grounds of confidentiality. The use of concessionality represents commercially sensitive information. Some believe that donors should invalidate this reasoning, since real trade secrets are not written in contractual agreements. In addition, donors manifest their difficulty in tracking financial contributions to collective investment vehicles (CIVs) and matching them with the reported outflows. Transparency on the use of public subsidies is highly demanded in order to avoid the risk of commercial capture, but the influence donors might exert in blended finance decreases along the delivery chain.

A completely decentralised monitoring system may increase the risk of fragmentation and poor data quality. Although more than 60% of blended finance funds and facilities surveyed by the OECD[11] have identified a team or unit responsible for monitoring and evaluation, this was most often part of line management rather than being independent and reporting directly to the Board of Directors (or equivalent). The survey confirmed that blended CIVs managed directly by donors, rather than by their DFIs or private intermediaries, will more likely benefit from clearly identified resources for monitoring and evaluation. Nonetheless, the more impact-oriented investors typically identify a person in charge of monitoring environmental, social, governance (ESG) aspects of the investments and impact reporting. In a minority of cases (13%), there is no internalised monitoring and evaluation function, but the responsibility lies with each investment or project manager.

The institutionalisation of the monitoring and evaluation function depends heavily on the governance system. As such, blended finance instruments led by multi or bilateral public investors can usually rely on a pre-existing monitoring and evaluation unit, whereas private managers are less likely to have such in-house competence. All MDBs and DFIs have an internal, but not necessarily independent, team in charge of monitoring and evaluation, which ensures compliance with the monitoring requirements and mandates evaluations as part of the wider institution's learning approach. Bilateral DFIs are progressively gearing up their results measurement capacity, but resources for in-depth evaluations are not yet a natural part of the investment cycle in the same way as financial audits. In some governments, e.g. the UK and the Netherlands, the branch in charge of development cooperation has established contractual arrangements on a joint evaluation strategy with their blending arms, even when these are not under their direct supervision.

The lack of harmonisation in monitoring tools hampers comparability

Recent studies have highlighted the many differences regarding monitoring and reporting processes among DFIs, thereby making data comparisons difficult (Lemma, 2015[3]), and several efforts have already been undertaken to harmonise the monitoring and evaluation practices of development finance providers.

To promote consistency among MDBs, the Evaluation Cooperation Group (ECG) defined 'Good-Practice Standards for the Evaluation of Private Sector Investment Operations' in

[11] In 2017, the OECD surveyed 167 facilities and 189 funds. The full list can be found in Annex B in (OECD, 2018[1]).

2001 and revised it four times since (ECG, 2006[4]). In 2013, a total of 25 IFIs agreed to work towards harmonising their development result indicators to reduce variations in data and the reporting burden on clients, and to facilitate learning.[12] Attempts at deeper coordination have also been registered at the regional level. For instance, the Mutual Reliance Initiative, agreed between EIB, AFD and KfW, allows the banks to use each other's supervision and monitoring systems to reduce the administrative burden on partner countries. Furthermore, the DFI's enhanced principles on blended concessional finance in private sector projects (DFI Working Group, 2018[5]) refer to promoting high standards (principle 5). In practice, with regards to evaluation, this means that DFIs/MDBs apply the same standards to projects involving blended concessional finance as to other operations.

Private managers have also acknowledged the value of standardised metrics. As an example, IRIS,[13] an initiative managed by the Global Impact Investing Network, proposes a free online taxonomy on how to measure the social and environmental performance of investments. These definitions underpin the Global Impact Investment Rating System (GIIRS), an impact rating tool that assesses companies and funds based on their social and environmental performance. Private managers often equate 'impact reporting' with evaluation in spite of the difference between the two concepts.[14] While 19 out of 24 private managers surveyed by the OECD claimed having performed at least one evaluation, only three reports could be located online and donors mostly initiated them.

Monitoring schemes should capture information in a traceable, versatile and, where possible, quantifiable manner, while meeting the capabilities at the investee level, if necessary through technical assistance. Concessional resources are thus crucial to ensuring evaluability, since capacity building for small clients is often needed to produce baseline information. Declarations from implementing partners may raise trust issues but are still the preferred source, for reasons of efficiency and as a way to stimulate ownership. Hence, the quality of data available largely depends on the nature, capacity and commitment of the investees.

The interest of a broad alignment (at least, on a set of core indicators) is widely recognised by development practitioners in order to ensure the consistency and comparability of data produced, whilst preserving the opportunity for elective adaptations depending on the instrument's specificities. Analysis at an aggregated level (e.g. institutional or fund level) remains crucial in proving effectiveness to public and private investors. However, asset managers face a recurring trade-off between the granularity and precision of their results-based management tools versus comparability of information at industry or sector level.

Ultimately, the usefulness of such efforts hinges on the availability of standardised data, the quality of data, and on the willingness of blending institutions to publicly engage in such comparisons.

3. Key terms, evaluation standards, criteria and methodological approaches

A number of key terms, evaluation standards and criteria are often used as references when organising and conducting development evaluations (OECD, 1991[6]; OECD, 2002[7];

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

[13] https://iris.thegiin.org/

[14] See (OECD, 2002[7]) for a definition of evaluation.

OECD, 2010[8]). This provides for a common language, which can be useful when assessing and comparing evaluations independent of the type of instruments analysed. While evaluations of blended finance could benefit from applying a common language, the development and use of such a language relevant for blended finance evaluations is still a work in progress.[15]

Objectives and evaluation design

Blended finance evaluations pursue different objectives. For example, they may seek to shape investments in order to maximise impact, to guide future investment strategies, to assess the developmental impact of the intervention[16] or to ensure the efficient use of public resources. The evaluations may also consider whether the selected financial instrument and mode of financing were the most relevant in the given context. In the next section these issues will be discussed further.

In conceptualising such evaluations, it is important to tailor the evaluation design to the different types of blended finance instruments, being either debt, equity or technical assistance. The reason is that outputs, outcomes and impacts cannot be analysed without considering the inputs used to achieve these outcomes. For example, evaluation scope and judgement criteria need to be adjusted depending on the timeframe that different types of blended finance instruments need in order to achieve commercial sustainability.

Methods

On the methodological side, an integrated approach (Goertz, 2017[9]) is needed to reflect the complexity of blended initiatives involving a diversity of contexts, objectives, instruments and actors. The decision on the most suitable methods and combination thereof – which could also include various quantitative and qualitative tools - should obviously be based on the evaluation questions. An assessment of the alignment of blended finance interventions with international or national development strategies will, for instance, require other methods than a systematic assessment of the additionality and impact of blended finance funds and facilities, where a theory-based approach would usually be recommended. A theory-based approach is a way to structure an evaluation, without ruling out the use of any particular evaluation method (White, 2009[10]; Stern et al., 2012[11]; White and Phillips, 2012[12]). In such an approach, a theory of change needs to be formulated to outline how the blended finance intervention intends to reach its results and to assess "the assumptions underlying the evaluated intervention's causal chain from inputs to outcomes and impact" (White, 2009[10]).[17]

A blended finance intervention offering financial support to microfinance institutions would, for example, consist of loans and capacity development as inputs and the loans granted to MSMEs as outputs. One expected outcome may be that population groups gain

[15] An introduction to the discussion of various concepts is provided in Koenig and Jackson (2016) and (OECD, 2016[26]).

[16] In particular, the evaluation of the impact of blended finance has led to discussion. For an analysis of similarities and differences between impact measurement and evaluation, see for example (Vo and Christie, 2018[40]). For a proposal on an impact measurement standard, see (Ruff and Olsen, 2018[41]).

[17] Theory-based evaluation has been the subject of a comprehensive literature, which will not be reviewed here.

access to the financial sector or have additional funding at their disposal. Other outcomes could be new products for the target group and financial sector development. At impact level, increased employment or economic growth are expected effects. Only a theory-basis in the evaluation of blended finance allows to answer how and why the observed results have been achieved.

Analysing cross-case causal inference in blended finance evaluations correctly is challenging because it would be necessary to compare a situation with and without a blended finance fund or facility in place - ideally in the same context. Such comparisons are difficult to undertake. Since blended finance can have a broader impact (e.g. through the development of financial markets) and funds are not allocated randomly but based on criteria such as credit ratings and the availability of fundable projects, it is challenging to create a valid control group. Hence, experimental designs or randomised control trials (RCT) that intend to minimise selection bias are often not feasible due to the nature of activities being implemented (Carter, P, Van de Sijpe, N. and Calel, 2018[13]).

If treatment and control groups of blended finance interventions cannot be selected randomly - e.g. DFIs decide where to open a branch and whether to grant a loan, end beneficiaries purposefully decide to apply for a loan - quasi-experimental designs can offer ways to artificially construct a control group, often ex-post (Linden and Adams, 2012[14]). The constructed control group allows determining what a situation would have looked like without the blended finance intervention, thereby determining the intervention's effect. Compared to RCTs, quasi-experimental methods have the advantage that they can be carried out while an intervention takes place or after an intervention has been implemented. However, in a quasi-experiment, the control and treatment groups differ, not only in terms of the treatment they receive, but also in other ways, and researchers must try to statistically control for as many of these differences as possible. Yet the risk remains that unknown variables causing the difference between both groups escape the researcher's attention, leading to biased results (Deaton, 2009[15]; Banerjee, 2011[16]).

In the case of blended finance evaluation, the construction of a control group is particularly challenging when all groups are already financed through different funding sources. Therefore, the impact of blended finance is often not about access to financial means, but about additional sources of finance.

In case a control group cannot be constructed, or ideally as an addition to a quasi-experimental design, evaluators need to employ qualitative approaches allowing for causal inference.[18] One possibility is comparative case studies, e.g. using qualitative comparative analysis (QCA) (Befani, 2016[17]).

Comparative case studies are particularly suitable for blended finance interventions that are performed across different contexts and for which "how" and "why" questions should be answered (Goodrick, 2014[18]). Such qualitative methods can help to identify new or previously missing variables and thus avoid omitting variables in the formulation of hypotheses. This identification of variables can also help to draw conclusions about the underlying causal mechanisms (Bennett, 2004[19]; Stern, 2015[20]). Comparative case studies can be relevant in assessing the determinants of effectiveness across different projects financed through blending, such as the influence on outcomes by government policies and markets or the influence on accompanying technical assistance. Further, they

[18] See (White and Phillips, 2012[12]) for an overview.

are relevant in analysing variations in performance among blended finance approaches and other forms of development financing.

For within-case causal inference, case study analysis using contribution analysis and process tracing can also serve to uncover causal relationships (Befani, B. and Mayne, 2014[21]), such as the impact of blended finance programmes on the development of markets. Unlike experimental designs, these qualitative approaches do not rely on counterfactual causation, but analyse the (intervening) causal mechanism or process that generates a certain outcome. Sometimes this is referred to as generative causation (Stern, 2015[20]). When relying on a solid theory of change and a mixed method approach, some authors suggest that contribution analysis may be "stretched" to estimate the size of effects in a quantitative manner (Ton et al., 2019[22]). Process tracing essentially explores a series of interlocking events or facts (mechanism parts) that together can explain the outcome. The researcher's task is to verify each of these mechanism parts. Advocates of such approaches argue that the analysis of causal mechanisms can provide a better understanding of the inner workings of complex programmes, such as blended finance, and might be more suitable in contexts with small sample size (e.g. comparing results between a few countries) than experiments.

These approaches are related to realist evaluation approaches, which have as their aim to investigate 'what works, for whom in which circumstance.'[19] As noted later in the paper, many evaluations of blended finance refer to contribution analysis in the discussion of their methodological approach.

Specifically to assess effects of blended finance interventions at system level, such as general-equilibrium effects, statistical modelling can further be used in blended finance evaluation. Input-output modelling is, for example, suggested to estimate the impact of investments on economic transformation (e.g. effects on employment and economic growth) based on firm-level data. (Lemma, 2018[23]) proposes a set of 13 indicators to assess ex ante the transformation potential of DFI investments. At the same time, other methods – such as those introduced above - are needed to measure intended and unintended impacts on level of beneficiaries. As different approaches and methods can and should be applied in blended finance evaluations, evaluators need to be transparent about the nature of the evidence on which conclusions are based.

4. Additionality in blended finance evaluation

Definitions

The OECD definition (see section 1 of this paper) requires that blended finance mobilises additional finance, and that the mobilised funds are used for sustainable development. The first requirement is often referred to as financial additionality and the second as development additionality.[20] The concept of additionality is the core of the mandate of most

[19] The realist methodology studies what the mechanisms of an intervention are like (what makes an intervention work?), the context in which these may be triggered (what are the conditions in which an intervention is effective?) and how these mechanisms bring about certain outcomes. See e.g. (Pawson and Manzano-Santaella, 2012[42]) for an introduction to various perspectives on realist approaches to evaluation.

[20] Development additionality is defined as "development results that could not have been achieved without the mobilisation of commercial capital" (OECD, 2018[1]).

DFIs and other blending entities, and there have been several attempts to define different aspects of additionality (Sjö and Flygare, 2008[24]; Heinrich, 2014[25]; OECD, 2016[26]; EBRD, 2018[27]). Clarifying the way these concepts are defined and assessed should therefore be a key element of efforts to improve the monitoring and evaluation of blended finance.

Financial additionality refers to situations, where finance is mobilised and an investment is made that would not have materialised otherwise,[21] while development additionality refers to the outcome and impact of the investment that goes beyond what would have been achieved in the absence of additional finance (e.g. job creation, protected environment). Development additionality can be achieved not only through financial contributions, but also through non-financial elements (e.g. technical assistance, transfer of technology, management, etc.).

In praxis, the two concepts are closely related. It can hardly be justified to use public funds for a programme or a project that does not have a reasonable chance of delivering development additionality, nor would it be justifiable to use public funds for investments that the private sector would have made in any case. Financial additionality is also not a guarantee for development additionality. Only the presence or likely presence of both forms of additionality can justify the use of blended finance instruments (see Table 1).

Table 1. Blended finance and financial and developmental additionality

		Developmental additionality	
		Yes	No
Financial additionality	Yes	Blended finance should be used	Public sector should not engage
	No	Private finance should be used/encouraged	Public sector should not engage

The definitions for blended finance and additionality vary across stakeholders. The focus of the DFI blended finance guidelines is on blending concessional funds with non-concessional resources to support private sector operations (DFI Working Group, 2018[5]). The OECD approach to blended finance is somewhat different as it focuses on combining financing sources that have a developmental mandate with those that have a commercial mandate, while the investment could be either private or public (OECD, 2018[1]). Hence, what the OECD defines as blended finance would be classified as a form of commercial mobilisation by development finance providers, i.e. attracting commercial actors that otherwise would demonstrably not invest in a sustainable manner in developing countries.

The definitions and approaches to prove additionality diverge between the DFIs. For instance, the Evaluation Division of the EIB considers an intervention to be additional if it facilitates or enhances a project from the public welfare (development) perspective in a way that the market alone would not allow, or at least not to the same extent, or in the same timeframe. This implies (1) identifying the market failure problem the investment seeks to address, (2) qualifying the EIB's influence on the project design (i.e. the extent to which the investment addresses the market failure), (3) verifying that private sector actors are not

[21] As noted by (Carter, P, Van de Sijpe, N. and Calel, 2018[13]), this definition requires that the offered finance does not crowd-out private investors.

crowded out. The EBRD uses a slightly different approach to ensure additionality. Additionality is one of the three core operating principles of the EBRD (together with sound banking and transition impact) and has since extended from a primarily financial concept to include non-financial aspects, such as lower risk or improved quality, as well. The EBRD defines additionality as an input that the Bank contributes to each project, which is expected to foster development impact through its influence on project design and implementation. Additionality is assessed at an early stage, based on the following dimensions: (1) providing financing at reasonable terms (e.g. appropriately priced); (2) sharing specific attributes (e.g. dialogue with government); and (3) structuring conditionalities (e.g. adherence to ESG standards). Nonetheless, many view the operationalisation of additionality at project level as the least robust out of the three key operating principles of the Bank. The presentation of non-financial additionality is seen as overlapping with transition impact. The lack of an identified functional responsibility at the organisational level shows that additionality has not received sustained or coordinated focus, also limiting the degree of internal accountability. Consequently, the (EBRD, 2018[27]) concludes that "additionality is not a definition against which performance can be rigorously measured - assessment of additionality is a judgment call".

Ultimately, blending should be resorted to when it is recognised as the preferable way to fund a specific public problem. Development finance providers must ensure that allocating scarce concessional resources for blending does not draw them away from other types of intervention that may be more effective in a given circumstance. The level of perceived risk (e.g. when moving towards investments in fragile contexts) is intrinsically linked to the concessionality needed, as entering frontier markets requires additional incentives and risk mitigation to attract private sector investments. Hence, blended finance might require large amounts of concessional funding in such contexts, which increases the likelihood of drawing funds away from other types of intervention.

Three key challenges

The above definitions lead to three key challenges that evaluators of blended finance are likely to face: determining mobilisation, proving additionality and showing efficiency. These challenges will be discussed in the following. Additional challenges, which will not be discussed here include attempts to assess the indirect effects, as required in the DAC definition of the evaluation criterion for impact (OECD, 1991[6]),[22] and the time horizon of both direct and indirect effects.

First, evaluators have to assess whether and to which degree additional finance was mobilised. The OECD has developed a methodology to measure the amounts of private finance mobilised by Official Development Finance (Benn, Sangaré and Hos, 2017[28]). Showing causality remains a critical issue, as the methodology assumes that the private investor would not have engaged in the blended finance activity without the involvement of the public sector. This assumption may potentially lead to overestimation of the mobilised amount (OECD, 2018[29]). In practice, mobilisation is often assumed rather than observed. The difficulty of establishing causality by quantitative methods has led to recommendations of using other methods, like process tracing (see section 3 of this paper) or more pragmatic approaches. Several practical suggestions have been made in the literature to assess financial additionality. Examples of criteria for the ex-ante assessment

[22] A proposed method for assessing the indirect effects is provided in (MacGillivray, A., Kim, R., van Moorsel, T. and Kehoe, 2017[43]).

of additionality include company level factors such as insufficient funds to self-finance the project, the lack of knowledge or competencies to design and implement a poverty-reducing business model, and the lack of duplication with other possible forms of donor support (Heinrich, 2014[25]). Ex-post assessments of additionality may use criteria related to the context in which financing is provided (for example, whether foreign direct investment in a given country or sector is small and commercial banks are not active) or criteria related to the characteristics of the venture, such as the access of participating entities to finance (IOB, 2009[30]; Escalante, D., Abramskiehn, D., Hallmeyer, K. and Brown, 2018[31]). (Carter, P, Van de Sijpe, N. and Calel, 2018[13]) shows that neither quantitative nor qualitative methods are capable of delivering reliable results for the evidence of additionality. Consequently, they propose a probabilistic approach to additionality that is less focused on 'measuring' additionality but rather asks under what circumstances it is more likely to exist. Blended instruments are by definition partnerships where each player wants to claim their own additionality, thereby generating a fictitious competition and, possibly, duplication of efforts amongst partners. Instead, all partners should ideally assess additionality jointly, looking at the blended mechanism holistically. This would obviously require a common understanding and definition of additionality for all partners.

Second, evaluators have to assess whether development additionality was achieved. This is close to a traditional impact evaluation, where some of the methods mentioned above, or a combination of them, may be considered. Evaluators have to isolate the effects of blended finance towards the defined purpose of the intervention. Examples of such purposes include the contribution to the attainment of SDGs, job creation, improvement in product quality, or the scaling up or geographical expansion of activities. The financial rate of return is sometimes referred to as an indicator of the success of an intervention, but it does not provide much information about development additionality or 'what worked' and 'why'. Evaluations of blended finance instruments are often more complex than a project evaluation, because they typically have to cover both programme and project level of a blended finance instrument (e.g. a comprehensive evaluation of a fund requires to evaluate the overall portfolio and the operations it finances). Even if monitoring data is available (including baseline and end-line data), such 'before-and-after' assessments may not provide the necessary information about the effects of an intervention because a causal link is hard to establish, especially when targeting economic transformation or market creation. Since most development finance providers are mandated to transform sectors by promoting new business models in nascent or failed markets, it is insufficient to assess success only at the project level. When subsidising a private actor, donors and investors should also study the effects at system level. While the direct and unexpected impacts can only be ascertained at the level of the end beneficiaries, general equilibrium effects should be included in addition in order to verify if economic transformation occurred.

Third, while the first two challenges relate to the effectiveness and impact of interventions, evaluators could be asked to assess whether the intervention was efficient, including in comparison to potential alternatives.[23] In light of the various blending instruments and the incentives they provide for private sector involvement, this is obviously an important question. The main arguments for public intervention are that there are additional social returns beyond the expected private returns or that the private sector overstates the risks or

[23] See (OECD, 1991[6]) for a definition of the DAC evaluation criterion on efficiency, which requires that it is assessed whether the (OECD, 2002[7]) programme or project is "implemented in the most efficient way compared to alternatives".

underestimates the private returns of the interventions (Barder, O. and Talbot, 2015[32]). As Barder and Talbot note, there is no case for blending (or public intervention more in general) if the combination of risks and returns is competitive to alternative investment opportunities. In this case, blending would just be a subsidy and lead to extra profit for the private investors and/or crowd out potential private investors. There would be no financial additionality. The relevance of individual blending instruments depends on the problem to be addressed, and the applied instruments should reflect the specific relationship between expected returns and risks, which depend on the context and vary between sectors. If the aim is to reduce the perceived (high) risks for private investors, this could be achieved by providing information, participating in the management of the project, providing guarantees, providing equity, or participating through subordinated loans. Concessional finance may also raise private investors' expected returns by lowering the project costs. Institutions often report the catalytic impact or leverage effect as proof of additionality, with higher leverage rates pointing to higher effectiveness. However, there is a need to be careful with such an interpretation: a high leverage effect can suggest that a project is already highly attractive for the private sector, giving reason to question the degree to which the blending mechanism is additional. There is, nevertheless, a possibility that the role of the financial institution in the management of the programme or project sufficiently lowered the (perceived) risks for private investors. The evaluation of the relevance and efficiency of the applied blending instrument will thus have to be based on a clear identification and assessment of the problem which blending was meant to solve in the specific case.

In addition, there is no aggregate monitoring and reporting on additionality so far. In this respect, a harmonised approach that is applied by all DFIs would be necessary (EBRD, 2018[27]). DFIs and other actors should include monitoring data that allows to verify the additionality, e.g. through a monitoring of interest rates on commercial loans, risk ratings or market developments. Such data would allow continuously assessing the additionality of the investments and, if necessary, lead to a reduction in concessional funding, which could free up funds to be used elsewhere. Given that the additionality assessment is usually based on a binary judgement, i.e. on yes or no, there is a lack of relative values and thus no possibility of appreciating the scale of a project's additionality. In this regard, a more comprehensive framework in assessing financial additionality would allow for comparing different projects with each other.

5. A brief review of applied evaluation approaches to blended finance

This section provides a summary of key challenges emerging from a review of selected evaluations and some other studies which address the effects of blended finance interventions. The purpose of the review was not to assess the quality of individual evaluations or the potential effects of blended finance interventions, but to make a review of applied evaluation approaches to blended finance and to identify some key challenges.[24] A purposive sample[25] of completed evaluations and relevant studies was covered by the

[24] Note that some of the identified key challenges are not specific for evaluations of blended finance but can also be found in evaluations of other types of interventions.

[25] A list of the reviewed evaluations and other studies is annexed to this paper. It is generally difficult to establish an overview of existing evaluations of blended finance. The OECD DAC resource centre (DEREC) contains some evaluations (see http://www.oecd.org/derec/), but not all evaluations are

review, which focused on the evaluations' (i) purpose and scope; (ii) approach and methodology; (iii) assessment of the monitoring and evaluation system.[26] The review was only based on evaluation reports and the reports of the selected studies. Background documents, including Terms of Reference, were not consulted and no interviews were made. Due to the tentative nature of the findings, the review does not lead to specific recommendations. The review of the evaluations and the studies is summarised in the following.

Purpose and scope

Evaluations referenced existing evaluation terminology, standards and criteria to varying degrees and differed in the use of concepts and definitions. Some evaluations referred to the DAC's evaluation criteria,[27] while other evaluations applied evaluation criteria that seem to have been developed in the individual evaluations. In some cases, this implied a mixture of DAC evaluation criteria and other criteria. Several evaluations did not make a clear distinction between concepts of 'outcome' and 'impact'. The two concepts were often used interchangeably,[28] and 'impact' was used in different ways. This had implications for the approaches of the evaluations and made comparisons difficult. The review of the evaluations also showed that there was not a common understanding of various dimensions of additionality.

In most evaluations, the purpose (accountability and/or learning and for whom) of the evaluations was not clearly defined. This probably reflects the complex management and organisational structure of most blended finance arrangements, where different actors may have different perspectives and interests in the evaluations. The evaluations reviewed focused on internal accountability rather than on generating broader lessons on the utility of various approaches to blended finance. Thus, the focus was only to a limited extent on learning and 'what worked in which context'.

The evaluations reflected the different purposes of the interventions. Several evaluations focused on financial returns and many on job creation. Other evaluations had a broader perspective, which included assessments of various elements such as improvements in ESG indicators, technology transfers and demonstration effects. Some evaluations also assessed the effects beyond the primary target group and considered potential effects at sector level or for specific value chains.[29] In some cases, assessments of broader effects comprised a potential strengthening of local financial markets (European Commission, 2016[33]; European Investment Bank, 2017[34]) or institutional and policy change (European Commission, 2016[33]). Only a few of the evaluations considered, however, the potential crowding-out of other actors due to the use of blending or the coherence with other

shared on DEREC. Finally, as indicated in section 2, a number of evaluations are not made publicly available.

[26] This review should not be compared to a meta-evaluation, which would require a more systematic approach (OECD, 2002[7]).

[27] See e.g. (SECO, 2013[50]); (Smith et al., 2014[36]); (EU and AETS, 2015[51]); (DEval, 2017[52]).

[28] For exceptions, see (Smith et al., 2014[36]); (DEval, 2017[52]); (Ruben, R., & Tholen, 2017[38]).

[29] See e.g. (Smith et al., 2014[36]) and (Lerner et al., 2016[53]).

activities.[30] The focus on ESG standards and the SDGs seems only gradually to be incorporated in the interventions,[31] and there were few examples of assessments of trade-offs between financial returns, ESG considerations and development effects.[32] Finally, some evaluations, but not all, assessed the level of increased tax payments in developing countries as a result of the interventions.[33]

Evaluations varied in the extent to which they included assessments of the justification for the selection and design of specific interventions. Such assessments could – as indicated above - include the relevance of the blended finance projects for the stated development objectives (e.g. reducing poverty, job creation, ESG outcomes), choice of partner countries and alignment with their priorities (e.g. low-income countries), choice of sectors (e.g. finance or agriculture sector), and alignment with the SDGs. A common approach was to make a portfolio analysis and to assess its justification.[34] Some evaluations considered some broad arguments for the justification of the interventions, such as the absence of mature markets (Slob et al., 2017[35]), and several evaluations mentioned the existence of significant risk as a justification for blending. Very few evaluations, however, made an attempt to analyse various risk factors in detail[35] or the incentives of various actors.[36]

When efficiency was analysed, focus was usually on managerial and administrative efficiency, while only a few evaluations assessed whether the chosen blending instrument was the most efficient in the specific context.

Sustainability, including financial sustainability, and time lags of the effects were to varying degrees analysed in the evaluations.[37] A few evaluations explicitly mentioned that some of the effects are associated with time lags,[38] which implied that an evaluation just after the completion of the intervention may not provide a full picture of the effects of the intervention.

Approach and methodology

The majority of evaluations constructed or reconstructed theories of change (TOCs).[39] In most cases, TOCs do not seem to have been developed as part of the programme design. TOCs were typically developed to structure the whole evaluation, but there were also

[30] For an exception, see (European Commission, 2013[44]).

[31] See e.g. (DEval, 2017[52]); (Ruben, R., & Tholen, 2017[38]).

[32] See e.g. (DEval, 2017[52]) and (Spratt, 2012[53]).

[33] An example is (Spratt, S., O'Flynn, P. and Flyn, 2018[37]).

[34] See e.g. (Spratt, S., O'Flynn, P. and Flyn, 2018[37]).

[35] See for instance the distinction between different risk factors in (FMO, 2016[46]).

[36] For an exception, see (IEG, 2017[47]).

[37] Sustainability is assessed in several evaluations, including (EU and AETS, 2015[51]), (DEval, 2017[52])and (Spratt, S., O'Flynn, P. and Flyn, 2018[37]), while (FMO, 2015[48]) is an example of an attempt to analyse long-term effects.

[38] See e.g. (Lerner et al., 2016[53]).

[39] See e.g. (European Commission, 2016[33]), (DEval, 2017[52]) and (Slob et al., 2017[35]) for examples of constructed theories of change.

examples of evaluations that formulated TOCs for the individual blending projects.[40] Some referred to evaluations as being theory-based, but the 'theories' and assumptions to be assessed were not always made explicit. There were several references to contribution analysis, while other methodological approaches were only referred to in a few evaluations. In some evaluations formally applying contribution analysis, it was not clear to which degree contribution analysis had guided the evaluations in practice.

Several evaluations referred to mixed-methods approaches, but the way these were applied differed very much. In several cases it was not clear how the different methods could mutually complement each other.[41] Quantitative evaluation methods were used to a limited extent, which probably reflects methodological challenges, but also lack of data, including baseline data. In the reviewed sample, experimental and quasi-experimental designs were only used in a couple of evaluations,[42] while other evaluations used benchmarking.[43] The most frequently used methods appeared to be stakeholder interviews in combination with portfolio analysis, document and literature reviews.

Case studies, in most cases assessments of individual projects, were used in a number of evaluations, but the analytical approach was not always made explicit. This applied to sampling but also to the collection of case study information, comparisons and attempts at generalisation. The number of cases varied a great deal from around 10 cases to more than 200 cases.[44]

The evaluation did not apply a common approach to assessing additionality, and various dimensions of additionality were considered in the evaluations. The assessment of financial additionality varied in approach, but evaluators mainly based assessments of financial additionality on available monitoring data, interviews, surveys or qualitative case studies.[45] The leverage factor was widely used as an indicator for financial additionality and in several cases without critical reflections on its relevance. Some evaluations referred to ex-ante assessments of financial additionality but did not always assess the relevance and quality of these ex-ante assessments. Only a few evaluations addressed various dimensions of financial additionality (e.g. both mobilised finance and the terms of the finance), and whether blending had changed risk or perceived risk. In a couple evaluations, comparison with relevant research literature was used as an approach.[46]

[40] (Smith et al., 2014[36]) is an example of an evaluation which makes a clear methodological distinction between evaluating the programme and the individual projects.

[41] Examples of very systematic applications of a mixed-methods approach are (DEval, 2017[52]) and (Ruben, R., & Tholen, 2017[38]).

[42] See e.g. (Ruben, R., & Tholen, 2017[38]).

[43] See e.g. (Planet rating, 2014[49]).

[44] (Slob et al., 2017[35]) covered 20 cases, while (Smith et al., 2014[36]) and (Planet rating, 2014[49]) covered 11 cases, (European Commission, 2016[33]) analysed 32 cases, (FMO, 2015[48]) 36 cases and (DEval, 2017[52]) 216 cases.

[45] See e.g. (European Commission, 2016[33]) and (Slob et al., 2017[35]).

[46] See e.g. (Ruben, R., & Tholen, 2017[38]).

Methods for assessment of outcome and impact (development additionality) were described to varying degrees, and a simple 'before-and-after' method was widely used based on available data from indicators or information collected as part of the evaluation.

Only a few evaluations contained detailed reflections on the relevance and limitations of the applied methodology.[47] In spite of this, several evaluations contained relatively clear findings and conclusions on outcomes and impact.

Monitoring and evaluation system

Many – but not all - evaluations assessed the monitoring and evaluation system, including availability of information, relevance, reliability and measurement of indicators as well as financial reporting.[48] Data limitations, including lack of data, were mentioned as a challenge in several evaluations. This included lack of baseline data and comparable data over time due to changing monitoring practices (Smith et al., 2014[36]; Slob et al., 2017[35]; Spratt, S., O'Flynn, P. and Flyn, 2018[37]). Some evaluations also reported that activities were new and monitoring data were therefore not yet available (Smith et al., 2014[36]; European Commission, 2016[33]; Ruben, R., & Tholen, 2017[38]) and that the large number of indicators was seen as a challenge (Spratt, S., O'Flynn, P. and Flyn, 2018[37]). Limited access to information due to confidentiality and policies on disclosure of information was recurrently mentioned.

6. Concluding remarks and potential next steps

There is an urgent need for a better understanding of the potential role of blended finance in achieving the SDGs. Blended finance comprises various instruments and modalities, and monitoring and evaluation will be critical for establishing an evidence base about 'what works in which context'. The practice of evaluation of blended finance is still under development, and this paper highlights a number of challenges, which will have to be addressed.

First, the paper clearly shows that institutional and organisational factors influence evaluation practice. There is a significant variety in evaluation practice, including access to financial resources for evaluation, evaluation capacity, and degree of independence. It is important that a shared evaluation culture is established. Although some improvements have taken place recently, access to information and policies for disclosure of information are still challenges for evaluation. As mentioned, it is difficult to get an overview of completed evaluations, and it could be considered to create a specific sector or theme code in DEREC where evaluations could be uploaded. **Efforts should also be made to ensure that evaluation reports from DFIs and private intermediaries will, as a general principle, be publicly available.**

Second, the lack of a common language and terminology is a clear finding from the brief review of a sample of evaluations. Key concepts such as additionality and impact are used in different ways, which makes comparisons difficult. The OECD DAC 'Glossary of Key Terms in Evaluation and Results-Based Management' (OECD, 2002[7]) has greatly contributed to harmonising the language on monitoring and evaluating development aid. Some evaluations use the DAC evaluation terminology, but **it may be necessary to adapt**

[47] See e.g. (DEval, 2017[52]) and (IEG, 2017[47]) for exceptions.

[48] See e.g. (Smith et al., 2014[36]); (DEval, 2017[52]); (Spratt, S., O'Flynn, P. and Flyn, 2018[37]).

and develop existing terminology to make it relevant to blended finance evaluation. A similar tool (or an addendum to the existing glossary) targeting the field of development finance could thus be proposed.

Third, the review of the sample of evaluations also shows a wide diversity in applied evaluation approaches. This relates to the formulation of evaluation criteria as well as the choice of methodological approaches. The variety in approaches can to some extent be explained by differences in mandates and instruments, but a standardisation of approaches, where relevant, would facilitate comparison and learning. An example is that various approaches exist to assessing impact, and it would be useful to create an overview with the aim of ensuring some comparability. There may also be a need for **more detailed assessments of the relevance of individual evaluation methods for specific blended finance instruments and interventions**.

Fourth, evaluations of blended finance are highly dependent on the existence of robust monitoring systems. The quality of these systems varies, and efforts should be made to **improve access to information and to ensure availability of comparable monitoring data**. As mentioned, important work is ongoing on monitoring and results measurement of private sector engagement. This work should be continued and coordinated with efforts to strengthen the evaluation of blended finance.

Fifth, there is a strong need for learning more about 'what works and in which context'. The existing practice was assessed based on a brief review of a purposive sample of completed evaluations and other studies. The review only focused on evaluation approaches, but **a systematic review of existing evidence with a focus on achieved results by various blended finance instruments** would be an important contribution to the further discussions on the potential role and relevance of blended finance for the achievement of the SDGs.

The 2030 Agenda has expanded the number and diversity of financial actors, and addressing the above-mentioned five challenges of evaluation of blended finance will therefore require building consensus across a diverse and multilayered set of stakeholders. Historically, the OECD Development Assistance Committee (DAC) has played an important role in promoting the quality of the evaluation of development assistance. The same should now happen in the development finance field, where an important step in that direction has been taken by the adoption of the OECD DAC Blended Finance Principles for Unlocking Commercial Finance for the SDGs.

Annex: List of reviewed evaluations and other studies

ADB (2014), "Real-Time Evaluation of ADB's Initiatives to Support Access to Climate Finance Evaluation Independent", Corporate and thematic evaluation studies of ADB.

CDC (2015), What was the impact of CDC's fund investments from 2004 to 2012, http://www.hbs.edu/faculty/Pages/item.

Dalberg (2011), Report on Support to SMEs in Developing Countries Through Financial Intermediaries, Dalberg.

DEval (2017). Evaluation of the developpp.de programme. German Institute for Development Evaluation.

ECORYS (2016), Effectiveness study of Fund Emerging Markets for Development Cooperation, FMO. Anonymised Executive Summary.

European Commission (2016), Evaluation of Blending, ADE SA. https://ec.europa.eu/europeaid/sites/devco/files/evaluation-blending-volume1_en.pdf.

European Commission (2013), Mid-Term Evaluation of the Neighbourhood Investment Facility under the European Neighbourhood and Partnership Instrument (ENPI) 2007-2013.

EIB (2017), Evaluation of EIB Intermediated Lending through the Investment Facility in ACP Operations Evaluation. European Investment Bank.

EU and AETS (2015), Evaluation of Western Balkans Investment Framework (WBIF).

FMO Strategy Department (2015) Annual Sector Evaluation – Energy.

FMO Strategy Department (2016) Ex-post effectiveness study: (M)SME Financing in a fragile African state – Evaluation executed by A2F Consulting.Anonymised Executive Summary.

FMO Strategy Department (2016), Investments in Infrastructure, Manufactoring and Services.

GEEREF and EIB (2016), GEEREF IMPACT REPORT 2015: Catalysing EUR 10 billion of clean power.

IEG (2017). The International Finance Corporation's Approach to Engaging Clients for Increased Development Impact. An Independent Evaluation. Independent Evaluation Group, World Bank.

Independent Evaluation Office of the GEF (2017), Evaluation of GEF's Engagement with the Private Sector Volume I and II

Lemma, A. F. (2018). Measuring the Potential Contribution of DFIs to Economic Transformation. SET. Supporting Economic Transformation.

Lerner, J. et al. (2016), "Impact of Early Stage Equity Funds in Latin America. Multilateral Investment Fund. Member of the IDB Group.

Oxfam (2013), Impact report. Result of Impact study. Vision Fund Ghana.

Paniagua, G. and A. Denisova (2012), "Meta-Evaluation on Job creation effects of private sector interventions" Washington DC: World Bank.

Planet rating (2014), Impact Assessment of the REGMIFA Intervention on Partner Lending Institutions Regional MSME Investment Fund for Sub-Saharan Africa, TA FACILITY.

Pompa, C. (2013), Understanding challenge funds. ODI, London. UK.

Ruben, R., & Tholen, J. (2017), Assessing IDH's contribution to public good impacts at scale (2016-2020), Wageningen University & Research.

SECO (2013). Development Effects of SIFEM's Investment Interventions. Independent Evaluation. State Secretariat for Economic Affairs. Bern. Dalberg.

Slob, A. et al. (2017), Evaluation FMO Access to Energy Fund.

Smith, R. et al. (2014), "Evaluation of the Responsible and Accountable Garment Sector Programme Final Report". Oxford Policy Management.

Spratt, S. and L. Collins (2012), Development Finance Institutions and Infrastructure: A Systematic Review of Evidence for Development Additionality. Institute of Development Studies and Private Infrastructure Development Group.

Spratt, S., O'Flynn, P. and Flyn, J. (2018). DFIs and Development Impact: an evaluation of Swedfund. Expertgruppen for Bistandsanalys. 2018:1.

References

Banerjee, A. (2011), "Poor economics: A radical rethinking of the way to fight global poverty, Public Affairs.". [16]

Barder, O. and Talbot, T. (2015), "Guarantees, subsidies, or paying for success? Choosing the right instrument to catalyze private investment in developing countries.", No. 402, Center for Global Development,. [32]

Befani, B. and Mayne, J. (2014), "Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation.", pp. IDS Bulletin 45(6), 17-36.. [21]

Befani, B. (2016), *Pathways to change: Evaluating development interventions with QCA. Expert Group for Aid Studies.*. [17]

Bennett, A. (2004), *Case study methods: Design, use, and comperative advantages", Models, numbers, and cases: Methods for studying international relations*. [19]

Benn, J., C. Sangaré and T. Hos (2017), "Amounts Mobilised from the Private Sector by Official Development Finance Interventions Guarantees, syndicated loans, shares in collective investment vehicles, direct investment in companies and credit lines", https://www.oecd-ilibrary.org/docserver/8135abde-en.pdf?expires=1532438903&id=id&accname=ocid84004878&checksum=543F0A0796619DF343210C316FCD453A (accessed on 24 July 2018). [28]

Carter, P, Van de Sijpe, N. and Calel, R. (2018), "The Elusive Quest for Additionality", No. Working Paper No. 495, Center for Global Development, Washington DC. [13]

Deaton, A. (2009), "Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development", The National Bureau of Economic Research, http://dx.doi.org/10.3386/w14690.. [15]

DEval (2017), "Evaluation of the developpp.de programme". [52]

DFI Working Group (2018), *DFI Working Group on Blended Concessional Finance for Private Sector Projects: Joint Report (October 2018 Update)*, https://www.adb.org/sites/default/files/institutional-document/457741/dfi-blended-concessional-fiance-report.pdf (accessed on 07 January 2019). [5]

EBRD (2018), *Additionality in the EBRD - Review of Concept and Application*, http://www.ebrd.com. [27]

ECG (2006), "MDB-ECG Good-Practice Standards for the Evaluation of Private Sector Investment Operations. Third Edition". [4]

Escalante, D., Abramskiehn, D., Hallmeyer, K. and Brown, J. (2018), *Approaches to assess the additionality of climate investments: Findings from the evaluation of the Climate Public Private Partnership Programme.*, Climate Policy Initiative.. [31]

EU and AETS (2015), *Evaluation of Western Balkans Investment Framework*. [51]

European Commission (2016), *Evaluation of Blending - Volume I – Main Report*. [33]

European Commission (2013), *Mid-Term Evaluation of the Neighbourhood Investment Facility under the European Neighborhood and Partnership Instrument*. [44]

European Investment Bank (2017), *Evaluation of EIB Intermediated Lending through the Investment Facility in ACP Operations Evaluation*. [34]

FMO (2016), *Infrastructure Development Fund*. [46]

FMO (2015), *ANNUAL SECTOR EVALUATION 2015 - Energy*. [48]

Goertz, G. (2017), *Multimethod Research, Causal Mechanisms, and Case Studies*, Princeton University Press, http://dx.doi.org/10.1515/9781400888115. [9]

Goodrick, D. (2014), "Comparative case studies: Methodological Breifs", http://dx.doi.org/10.1007/bf01322868. [18]

Heinrich, M. (2014), *Demonstrating addtionality in private sector development initiatives*, Edward Elgar Publishing, http://dx.doi.org/10.4337/9781781951194.00013. [25]

IEG (2017), *The International Finance Corporation's Approach to Engaging Clients for Increased Development Impact. An Independent Evaluation*, Independent Evaluation Group, World Bank.. [47]

IOB (2009), *Investing in infrastructure. Evaluation of the LDC Infrastrucuture Fund*, http://dx.doi.org/10.2172/125028. [30]

Lemma, A. (2018), "Measuring the Potential Contribution of DFis to Economic Transformation.", http://dx.doi.org/10.1080/14631379708427881. [23]

Lemma, A. (2015), "Development Impact of DFIs: What are their impacts and how are they measured?", *ECONOMIC AND PRIVATE SECTOR PROFESSIONAL EVIDENCE AND APPLIED KNOWLEDGE SERVICES* , http://partnerplatform.org/eps-peaks (accessed on 26 November 2018). [3]

Lerner, J. et al. (2016), *Impact of Early Stage Equity Funds in Latin America*. [53]

Linden, A. and J. Adams (2012), "Combining the regression discontinuity design and propensity score-based weighting to improve causal inference in program evaluation", *Journal of Evaluation in Clinical Practice*, Vol. 18/2, pp. 317-325, http://dx.doi.org/10.1111/j.1365-2753.2011.01768.x. [14]

MacGillivray, A., Kim, R., van Moorsel, T. and Kehoe, A. (2017), *Measuring Total Employment Effects: a lean data methodology for a portfolio of investments in developing countries.*, Organisation for Economic Co-Operation and Development (OECD), http://dx.doi.org/10.1787/dcr-2017-graph13-en. [43]

OECD (2018), *DAC methodologies for measuring the amounts mobilised from the private sector by official development finance interventions*, https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/DAC-Methodologies-on-Mobilisation.pdf (accessed on 07 January 2019). [29]

OECD (2018), *Making Blended Finance Work for the Sustainable Development Goals*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264288768-en. [1]

OECD (2018), *Measuring the results of private sector engagement through development co-operation*, OECD, Paris, https://www.oecd.org/dac/results-development/docs/results-pse-results-workshop-apr-18.pdf (accessed on 03 December 2018). [39]

OECD (2016), *PRIVATE SECTOR PEER LEARNING Peer Inventory 1: Private Sector Engagement Terminology and Typology Understanding Key Terms and Modalities for Private Sector Engagement in Development Co-operation*, http://www.oecd.org/dac/peer-reviews/Inventory-1-Private-Sector-Engagement-Terminology-and-Typology.pdf (accessed on 07 January 2019). [26]

OECD (2010), *DAC Guidelines and Reference Series Quality Standards for Development Evaluation*, https://www.oecd.org/development/evaluation/qualitystandards.pdf (accessed on 07 January 2019). [8]

OECD (2002), *Glossary of key terms in evaluation and results based management*, https://www.oecd.org/dac/evaluation/2754804.pdf (accessed on 07 January 2019). [7]

OECD (1991), *DAC Principles for Evaluation of Development Assistance*, https://www.oecd.org/development/evaluation/2755284.pdf (accessed on 07 January 2019). [6]

Pawson, R. and A. Manzano-Santaella (2012), "A realist diagnostic workshop", *Evaluation*, Vol. 18/2, pp. 176-191, http://dx.doi.org/10.1177/1356389012440912. [42]

Planet rating (2014), "Impact Assessment of the REGMIFA Intervention on Partner Lending Institutions Regional MSME Investment Fund for Sub-Saharan Africa, TA FACILITY". [49]

Romero, M. J. and Van de Poel, J. (2014), *Private finance for development unravelled. Assessing how Development Finance Institutions work.*, Elsevier, http://dx.doi.org/10.1016/b978-0-12-391058-5.00016-3. [2]

Ruben, R., & Tholen, J. (2017), *Assessing IDH's contribution to public good impacts at scale (2016-2020)*, Wageningen University & Research.. [38]

Ruff, K. and S. Olsen (2018), "The Need for Analysts in Social Impact Measurement", *American Journal of Evaluation*, Vol. 39/3, pp. 402-407, http://dx.doi.org/10.1177/1098214018778809. [41]

SECO (2013), *Development Effects of SIFEM's Investment Interventions. Independent Evaluation. State Secretariat for Economic Affairs.*, Dalberg, Bern. [50]

Sjö, B. and S. Flygare (2008), "Strength Evaluation of Interior Slab-Column Connections", *ACI Structural Journal*, Vol. 105/6, http://dx.doi.org/10.14359/20097. [24]

Slob, A. et al. (2017), *Evaluation FMO Access to Energy Fund*. [35]

Smith, R. et al. (2014), *Evaluation of the Responsible and Accountable Garment Sector Programme Final Report*. [36]

Spratt, S., O'Flynn, P. and Flyn, J. (2018), *DFIs and Development Impact: an evaluation of Swedfund. Expertgruppen for Bistandsanalys.*. [37]

Spratt, S. (2012), *elopment Finance Institutions and Infrastructure: A Systematic Review of Evidence for Development Additionality.*, Institute of Development Studies and Private Infrastructure Development Group.. [45]

Stern, E. (2015), "Impact Evaluation. A guide for Commissioners and Managers. Bond for International Development.", *Nursing Standard*, http://dx.doi.org/10.7748/ns.14.3.26.s39. [20]

Stern, E. et al. (2012), *Broadening the range of designs and methods for impact evaluations*, http://dx.doi.org/10.22163/fteval.2012.100. [11]

Ton, G. et al. (2019), *Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?*, http://www.ids.ac.uk/cdi (accessed on 22 January 2019). [22]

Vo, A. and C. Christie (2018), "Where Impact Measurement Meets Evaluation", *American Journal of Evaluation*, Vol. 39/3, pp. 383-388, http://dx.doi.org/10.1177/1098214018778813. [40]

White, H. (2009), *Theory-based impact evaluation: principles and practice*, http://dx.doi.org/10.23846/wp0003. [10]

White, H. and D. Phillips (2012), *Addressing attribution of cause and effect in small impact evaluations: towards an integrated framework*, http://dx.doi.org/10.23846/wp0015. [12]


OECD
BETTER POLICIES FOR BETTER LIVES