FERDI CHAIR Impact Investing

Gouvernement Princier
PRINCIPAUTÉ DE MONACO

Report
June 2026

What Impact Funds in Africa Really Measure: An Analysis of Impact Evaluation Practices

Report for the Impact Investing Chair at FERDI. With financial support from the DCI of the Government of the Principality of Monaco.

Florian Léon, Research Officer, FERDI

Zeïneb Benaïssa, Research Assistant, FERDI

LA FERDI EST UNE FONDATION RECONNUE D'UTILITÉ PUBLIQUE.
CETTE PUBLICATION A BÉNÉFICIÉ D'UNE AIDE DE L'ÉTAT FRANÇAIS AU TITRE
DU PLAN D'INVESTISSEMENT FRANCE 2030 PORTANT LA RÉFÉRENCE ANR-16-IDEX-0001.

Executive Summary

Context and Objectives

Impact investment funds seek to maximise the economic, social and environmental impacts of their investments, subject to financial sustainability constraints. To justify potentially lower financial returns, these funds must be able to demonstrate the reality of their non-financial impacts. Impact evaluation is therefore a central dimension of their activity — and yet it remains largely underdeveloped in practice.

This report, produced by the FERDI Impact Investing Chair, has three objectives. First, to clarify what impact evaluation is and how it can be implemented by investment funds. Second, to analyse the effective practices of impact funds operating in Africa, using an original evaluation grid applied to 250 funds. Third, to propose concrete avenues for improving these practices, without losing sight of the fact that these structures' primary mission remains to invest.

What Impact Evaluation Is — and What It Is Not

Impact evaluation encompasses the entire process of assessing the effects of an investment on its beneficiaries and, more broadly, on society. It is distinct from simple measurement, which refers more specifically to the step aimed at analysing the causal relationship. In practice, the two terms are often conflated. In the context of our study, we focus on impact evaluation, considering its various stages, of which measurement is an essential milestone.

This process is structured around four complementary steps. First, the theory of change identifies the mechanisms through which an investment is expected to produce effects, distinguishing between inputs, outputs, outcomes and impacts. Second, indicators translate these mechanisms into measurable quantities, balancing precision with data collection feasibility. Third, attribution seeks to establish that observed changes are genuinely due to the investment and not to other factors. This is the most demanding step, ideally requiring the construction of a counterfactual, i.e. a hypothetical situation representing what would have happened in the absence of the intervention. Finally, the use of results determines the real value of the evaluation: a measurement that does not feed into decision-making, learning or communication with stakeholders remains a formal exercise with no real impact.

It is worth noting a complexity specific to investment funds, often absent from public policy evaluations: the question of additionality. When a state implements a public policy, it acts alone — if the policy does not exist, no one substitutes for it. In contrast, if a fund does not invest in a company, another investor might do so in its place. It then becomes difficult to attribute the entirety of the observed impact to the fund itself. This reality calls for a careful distinction between the impact of the investee company and the fund's own impact, and is a reminder that additionality is not a given but a hypothesis to be verified.

The Practices of Impact Funds in Africa: A Concerning Diagnosis

To analyse the practices of impact funds operating in Africa, this report constructs an original evaluation grid structured around five dimensions: the general analytical framework, the definition of indicators, impact attribution approaches, reported results, and transparency. This grid, comprising 36 binary

questions, is applied to 250 funds identified by Léon and Rabary (2025), based on a systematic documentary review of their websites, annual reports and public communications.

The results paint a concerning picture. The average score across all five dimensions is low, and practices vary considerably from one fund to another. Three findings stand out.

First, the theory of change is present in a majority of funds, but often remains superficial. While many structures mention an intention to measure impact, few actually formalise the mechanisms linking their investments to expected effects.

Second, the definition and monitoring of indicators are highly uneven. Output indicators are widely used, but outcome and impact indicators — which require time-based monitoring and reflection on causality — remain rare. Thus, impact indicators are defined by only 6.8% of the sample, and outcome indicators by 38.4%. Indirect effects and environmental dimensions are even more frequently absent.

Third, and most strikingly, impact attribution is virtually absent from observed practices. Fewer than one in ten funds uses an approach incorporating a counterfactual. The vast majority relies on before-after comparisons or indicator tracking without a reference group, which does not establish a causal link between the investment and the observed effects.

The empirical analysis also reveals that certain fund characteristics are correlated with more advanced practices. Funds headquartered in Europe display the highest scores, which may partly be explained by the growing regulatory pressure on non-financial reporting in that region. Fund size and age also appear to play a positive role, though the effects are not always significant. In contrast, African footprint (i.e. the share of investments made on the continent) is not associated with better practices, suggesting that geographic proximity is insufficient to improve evaluation rigour.

These results can be explained by several structural hurdles. The cost of impact evaluation is real and often perceived as disproportionate relative to the expected benefits. Teams frequently lack methodological expertise. Access to data, particularly that needed to construct a comparison group, remains limited in the African context. Finally, incentives to produce rigorous evaluations remain insufficient: funders rarely require in-depth analysis, and the risks associated with impact washing remain poorly sanctioned.

Six Avenues for Improving Practices

The current situation is not inevitable. Six action pathways, complementary and progressive, can help embed impact evaluation at the core of funds.

The first consists of raising awareness across the entire ecosystem (managers, funders, stakeholders) about the importance of impact evaluation, emphasising that it is not a constraint but a tool for differentiation and credibility. This responsibility falls primarily on professional associations and donors.

The second avenue is the development of a methodological framework adapted to the African context. Rather than imposing a single method, this framework would help funds identify the most appropriate

approach for their situation, by matching the targeted impact, available data, mobilisable budget and the most suitable method. Its design should be collective, involving funds, experts and funders.

The third avenue concerns the application of basic methodological principles: a clearly defined theory of change, indicators aligned with priority impacts, integration of a simplified notion of counterfactual, and transparency regarding methodological choices and their limitations.

The fourth avenue concerns improving access to data, particularly data on non-beneficiary entities, which are essential for constructing a comparison group. Datasets exist (World Bank portal that get access to many micro datasets, open administrative data, satellite analyses) but their existence is often unknown to practitioners. A directory of available sources by impact theme would be a valuable tool.

The fifth avenue is financial and technical support for funds. Grants or dedicated funds could cover part of the initial costs related to recruitment or training, under the impetus of development finance institutions. Partnerships with academic structures or specialist experts could complement this support.

The sixth and final avenue invites funds to adopt a progressive and iterative approach, beginning with pilot projects in sectors where indicators are well defined, without trying to measure everything from the outset. The important thing is to engage in a process of continuous improvement, accepting that early evaluations may be imperfect.

These six avenues are not independent of one another. Their effectiveness rests on a collective dynamic in which every actor in the ecosystem — funds, donors, experts, academic actors — has a role to play. It is on this condition that impact evaluation can become, in Africa as elsewhere, a genuine lever of transformation rather than a mere formal obligation.

Table of Contents

List of Tables

List of Figures

List of Boxes

List of Examples

Abbreviations and Acronyms

AUM — Assets Under Management

CSRD — Corporate Sustainability Reporting Directive

DAC-OECD — Development Assistance Committee of the Organisation for Economic Co-operation and Development

ESG — Environmental, Social and Governance

EVRI — Environmental Valuation References Inventory

IADB — Inter-American Development Bank

FCDO — Foreign, Commonwealth and Development Office

FERDI — Foundation for Studies and Research on International Development

GIIN — Global Impact Investing Network

DFI — Development Finance Institution

IRIS+ — Impact Reporting and Investment Standards

MOPAN — Multilateral Organisation Performance Assessment Network

SDG — Sustainable Development Goal

SME — Small and Medium-sized Enterprise

SROI — Social Return on Investment

ToC — Theory of Change

SRR — Social Rate of Return

Glossary

Additionality: The capacity of an investor to make possible an investment that would not have occurred without their intervention. An investment is considered additional when it finances projects or companies that would not have found financing from other sources, or could only have done so on less favourable terms. Additionality should not be confused with impact: an investment can be additional without generating significant non-financial effects, and vice versa.

Impact Attribution: The process of establishing that observed changes in the selected indicators are genuinely attributable to the investment made, and not to other external factors. Attribution relies on the construction of a counterfactual and is the most demanding step in impact evaluation.

Counterfactual: A hypothetical situation representing what would have happened in the absence of the investment. The counterfactual cannot be directly observed: it must be reconstructed from the evolution of a comparison group with characteristics similar to those of the investment's beneficiaries. It constitutes the foundation of any rigorous impact evaluation.

Impact Evaluation: A comprehensive process aimed at assessing the effects of an investment on its beneficiaries and, more broadly, on society. It encompasses all steps from defining the theory of change to using the results, including collecting indicators and attributing impact. It is distinct from simple impact measurement, which refers more specifically to the quantification of observed effects.

Ex-ante Evaluation: An evaluation conducted before the investment is implemented, aimed at estimating the expected non-financial effects. Like financial forecasts, it is based on assumptions and allows for the establishment of a baseline scenario against which actual results can be compared.

Ex-post Evaluation: An evaluation conducted after the investment has been made, aimed at analysing the non-financial effects actually produced. Comparing the ex-ante and ex-post evaluations enables the identification of discrepancies and an understanding of their causes, in a logic of learning and continuous improvement.

Output: The direct and immediate result of a funded activity, generally observable in the short term. In the theory of change, outputs are positioned between inputs and outcomes.

Impact: The lasting effect produced by an investment on beneficiaries and their environment, beyond direct and short-term effects. Impact differs from outcomes in its temporal dimension (it occurs in the medium or long term) and its scope (it may affect indirect stakeholders as communities, environment, markets). In this report, the term is used broadly to encompass economic, social and environmental dimensions.

Impact Washing: The practice of claiming positive social or environmental impact without genuinely demonstrating or substantiating it. By analogy with greenwashing, impact washing refers to a form of misleading communication that harms the credibility of the impact investing sector as a whole.

Indicator: A variable used to measure a dimension of the theory of change — input, output, outcome or impact. A good indicator should be as close as possible to what one wishes to measure, collectible at an appropriate frequency, and, where possible, harmonised with sector standards (such as IRIS+).

Input: A resource mobilised to implement an investment. In the theory of change, inputs represent the starting point of the causal chain. Example: the capital invested by a fund in a company.

Impact Investment: An investment made with the explicit intention of generating measurable positive social and/or environmental effects, while seeking a financial return. Impact investment is distinguished from traditional investment by the deliberate consideration of non-financial dimensions, and from philanthropy by the requirement for a financial return.

Impact Measurement: The quantification of the effects of an investment through the collection and monitoring of indicators. Impact measurement constitutes a step within the broader impact evaluation process, providing the necessary data.

Outcome: An intermediate effect produced by an investment, situated between outputs and impact in the theory of change. Outcomes are generally observable in the medium term and directly concern the investment's beneficiaries.

Social Return on Investment (SROI): A method for valuing the social impact of an investment that seeks to express non-financial effects in monetary terms. The SROI enables the cost of an investment to be weighed against the social value it generates, facilitating comparison between different types of returns — financial and non-financial.

Theory of Change (ToC): An explicit representation of the mechanisms through which an investment is expected to produce effects on its beneficiaries and their environment. It decomposes the causal chain into four successive elements — inputs, outputs, outcomes and impacts — and identifies assumptions and risks that may affect transmission between each link. The theory of change constitutes the starting point of any rigorous impact evaluation process.

Impact Thesis: A strategic framework defining a fund's vision of the impacts it seeks to produce through its investments. The impact thesis specifies areas of intervention, target populations, mechanisms for creating non-financial value, and pursued impact objectives. It constitutes the guiding thread of the investment strategy and the anchor point of the theory of change.

Introduction

Impact investment funds seek to maximise the economic, social or environmental impacts of their investments, subject to financial sustainability constraints. These investors may accept a reduction in their financial return in exchange for impacts. In order to compare these two types of returns — financial and non-financial — it is important to have metrics for each. While there is consensus on how to measure financial returns, debates on the best way to measure the non-financial effects of investments remain wide open. The question of impact measurement is therefore a crucial dimension of the activity of these funds.

Impact measurement has two different meanings depending on the audiences referring to it.

On the one hand, academic economics research refers to the evaluation of public policies and social programmes, and to experimental and quasi-experimental methods. This body of work has taken a central place in the field of research, particularly in development economics and labour economics. These studies have sought more rigorous approaches to identify the causal effects of a public policy. This movement has given rise to a "credibility revolution" (Angrist and Pischke, 2010) that has permeated the entire profession, moving from correlation to causality.

On the other hand, practitioners (including impact investors) have a more flexible and contextual view of impact evaluation. Their approach favours qualitative methods, such as case studies or change narratives, the monitoring of a few key indicators, and adapts to the specificities of each investment.

The challenge for impact investors is to find the right balance between rigour and practicality. Researchers are capable of producing analyses in which impact attribution is clearly established. However, these studies are costly and difficult to implement for investors. The objective of this report is to understand the practices of impact investors operating in Africa in order to find solutions for improving the rigour of their analyses. With the perspective of economists working on impact investing, this report also aims to better understand how impact investors can draw inspiration from public policy evaluation to enrich their practice.

The first part presents what impact measurement is. It aims to recall its rationale (why measure impact?) but also the different stages of the process and the way to integrate it into the functioning of a fund (how?). The second part is an exploration of the practices of impact investors operating in Africa in order to better understand their approaches to impact evaluation. After constructing a grid for assessing the methods in place, we analyse these for the 250 funds operating on the African continent (located in Africa or elsewhere). Finally, in a third part, we present some avenues for opening up areas for improvement of existing practices. Our ambition is to show that funds, without becoming impact measurement experts, can adjust their practices to make them more rigorous.

Part 1 – Impact Evaluation for an Impact Fund

This first part aims to provide an overview of the key principles of impact evaluation before turning to the practices of impact investment funds operating in Africa (in the following section). This first part of the report is divided into four sections, each addressing a central question. The first section examines the reasons justifying impact evaluation ("*why measuring impact?*"). The following section is dedicated to clarifying the general principles of impact evaluation.[1] The third section presents the implementation of impact measurement ("*how?*"), drawing on a fictitious case study. Finally, the last section focuses on the use of impact analysis results ("*for what purpose(s)?*").

1.1. The Rationale for Impact Measurement

1.1.1. Good Reasons to Measure Impact

Impact funds reconcile financial returns with the impact of their investments. It is therefore essential for them to measure impacts in addition to financial returns, if only to justify a reduced financial return. Nevertheless, this argument does not exhaust the reflection on impact measurement. Establishing a robust measure of impact is a costly procedure. Funds, even impact funds, may therefore be tempted to proceed quickly and provide an imperfect measure. Nevertheless, measuring impact correctly is beneficial for two main reasons (Roor and Maas, 2024): (i) proving impact; (ii) improving practices.

1.1.1.1. Proving Impact

This first justification aims to demonstrate the impact of investments to stakeholders, particularly funders and beneficiaries.

At the macroeconomic level, this involves justifying the investment, ensuring transparency and reinforcing the credibility of the impact investing sector. This can attract traditional investors by giving them confidence in the sector and its specificities, thus enabling capital within the sector to grow and resources available to achieve various impact objectives to increase. It also enables a distinction between genuine impact funds and those that merely claim to be, thereby avoiding "impact washing[2]".

At the microeconomic level, producing an impact analysis is essential vis-à-vis funders. It enables the justification of a reduced financial return in exchange for a social return on investment[3]. The social return on investment can thus be weighed against the financial return. A fund can thus easily justify certain investments with reduced financial returns if they have high non-financial impacts. Capital owners sensitive to these social and environmental aspects will thus be more inclined to support this

[1] The terms "*impact assessment*" and "*impact measurement*" are often used interchangeably. However, there is a difference in that assessment refers to the overall process, whereas measurement focuses more specifically on quantification (data collection and monitoring, and measuring the effect). Whenever possible, we maintain this distinction in the report.

[2] Impact washing refers to companies that claim to have a positive impact on social or environmental issues without actually taking meaningful action.

[3] The term "social return" (or "social return on investment") is used in a broad sense in this report to refer to all non-financial impacts—whether economic, social, or environmental—that benefit society (the community) as a whole. It should not be understood as referring solely to social impacts.

type of investment and impact funds. These funders will also be able to justify to their own governance (i) the alignment of sought and obtained impacts with their mandate, (ii) the possible weakness of the financial return sought or obtained, and (iii) the level and relevance of any subsidies provided to the financed investment vehicle or entrepreneurial target.

Beyond the discussion with fund investors, this analysis is useful within the structure to understand the different expected returns across the various investments made. Furthermore, this priority given to impact measurement can also change the life of the fund by modifying team incentives. Having robust metrics for impact enables the construction of an incentive framework that relies not only on financial returns but also on non-financial returns, in line with the missions the fund sets itself.

1.1.1.2. Improving Practices

The second justification aims to use evaluation as a tool for learning and continuous improvement. Impact evaluation enables a better understanding of what happened (formative approach) by providing feedback on the effects of the investment and comparing them with expectations. Investment funds, during the study of an investment file, establish forecasts not only on financial performance but also on the impact of the investment (ex-ante evaluation). Evaluating the real impact observed ex-post and comparing it with forecasts first enables a check on whether what was planned actually happened, and, if not, initiates reflection on explanations for this finding. Indeed, a difference (positive or negative) between ex-ante and ex-post evaluations may be due to exogenous or endogenous factors. By analysing the possible reasons behind observed differences, the fund can better understand the factors influencing impact performance.

Understanding what happened is only useful if the lessons drawn are subsequently exploited to improve the fund's operations. From the analysis of reasons explaining deviations between expectations and achievements, the fund can change its practices to prevent possible risks or make the most of opportunities. This continuous learning is often carried out for financial aspects. Nevertheless, for an impact fund, it is wise to continually compare indicators related to non-financial dimensions against predefined objectives without waiting for the collaboration between the fund and the company to end. Indicators relating to the progress of the intervention, such as the frequency of data collection or the delays between decisions and their implementation, are also concerned. It is more efficient to quickly identify what is not working in order to adjust resources accordingly than to wait until the end of the contract to realise that several years' work has not followed the expected trajectory.

This individual process also serves as an example for other projects or other funds wishing to invest in the same sectors and achieve the same objectives. The communicated experience enables proof that the corresponding model is viable, knowledge of what works or does not, and awareness of what could make an investment obsolete. Learning would thus benefit the fund itself, but also the sector as a whole.

Box 1: Ex-ante and Ex-post Evaluations: Principles, Differences and Common Points

Impact evaluation can precede the investment (ex-ante evaluation) or follow it (ex-post evaluation).

Ex-ante evaluations (a priori, prospective) serve to establish the expected effects of an investment. Like financial forecasts, they are built on assumptions to provide a credible picture of the non-economic effects the investment will generate (number of jobs created, tonnes of CO_2 avoided, etc.).

Ex-post evaluations (a posteriori, retrospective) serve to analyse what the impacts of the investment actually were once it has taken place. Like financial results, it is possible to compare expected forecasts (ex-ante evaluation) with achieved results (ex-post evaluation).

The ex-ante analysis enables the construction of a scenario and the ex-post analysis enables observation of what actually happened. Comparing the two enables the identification of discrepancies (positive or negative) in order to understand their causes. This phase is essential to ensure that impact evaluation is part of a permanent improvement of practices.

While the most common relationship is to start from ex-ante evaluations to judge ex-post results, ex-post analyses are themselves useful for constructing credible ex-ante evaluations. Indeed, it is sometimes difficult to know a priori the non-financial effects of an investment. Experts will therefore draw on existing impact evaluations, conducted either internally or externally. These studies, although having taken place in different contexts, enable the expected effects to be calibrated within a credible range. A good ex-ante analysis must be able to justify the parameters applied, and the use of previous studies is a good approach for this.

1.1.2. Good (and Not-So-Good) Reasons Not to Measure Impact

While impact evaluation is important for impact funds, there are good reasons not to conduct it. In an original article, Gugerty and Karlan (2018) articulate the good reasons not to measure impact. Their perspective differs from ours. The authors question the rationality of wanting to implement an impact evaluation for every public policy or project supported by donors. Their starting point is that this dynamic can lead to evaluations that are not very robust and therefore not very informative. This situation can be compared to that of many impact funds that produce analyses whose quality, and ultimately usefulness, can be questioned (see Part 2 of the report).

Gugerty and Karlan (2018) classify good reasons not to measure impact into four categories: (i) not the right tool; (ii) not now; (iii) not feasible; (iv) not worth it. We follow this typology to present the various arguments.

1.1.2.1. Not the Right Tool

The first argument advanced is that impact evaluation is not always the right tool to answer the question posed. Impact analysis as disseminated in the academic world is an exercise based on the collection and

analysis of quantitative data. However, investors may seek to evaluate the effect of an investment on aspects that are difficult to quantify. It is always possible to collect data to assess qualitative elements, but this often involves simplification and temporal monitoring of such indicators is often complex.

This argument highlights that impact evaluation often focuses on one particular dimension (for example, employment generated). However, investments almost never affect only one aspect; they have multiple impacts on different dimensions that may even move in opposite directions.

The choice to focus on a limited number of impacts falls within a framework (the theory of change, detailed below) that enables identification of the most probable non-financial effects. However, this choice may also be guided by other considerations, such as the ability to measure and monitor one metric rather than another. The risk is finding oneself faced with what economists call the lamppost paradox (Fitoussi, 2013). Studies only shed light on what can be illuminated (i.e. what is situated under the lamppost's light) even if these are not the most important aspects. Furthermore, resorting to a robust impact analysis implies defining a strict framework, which supposes that expected effects are known in advance. Yet it is possible that certain unforeseen effects will materialise during the project. The approach must therefore be able to adjust to be capable of including these unforeseen impacts.

The selection of a limited number of impacts for in-depth evaluation can finally lead to ignoring unmeasured negative impacts, whose intensity may cancel out the benefits of the former. Many policies or impact theses have contradictory dimensions. A classic example in environmental matters is the contradiction between carbon emissions and pollution (nuclear produces low-carbon energy but generates waste whose treatment creates major challenges) or between green energy, biodiversity and food (the production of biobased green energy has contradictions with food supply, and the coverage of vast areas with solar panels or wind turbines generates significant negative environmental externalities).

Another argument advanced by Gugerty and Karlan (2018) concerns the existence of indirect effects. A credible impact measure implies being able to identify direct effects without too much interference. However, certain investments may have effects not only on targeted entities but also on other actors. The existence of indirect effects can make analysis complex, especially if it is difficult to identify and account for these indirect effects[4].

These arguments do not imply a rejection of impact analysis. They underline that it is a method for evaluating the effects of an investment, but it is not the only existing approach. Other approaches can (must) be mobilised to clarify the effects (for example, by considering beyond the number of jobs generated, their quality) or to capture non-quantifiable, negative, indirect or unforeseen effects.

[4] Suppose a company in a given market receives significant support. To determine whether this investment has had an impact, it is helpful to compare it with the performance of other companies operating in the same market. However, if the funded company is a key player, a change in its operations can affect all other companies. For example, by offering a new, lower-priced product, it can reduce its competitors' business.

1.1.2.2. Not Now

Gugerty and Karlan (2018) advance a second series of arguments relating to the timing of measurement. The fundamental idea is that impact measurement must be carried out at the time of actual project implementation — neither too early (project not yet implemented) nor too late (project already completed or even generalised). Beyond this logical principle, the authors insist on the importance of properly constructing the theory of change, i.e. the logical chain that enables understanding of how the invested funds will generate impact in the longer term. During the analysis, it is important to ensure that this chain is coherent and that effects materialise at each stage.

This argument recalls the importance of properly aligning the timing of the evaluation with that of the project and ensuring that the impact evaluation is coherent with the constructed theory of change. Where possible, the impact measurement protocol should be structured from the earliest stages of the project in order to be best integrated into the process. However, just because the genesis of the investment did not address this point from the outset does not mean that an impact evaluation is either infeasible or uninformative. Methods exist to reconstruct counterfactual analyses after the project has been carried out. These can thus be implemented ex-post. The only major obstacle to conducting an unplanned evaluation is the fact that the necessary data were not collected throughout the project (particularly before and after). In this case, it is more difficult to conduct a convincing impact evaluation.

1.1.2.3. Not Feasible

The third series of arguments is undoubtedly the one most frequently put forward by impact funds.

The first argument concerns limited resources, particularly human resources, for carrying out this type of exercise. Funds can externalise or internalise this task (see Box 2 below). Internalisation allows for an impact evaluation for all investments according to a standardised approach. However, this approach requires having qualified staff dedicated to this task. Many funds operating in Africa are small structures (Léon and Rabary, 2024) that do not necessarily have the capacity to allocate a team to impact evaluation. There may also be resource access constraints. This may involve people trained in impact evaluation, or limited access to data, since impact evaluation can be relatively expensive in terms of data collection (surveys), even if current developments can increasingly allow this problem to be circumvented.

Finally, a good impact evaluation requires a stable protocol that is not always compatible with the life of a company. The latter often has to manage many unexpected events and is led to adjust its strategy and activities. Yet, in the absence of a well-established protocol, it is difficult to trace a clear causal relationship between the investment and the observed effects.

1.1.2.4. Not Worth It

Gugerty and Karlan (2018) address a final category of arguments — the absence of utility of impact evaluations. It should be kept in mind that their study is set in a different framework from ours (public policy evaluation) and in a knowledge production approach. They advance two arguments suggesting that (i) the answer to the question is already known and (ii) the results are not generalisable.

In the context of impact investors, this argument is sometimes put forward in a different form. Some funds draw on the principle of additionality to justify the absence of impact evaluation. On the grounds that they have invested in a company where no one else would have, and that the company has grown, they conclude that their investment had a positive impact. This amounts, however, to confusing additionality with impact. Additionality refers to the capacity of an investor to make possible an investment that would not have occurred without their financing (Carter et al., 2021). Impact, on the other hand, implies that this investment has produced non-financial effects — economic, social, environmental — on stakeholders (financed company, suppliers, customers, communities, etc.). Deducing that additionality, coupled with a favourable evolution of the company, is synonymous with impact is incorrect for two reasons. On the one hand, the positive dynamic of the company may result from the investment, but also from other factors, such as market growth. On the other hand, the notion of impact goes well beyond the economic and direct effect on the financed company, which makes impact evaluation all the more necessary.

Box 2: Who Should Conduct the Impact Evaluation?

Impact evaluation can be conducted internally or outsourced. There are arguments in favour of each approach.

Outsourcing impact evaluation — i.e. having it conducted by a team external to the fund — has several advantages. First, it does not require the fund to recruit professionals dedicated to this task. Thus, outsourcing is particularly useful during the fund's establishment phase, when activities are primarily focused on the investor's core business. Second, using an external team enables the development of specific evaluation tools that require deep mastery of impact evaluation. This approach is thus relevant when the fund wishes to improve its evaluation practices on a particular aspect and can thus work with specialists in this field. Third, outsourcing provides protection against criticism of result manipulation. Nevertheless, there are limits to this exercise, as the service provider is funded by the fund and the fund may always retain the freedom to present the most favourable results.

Internalisation involves conducting an evaluation using the fund's own team. The main advantage of internalisation is the ability to systematise impact evaluation within the fund. Thanks to a dedicated team, the fund can build a standardised approach that can be replicated for all of the fund's investments. To ensure that this internal evaluation is credible, it is important to ensure that the evaluation team is independent and has the means to carry out its task (human and material resources, access to data, etc.).

The two approaches are not mutually exclusive and can be combined. Outsourcing is particularly suitable for one-off evaluations or for technical aspects requiring specialised expertise, while internalisation enables the embedding of evaluation into the fund's daily operations and its systematisation across the entire portfolio.

1.2. General Principles of Impact Evaluation

1.2.1. Impact Thesis vs. Theory of Change

It is important to distinguish two related concepts that can sometimes be confused: impact thesis and theory of change.

- The impact thesis is a general framework aimed at orienting the activities of the investment fund by setting a general mandate. It therefore refers to the fund's policy and enables knowledge of its priorities. The impact thesis is then broken down into resources to mobilise and actions to be taken.
- The theory of change[5] (ToC) has a similar logic but applies to a specific investment. It enables the establishment of the causal chain for each investment that links the resources mobilised to the expected impacts.

The impact thesis therefore provides the general framework for the fund's action, while the theory of change enables the formulation of expectations for each investment. The expected impacts for the various investments produced in each theory of change must be aligned with the fund's impact thesis.

1.2.2. Ex-ante vs. Ex-post Evaluation

Another important distinction for impact evaluation rests on the difference between ex-ante and ex-post impact evaluations.

- Ex-ante impact evaluation seeks to identify, during the study of the investment file, the expected non-financial effects. This approach draws on working assumptions to quantify the possible impacts of the investment. It enables justification of the relevance of a financing file for non-financial dimensions and is thus essential for project appraisal.
- Ex-post impact evaluation takes place once the investment has been made in order to identify the various non-financial effects. It enables understanding of how the investment has concretely changed the situation.

Like financial analyses (expected cash flows for financial analysis, non-financial effects for impact analysis), the two approaches are complementary. In the context of this report, we focus on ex-post impact evaluations. Nevertheless, we bear in mind the importance of also having an ex-ante analysis (cf. Box 1).

1.2.3. An Impact, Many Impacts

The use of the singular for impact can be a source of confusion, as there are multiple impacts for each intervention, whether a public policy or an investment. We propose several important typologies to keep in mind.

[5] "Theory of change" is the most common term, but this approach is sometimes referred to as a logic model or a results chain. There are some subtle differences between these approaches, but they are not essential to our discussion.

1.2.3.1. Nature of Impact

This is undoubtedly the most well-known distinction. Impact can be multidimensional and affect various dimensions. Reference is generally made to:

- Economic impacts (jobs generated, economic activity, tax revenues, etc.)
- Social impacts (improved health, education, inclusion of populations, etc.)
- Environmental impacts (reduction of greenhouse gases and pollution, biodiversity, etc.)

However, an intervention may have effects beyond these dimensions by acting, for example, on governance aspects.

1.2.3.2. Mission Impact vs. Modality Impact

Mission impact corresponds to the fundamental objective of the company, i.e. the *raison d'être* of its activity. This is the direct contribution of the company to a societal or environmental challenge, independent of how it operates.

Modality impact concerns the way in which the company conducts its activities to achieve its mission. This encompasses its internal practices, its logistical, social or environmental choices, which can amplify or harm its overall impact.

To better understand the distinction, consider a textile company. This company may have an impact through the mission it sets itself, for example by seeking to reduce the ecological footprint of the clothes produced. It may also have an impact through the way it works, for example by favouring workers in reintegration and offering decent working conditions. In other words, a mission-driven company generally aims for impact, but this can be reduced (or amplified) by its practices. Conversely, a company without an explicit impact mission can have a strong impact through its operations.

1.2.3.3. Direct vs. Indirect vs. Induced Impact

A less well-known dimension consists of distinguishing direct effects from indirect effects. Direct effects are those that apply to the entities targeted by the intervention (households, companies).

However, limiting oneself to direct effects does not exhaust the reflection, since investments can have impacts beyond the explicit targets. It is therefore important to look at these additional effects, which can sometimes be more important than direct impacts. It is useful to distinguish between indirect effects and induced effects.

- Indirect effects are impacts generated along the value chain (upstream and downstream). For example, it is not uncommon for an investment to generate a limited direct employment effect but a strong indirect employment effect by stimulating the activity of suppliers and customers. Conversely, the gains obtained by the supported company may come at the expense of competitors, which tends to reduce the overall effect on employment.

- Induced effects are externalities that are not generated along the value chain but through other channels. For example, an investment may foster a better situation for the dependants of the company's employees.

It should be noted that the term "indirect effect" can cover the full set of indirect effects (in the strict sense) and induced effects. In all cases, it is important to take these investment externalities into account when measuring non-financial implications.

1.2.3.4. Positive vs. Negative Impact

A frequently overlooked dimension concerns the consideration of potentially negative effects. For example, the increase in production enabled by an investment may also induce harmful effects, such as soil artificialisation to extend the production unit or increased pollution. These negative effects may be marginal or substantial. In all cases, it is important to be aware of these negative impacts and not to limit the analysis to positive effects.

1.2.4. Impact of an Investment vs. Impact of a Fund

One final point deserving particular attention is the distinction between the impact of an investment and the impact of a fund.

- The impact of an investment encompasses the entirety of non-financial effects (economic, social, environmental) generated by an investment. This is the approach adopted in the academic literature on impact measurement (of public policies), often referring to a particular intervention in order to compare situations with and without intervention.
- The impact of an investment fund aims to evaluate the non-financial effects induced by the fund's entire investment portfolio.

Moving from one to the other would be easy if all the fund's investments generated similar, comparable impacts. In practice, however, this aggregation is more complex because each investment can have effects that differ considerably from one another. For example, investment A may help improve health indicators while investment B will have the primary impact of reducing water pollution. The analyst's challenge is therefore to provide a synthetic presentation of the fund's impacts. As we develop later in this report, methods exist for this.

1.3. The Four Steps of Impact Evaluation (of an Investment)

This section examines the implementation of impact evaluation, posing the question of how. The idea is to present the main steps of impact evaluation. Impact evaluation requires going through four successive steps, which we detail below. A fifth step, which is not directly related to impact measurement, is equally important: the use by the fund of the evaluation results, which we address in the following section.

1.3.1. Defining Expected Impacts

The first step consists of describing the impacts expected from the investment. The most common approach for linking the investment to impacts is to draw on a theory of change (ToC), whose general principle was mentioned earlier (subsection 1.2.1.).

The theory of change is a decomposition of the process of materialisation of the effects of an investment from the invested funds to the impact, identifying the various transmission chains of expected returns. Constructing a theory of change rests on several key steps. First, the long-term impact one wishes to achieve must be defined. Next, the intermediate results necessary to reach this impact must be identified, before describing the concrete actions that will be implemented to generate them. Finally, indicators are selected to measure progress made.

The theory of change is structured around four pillars as follows:

Inputs → Outputs → Outcomes → Impact

- Inputs represent the resources mobilised, notably the financing provided by the investor;
- Outputs are the products or services generated by these resources;
- Outcomes correspond to the changes observed in the short and medium term among beneficiaries;
- Impact designates the lasting and profound effects in the long term, such as poverty reduction or community empowerment.

This approach must also consider all stakeholders and potentially negative and indirect effects. A completely exhaustive approach is impossible, but it is important not to neglect the most important effects/actors.

Implementing a theory of change thus enables the decomposition of expected interactions and the identification of possible impacts. It also enables the identification of risks that may explain the absence of transmission between each element of the chain. In the example presented below, it is possible that the vehicle purchase (output) does not convert into better medicine distribution in rural areas (outcome) for various reasons: difficulty recruiting drivers, weakness of road infrastructure, rise in fuel costs inducing additional unforeseen expenses, insecurity in certain areas, etc.

Example 1: The Theory of Change

In this section, to make the elements presented more concrete, we draw on a fictitious case study that serves as a thread throughout. We consider an investment fund whose impact thesis consists of acting on the health of populations by supporting companies in the medical and paramedical sectors. In this context, this fund invests to support a pharmaceutical distribution company that operates primarily in rural areas. The fund's investment is intended to finance the purchase of more efficient refrigerated vehicles so as to expand the company's radius of action and thus supply medicines to pharmacies or health centres distant from urban centres. The theory of change in the context of this investment can be constructed as follows:

Input: Fund investment; Output: Purchase of refrigerated vehicles; Outcome: Increase in the medicine distribution radius; new health centres and pharmacies supplied with medicines; Impact: Improvement of health indicators in newly reached areas.

1.3.2. Defining, Collecting and Monitoring Indicators

From the theory of change, it is possible to characterise the indicators to collect to measure the various dimensions (outputs, outcomes and impacts). The definition, collection and monitoring of indicators require consideration of several aspects from the outset.

First, the selected indicators must be as close as possible to what one wishes to measure. In particular, for outcome indicators and even more so for impact indicators, it is sometimes difficult to explicitly define the indicators to consider. Take the example of an indicator often cited in impact fund reports: the number of jobs created. Some funds capture only direct job creation; a share separately presents direct and indirect creation, while others cumulate both into a single indicator. The same observation is made for full-time and part-time employment. To facilitate this work, there are initiatives to harmonise the main indicators retained by impact investors, such as IRIS+ (Box 3). Nevertheless, it is important to ensure that the metric considered is as close as possible to the aspects addressed by the theory of change.

A second aspect to consider concerns collection and monitoring procedures. The more precise and qualitative an indicator, the more difficult it risks being to collect the information at a high frequency. In the case of employment, if the analysis focuses on decent jobs, this implies measuring not only the number of jobs but also the conditions in which workers exercise their work and their remuneration. These data can be collected through employee surveys, which implies a lower frequency than if the evaluation only concerns total employment. There is therefore often a choice to be made between the depth of the indicator considered and the capacity to collect and monitor it. There is no general rule in this case. What matters is knowing what the priority is. If the emphasis is on the volume of jobs rather than the conditions of their exercise, then a more frequent but less precise measure may suffice. On the other hand, if the investment's primary objective is to improve working conditions, a better-adapted indicator will be needed, even if this means reducing the frequency of collection and monitoring.

Box 3: The IRIS+ Initiative

IRIS+ is the reference system developed by GIIN to help investors integrate social and environmental factors into their decisions and better measure their impacts. Free and continuously updated, it offers a set of metrics through the IRIS+ Core Metrics Sets where, for each indicator, a guide is allocated comprising its definition, collection or calculation advice, suggestions for combinations with other metrics, or the associated SDG. The definitions are based on the IRIS+ Taxonomy which offers widely recognised definitions. Its structure, divided into categories, enables the treatment of various themes (environment, education, health, governance, etc.), sub-themes (climate, air, forest, etc.) or dimensions (what, who, how, etc.). Each definition is accompanied by a set of information enabling the understanding of key terms, such as the conventions/standards/principles with which it is aligned or the objectives and project types to which it relates. In sum, IRIS+ provides investors with a practical and harmonised framework for defining their indicators, collecting data and communicating transparently on their impacts. Webinars and demonstrations are offered by GIIN to make the tool more accessible.

Finally, the collection and monitoring of indicators can draw on several sources. Schematically, there are two main sources:

- Internal (or primary) sources are those that come directly from the parties involved in the financing, notably the financed company but also other stakeholders (suppliers, customers). These primary data can be collected specifically for the impact evaluation or for other reasons but then used during the evaluation.
- Secondary sources encompass all external data that can be mobilised to conduct the evaluation. These data are particularly useful for constructing the counterfactual (see below). They can be extracted from different sources grouping together existing surveys, administrative data or more original data (such as satellite data). These external data are not collected to carry out the impact evaluation. In other words, unlike primary data specially calibrated to answer the evaluation, secondary data do not have the same status and the collected indicators are not always identical to those provided by primary data.

Example 2: Measuring and Monitoring Impact Indicators

In our fictitious example, the investment should improve the health indicators of populations living in areas newly supplied with medicines. It is possible to collect general health indicators such as the number of hospitalisations or morbidity and even mortality indicators. It is also possible to go further by drawing on the medicines distributed. For example, to observe whether the investment enables improvement of the distribution of medicines, particularly those sensitive to the cold chain such as injectable vaccines or insulin. Thus, the observed effects should be even more pronounced for the prevalence of the diseases concerned (diabetes, malaria, etc.).

Impact analysis can sometimes be constrained by the ability to collect primary data or the availability of secondary data that do not allow for the capture of the indicators that the evaluation wishes to measure.

Nevertheless, it should be noted that increasingly frequent access to administrative or survey data, as well as the use of original data (such as satellite data), tends to reduce this problem.

1.3.3. Impact Attribution

Impact attribution aims to ensure that the change observed in the chosen indicators is due to the investment made. A naive approach consists of comparing the indicator(s) considered before and after the investment, assuming that the entirety of the observed change is solely due to the intervention. This approach is, however, imperfect, as many events may have affected these indicators positively or negatively over time.

Example 3: Constructing the Counterfactual

To analyse the impact of the fund's investment, it is decided to study the evolution of hospitalisations due to malaria crises in health centres. Indeed, the deployment of refrigerated vehicles has enabled better distribution of anti-malarial vaccines through better cold chain maintenance. It is expected that this will materialise as fewer acute malaria crises in these areas. To carry out the impact analysis, the evaluators have access to the hospitalisation register of all the country's health centres (hospitals, dispensaries, etc.) over several years. They can thus identify the total number of hospitalisations due to malaria crises. Their objective is therefore to see whether this indicator has been reduced following the investment through better medicine distribution.

A simple approach would be to compare the evolution over time between newly served health centres and all other health centres in the country. This approach is better than a simple before-after comparison because it allows controlling for global factors (such as the implementation of a public policy). However, the health centres that benefited from the investment are particular because they are in rural areas and often less well equipped than urban hospitals. To account for this, a relevant counterfactual analysis will select health centres comparable to those that benefited from the extension of the company's activity.

Impact evaluation, to overcome this limitation, draws on a central notion: the counterfactual. The counterfactual is a hypothetical situation that would have occurred in the absence of the investment. The entire difficulty consists of reconstructing this situation which has, by definition, not occurred. The objective is to be able to compare the observed situation with this theoretical situation. To do so, analyses draw on the evolution observed within a group similar to that of the investment's beneficiaries (see Example 3 above).

The scientific literature on impact evaluation of public policies has developed many approaches to reconstruct the counterfactual. The common thread between these methods (Box 4) is the fact of not limiting the analysis to the beneficiary entity but of considering a so-called control group. In the absence of a control group, it will be very difficult to say whether the observed evolution (positive or negative) is due to the investment or to other unobserved factors (a public policy, for example). An essential point to retain is that a credible impact analysis cannot be constructed without recourse to data from other

entities that have not benefited from the investment. This aspect is essential for the collection and monitoring of indicators mentioned above.

Box 4: Impact Evaluation Methods

The academic literature proposes several methods for constructing a counterfactual and thus evaluating an impact. This counterfactual is also called a control group, in opposition to the treatment group formed by beneficiaries. The approaches differ in the way the counterfactual is identified.

Randomised Controlled Trial (RCT): The randomised controlled trial is a method that randomly constructs treatment and control groups. It is the random assignment that ensures the similarity of the groups and the extraction of impact by comparing them.

Matching Methods: Matching methods rely on the comparison between beneficiaries and non-beneficiaries with respect to a set of observable characteristics before intervention. There are several matching approaches that differ in how units are compared and how comparable non-beneficiaries are selected for each beneficiary.

Regression Discontinuity: Regression discontinuity exploits an eligibility rule that creates a threshold separating beneficiaries from non-beneficiaries. The method is applicable if units situated just below and just above this threshold are comparable, which can be verified graphically.

Difference-in-Differences: Difference-in-differences identifies the control group as the group of non-beneficiaries that, ex-ante, exhibits a similar evolution of indicators to that of the treatment group. The method assumes that, in the absence of the intervention, the two groups would have followed parallel trends.

Synthetic Control Method: The synthetic control method artificially constructs the control group from several units that are more or less comparable ex-ante to the treatment group. This approach is useful when no single control unit is, by itself, able to closely resemble the treatment unit.

When methods are properly implemented — i.e. when the assumptions they impose are respected — impact is calculated by comparing the average impact indicators of the two groups.

The methods presented above come from the literature on public policy evaluation. Applied to the context of investment funds, they encounter an additional difficulty specific to the nature of private investment: the question of additionality (Carter et al., 2021).

In the case of a public policy, the state is in a position of sole intervention. If the policy is not implemented, no other actor substitutes for it. Additionality is therefore verified by nature: the counterfactual is the complete absence of intervention. Evaluation methods can thus attribute the entirety of the observed change to the policy studied.

For an investment fund, this assumption no longer holds. If the fund does not invest in a company, it is possible — even likely in certain cases — that another funder, public or private, would have taken over. It then becomes very difficult to attribute the entirety of the observed impact to the fund itself: part of that impact might have occurred anyway.

This reality has two important implications for the impact evaluation of funds. On the one hand, it invites a distinction between two questions that are often conflated: "*What is the impact of the investment?*" and "*What is the fund's own impact?*". The first refers to the notion of counterfactual and what the investment has induced as change. The second relates to the notion of additionality, which comes in addition to the equation for a fund that must ensure that its contribution enabled the investment to take place.

This discussion is a reminder that additionality is not a given but a hypothesis to be verified, and that claiming the entirety of a company's impact without having asked this question amounts to overestimating one's own contribution. This difficulty explains why funds often apply an attribution rate (cf. Boxes 5 and 7) starting from the principle that only a share of the impact is due to their intervention.

1.3.4. Cross-Checking Analyses to Refine Results

The presentation of the three first steps of impact evaluation draws on a primarily quantitative approach in which indicators are quantified and compared with the counterfactual. Nevertheless, to properly understand the results obtained, this evaluation should be crossed with other analyses.

Example 4: Validating the Different Stages of the Theory of Change

The impact analysis has made it possible to highlight that the targeted areas experienced a reduction in the prevalence of malaria crises. To ensure that the observed effect is due to the investment, it is important to reconstruct the logical chain, i.e. in this example:

Evaluate whether the company increased its fleet of delivery vehicles (output); whether it distributed more medicines to targeted areas (outcome). These analyses do not necessarily require the same empirical rigour but are important to present to ensure that the observed effects are genuinely due to the investment.

Suppose we had not been able to observe an impact. This analysis could have identified the possible reasons for this absence of effect by decomposing possible blockages: inability to expand the vehicle fleet (too high a cost, insufficient supply, etc.); inability to deploy the fleet to reach targeted areas (deficient road infrastructure, lack of drivers, etc.).

On the one hand, it is important to reconstruct the causal chain from the investment to the impact to ensure that the observed effect is logically linked to the theory of change. It is expected that inputs will have effects on outputs and outcomes before having an impact. If the analysis does not confirm these first two links, it will become difficult to argue that the investment had the expected effect, or that it

was through a different channel than the one specified in the theory of change. This approach is particularly useful when the investment has not had the expected effects. Indeed, by reconstructing the logical chain, it is possible to understand the blocking points and act on them in the future (see Example 4 above).

On the other hand, it is useful to go beyond quantitative analysis to better understand how the investment induced changes. For this, a qualitative analysis is often welcome (see Example 5). It enables a better understanding of how changes came about, but also the identification of effects not initially perceptible. Furthermore, this enables the deepening of obtained impacts.

Example 5: Understanding Impact Through a Qualitative Approach

The impact analysis of the investment in the pharmaceutical distribution company highlights a positive effect not only on acute malaria crises but also on other health indicators. However, the company used the funds to increase the distribution of anti-malarial vaccines only. A more qualitative analysis is therefore implemented to understand this unexpected (positive) effect. Surveys of households living near the newly served dispensaries reveal that individuals took advantage of the vaccination appointment to address other health problems. Thus, other diseases could be treated, improving the overall health of the target populations.

1.3.5. Impact Measures and Financial Performance Measures

Before going further, it is useful to highlight the specificities that characterise impact measurement relative to financial performance measurement. Indeed, in reviewing impact reports from impact investors, it is possible to notice that great importance is given to the financial performance of funds and their beneficiaries (amounts invested, revenues generated). While their measurement is necessary, it does not replace impact evaluation. As stated above, the two processes differ on several points.

First, the object of evaluation differs. A financial performance measure assesses the company's performance. It therefore evaluates what the company can act on and what it directly produces. Considering the ToC, this corresponds to outputs, and perhaps outcomes if the evaluator is a fund wishing to monitor the performance of investee companies. Impact measurement analyses the effects on society and the environment and therefore considers the ultimate purpose of the ToC, which is impacts. Thus, depending on the chosen approach, different points of the ToC are evaluated.

Second, the processes specific to each type of measurement differ in their purpose. Both involve the definition of objectives as well as the measurement and monitoring of indicators. The financial performance measurement process is completed by their presentation or temporal comparison. That of impact measurement substitutes this with the use of methods to extract the causal effect. Thus, although the processes are similar at certain stages, the way in which metrics are exploited once measured differs.

Third, the approaches differ in their temporality. Financial performance measurement can be done through a short-term analysis. By choosing to evaluate the profitability of a company, one can compare the profit of one period with that of the next. An impact takes time to materialise, particularly for the environmental dimension. To conduct a rigorous impact evaluation, it is necessary to conduct a medium/long-term evaluation by analysing indicators over several periods.

Fourth, there is a specificity regarding the definition of metrics. Indeed, financial performance indicators are standardised (their calculation formulas are often international), while the definition of impact metrics can vary from one evaluator to another, as illustrated by the discussion around employment measurement. Thus, two funds evaluating their impact on employment may arrive at results that are difficult to compare, since the indicators actually capture two different things. We note, however, that in recent years, initiatives to standardise impact indicators have emerged and developed to address this problem.

Finally, another particularity concerns the nature of the evaluation object. For financial performance measurement, this object is initially quantifiable and monetary. Impact analysis often rests on qualitative, multidimensional or more difficult-to-quantify elements than financial data.

The cited differences make impact measurement and financial performance measurement non-substitutable processes. On the contrary, they are complementary, each providing different and useful information to the evaluated structure.

1.4. Using Impact Evaluation Results

Impact evaluation is only useful if it is fully exploited by the investment fund. As discussed above, the rationale for impact evaluation is to improve the fund's operations and to prove that it has generated impact. In this section, we revisit this dual purpose of impact evaluation. For the second justification, we distinguish between the analysis conducted at the investment level and that conducted at the fund level as a whole, although the two are closely linked.

1.4.1. Improving Practices: Comparing Results with Expectations to Draw Lessons

One of the main rationales for impact measurement is to improve the fund's operations by drawing lessons from what has worked and what has not worked as well. To do so, it is important to compare the results obtained and updated by the ex-post evaluation with the expectations during the file appraisal (ex-ante evaluation). Comparing these two levels of analysis is thus paramount in order to identify any significant discrepancies, positive or negative, between expectations and results.

This comparison implies that the metrics retained for the impact evaluation are similar, or at least comparable, between ex-ante and ex-post analyses. After identifying the differences, it will be necessary to interrogate those deemed significant. The in-depth analysis, presented in the previous section, is an indispensable step for uncovering what happened. Indeed, it will be useful to (i) explore the various links of the theory of change to analyse transmission mechanisms and (ii) draw on a more qualitative approach. These elements will enable the identification of factors that have influenced impacts. This analysis therefore requires good knowledge of the context in which the investment took place.

The feedback will only be useful if the lessons to be drawn are generalisable to other investments. It is possible that the non-achievement of expectations is due to an idiosyncratic phenomenon (e.g. natural disaster or civil conflicts) that is difficult to foresee and over which the fund has little control. On the other hand, it is possible that the analysis reveals more endogenous factors that could reappear in future investments. In this case, the analysis is particularly important.

It is important to emphasise that a failure can only serve to nourish future successes if the lessons drawn from it are fully exploited. This step can unfortunately be neglected because funds do not always seek to highlight their failures. Yet this is a mistake, as the risk is to multiply failures for similar reasons.

1.4.2. Proving Impact at the Investment Level: Financial and Non-Financial Returns

The second interest of impact evaluation consists of proving impact, i.e. enabling the presentation to stakeholders of the fact that the investment has had non-financial effects. Among the stakeholders, capital providers are particularly sensitive to this, as they accept reduced financial returns in exchange for notable social or environmental impact.

Box 5: Presentation of the SROI

The Social Return on Investment (SROI) is a method for valuing impact that seeks to express the non-financial effects of an investment in monetary terms. By assigning a monetary value to effects that are not traded on a market — improved health, reduced unemployment, welfare gains — the SROI enables the cost of an investment to be weighed against the social value it generates, akin to a classical financial return.

The SROI is calculated as the ratio between the net present value of the impacts produced and the total cost of the investment:

$$SROI = \frac{Net\ present\ value\ of\ impacts}{Total\ cost\ of\ investment}$$

An SROI of 3 means that for every €1 invested, €3 of social value has been created. To avoid overestimating impact, four corrective factors are applied to each identified effect before calculating their aggregate monetary value:

- Deadweight refers to the share of the effect that would have occurred anyway, in the absence of the intervention. It directly refers to the notion of additionality discussed in this report.
- Attribution measures the share of the effect genuinely attributable to the intervention, after deducting the contribution of other actors who acted simultaneously.
- Drop-off reflects the erosion of the effect over time: impacts produced in one year do not necessarily persist at the same level in subsequent years.

Once these adjustments have been applied, impact flows are discounted using a discount rate to bring them to current value, as with any financial return calculation.

Despite its appeal, the SROI raises several important difficulties. The monetisation of certain effects is delicate, even contestable. Furthermore, the adjustment factors rest on assumptions that can vary considerably from one study to another, making comparisons difficult. The final ratio can thus give a false impression of precision, masking the uncertainty inherent in each step of the calculation. Finally, the SROI does not fundamentally resolve the counterfactual question: it makes it an assumption — the deadweight — rather than an object of rigorous measurement, which distinguishes it from the quasi-experimental methods presented in Box 4.

In order to compare financial returns and impact, it is useful to have comparable metrics. The Social Rate of Return (SRR)[6] enables this comparison. The SRR is similar to an Internal Rate of Return (IRR) except that the profitability of an investment is not measured from the generated financial flows but by the non-financial benefits induced by the investment. The rate of return is defined as the rate that allows for a zero net present value. The formula is identical to that of an IRR except that the net present value

[6] As noted above, the adjective "social" refers more specifically to the benefits to society as a whole. The TRS therefore encompasses all non-financial dimensions. Specific terms can be used to distinguish between the different dimensions (for example, environmental rate of return).

is not constructed from cash flows but from reconstituted flows induced by the non-financial impacts generated.

The SRR can be added to the IRR to know the total rate of return of the investment. Thus, an investor can justify a reduced IRR in exchange for a higher SRR.

There are multiple approaches to measuring this social rate of return, of which the SROI is undoubtedly the best known (Box 5). Despite differences between methods, establishing the social rate of return implies three elements:

I. Attributing the observed non-financial changes to the investment;
II. Establishing the duration of the effect;
III. Monetising these impacts.

Impact analysis enables a response to the first step by offering an approach for evaluating the share of the observed impact induced by the investment in light of the considered counterfactual. The duration of the effect can be taken into account in the impact evaluation, although this is not mechanical. Its effect on the total value of the monetised effects will depend on the chosen discount rate[7].

On the other hand, the monetisation of measured impacts is not directly obtained from impact evaluation. It is, however, an essential dimension. To be convinced of this, suppose an investment of €1,000,000 that enables a reduction of 10,000 tonnes of CO_2[8] over ten years[9]. Starting from a price of €60 per tonne of CO_2, the social rate of return is 6.3%. Now suppose the tonne of CO_2 is valued at €30; the SRR drops sharply to 3.5%, significantly reducing the profitability of the investment. Conversely, if the tonne of CO_2 is valued at €90, the SRR rises to 9.2%. The monetisation of impacts is therefore a crucial step for moving from impact evaluation to its use in constructing a social rate of return that can notably be weighed against the IRR. As reflected in the literature on impact assessment, there are numerous approaches to monetizing economic, social, and environmental effects. The most common methods for monetizing impacts are based on individual preferences. There are two main approaches that rely on these preferences (Box 6). This range of methods highlights that no single approach is perfect, despite the validity of each in certain contexts. It is not uncommon for these approaches to yield figures that vary significantly from one to another. Thus, impact monetization should rely on the approach that appears most credible and best suited to the context. Indeed, much like impact assessment, it can be costly to develop a credible approach for monetizing an impact. One option is to

[7] The discount rate is the rate used to convert the value of future cash flows into present value. For example, if an investment generates $100 in cash flow next year, its present value cannot be $100 (if only because of inflation). The discount rate allows us to convert this future value into a present value. If the rate is 2%, then the 100 from next year is worth 98 today (100/1.02). Conversely, if the discount rate is 10%, then next year's 100 is worth only 91 today (100/1.10). This example highlights the importance of this rate in determining the weight given to future cash flows. The higher the discount rate, the lower the present value of future cash flows will be.
[8] In practice, an investment has multiple impacts, and the IRR is calculated by summing the "cash flows" associated with these various impacts. It is also important to factor in any potential negative effects to get a complete picture.
[9] This example also confirms that the effect of time is relatively minor, as the TRS remains largely unchanged whether we consider a 5-year or a 15-year period.

rely on assessments produced by others, as proposed by certain initiatives such as EVRI[10]. Nevertheless, one must be aware of the limitations of the chosen approaches and the importance of context.

Box 6: Monetisation Methods

Contingent Valuation Approach (Principle: Ask people for the subjective valuation of the good or service):

- Willingness to Pay: Directly asking people how much they are willing to pay or what compensation they consent to receive for a given good or service.
- Valuation Game: Directly proposing to people a set of goods and services whose prices are known, including the good to be valued. The ranking of the value of these products gives the range of possible values of the good of interest.

Revealed Preference Approach (Principle: Exploit behaviours to reveal preferences):

- Hedonic Pricing: Defines the value of a good from the costs of goods that interact with it.
- Travel Cost Method: Uses the travel cost incurred to access a good or service, for example a nature reserve.
- Restoration Cost Method: Uses the expenditure incurred to restore an ecosystem, after a natural disaster for example.
- Replacement Cost Method: Uses the amounts spent to reproduce an ecosystem that must be relocated elsewhere.
- Avoided Cost Method: Uses the costs that a good or service enables to be avoided to value it.

Other Approaches:

- Production Function Approach: Assesses a natural factor from the production it enables.
- Benefits Transfer: Recommends using value estimates from previous studies in similar contexts.

[10] The Environmental Valuation References Inventory (EVRI) is a database of studies providing economic values for natural assets and health. The tool was developed by Environment and Climate Change Canada—Canada's environmental ministry—along with experts and environmental organizations in the early 1990s. EVRI was implemented to facilitate the transfer of valuation data from one study to another. To help users easily navigate the more than 5,000 study summaries on the site, searches can be filtered by several criteria, including the continent or country where the study was conducted, the asset being valued, and the valuation technique used. More information at: https://evri.ca/en.

1.4.3. Proving Impact at the Fund Level: Aggregating Impacts

Impact funds are interested not only in the evaluation of their various investments but also in the total impact generated by the fund. Moving from an evaluation of individual projects to that of the portfolio as a whole therefore implies aggregating the impacts of each investment. This step is rarely easy due to the fact that each investment can induce very different impacts from one another. In this case, a common unit must be used for these various impacts in order to aggregate them. The most common approach consists of monetising the impacts.

1.4.3.1. Aggregating Similar Impacts

Before dwelling on the case of investments generating different impacts, it is useful to address the simpler case of a portfolio of investments generating comparable impact. It is particularly common for investment funds operating in Africa to have an objective of creating decent jobs. In other words, most impact evaluations present this result by identifying jobs created thanks to the investment. From this, aggregation is in principle relatively easy since it is simply a matter of summing all the jobs created by the investments. This situation can also be found for specialised funds that can thus have a common metric for all their projects (tonnes of CO_2 emissions avoided for a climate fund, for example).

While using a common metric for all evaluations facilitates aggregation, it is necessary to ensure that the selected indicator in each case is identical. The case of employment illustrates this difficulty. Funds generally seek to generate decent jobs. It is therefore necessary to agree on what a decent job means and how to qualify created jobs as decent. Moreover, created (decent) jobs can be temporary or permanent, full-time or part-time. Thus, simple aggregation becomes complex when these various dimensions come into play.

1.4.3.2. Aggregating Different Impacts

The difficulty of aggregation is greater when the impacts of the different investments in the portfolio differ from one another. Most funds operating in Africa are generalist funds (Léon and Rabary, 2024). They are thus led to intervene in varied sectors such as agriculture, health, environment or education. The impacts generated in these sectors are therefore also very different. It is therefore essential to find a means of bringing these effects together with a common unit. The most common approach is monetisation. Money is by definition a unit of account that enables comparison of goods or services with different values. Just as money enables the comparison of the price of fruit or a haircut, it offers the possibility of constructing equivalences between a tonne of CO_2 avoided and a year of life in good health. Of course, monetisation is an imperfect and debatable process, but it enables the gathering of the various generated impacts into a single figure.

An alternative approach consists of identifying representative indicators for each dimension of impact. This approach is often the one retained in impact reports because it enables the avoidance of the arbitrary choices that monetisation requires. It is in some sense an extension of the previous approach but considering different key indicators for each dimension. This approach therefore implies being able to reconstruct these few key indicators for each investment. This approach ultimately amounts to aggregating identical impacts but across several dimensions.

Part 2 – Evaluating the Practices of Impact Investors in Africa

The second part of this report focuses on the practices of impact investors operating in Africa with regard to impact evaluation. To do so, we have constructed an original database on the effective practices of funds. Before discussing the methodology and results, it is useful to highlight the originality of this work given the absence of similar studies.

2.1. Literature Review

Impact evaluation is one of the most studied fields in academic economics. There are very many articles on the evaluation of public policies. While work on this question on impact investors is less dense, it is by no means negligible. A simple search on an academic search engine reveals nearly 2,500 studies on the subject since the early 2000s, although Africa is very underrepresented in this research[11]. Furthermore, these works address multiple questions and there are few studies on the practices of impact investors.

We have identified a handful of articles evaluating what is done in terms of impact evaluation by ecosystem actors seeking to generate positive effects: impact investors without geographic restriction, social entrepreneurs, ethical banks.

The two works closest to ours are those of Loveridge (2016) and Roor and Maas (2024). Loveridge (2016) reviews impact measurement practices based on a documentary analysis and expert consultations. Roor and Maas (2024) conducted a systematic review of 141 scientific articles on impact measurement of investments. These two studies highlighted that the most widely used impact measurement tools are analytical frameworks, the ToC, reporting frameworks and the SROI. Thus, the most widely used tools relate to the stages of objective definition and use of results in the impact evaluation process. Regarding impact attribution, it has been observed that investors generally resort to simple before-after investment or reality-target comparisons (Jalila, 2022; Kocornik et al., 2021; Salazar et al., 2012). It has also been noted that measurement focuses primarily on inputs and outputs such as the resources mobilised or the number of investments (Ormiston and Seymour, 2011; Salazar et al., 2012). Loveridge (2016) reports the result of a GIIN survey: impact objectives are very important to investors, but only 15% of the 150 respondents affirm evaluating them.

Existing work often comes from analyses conducted on small samples (fewer than twenty observations), not representative of all actors. Moreover, to our knowledge, there is no work focusing on the impact investing sector in Africa. Our research proposes to fill this gap by studying the impact evaluation practices of all impact investors operating on the continent. To do so, we construct an evaluation grid and a database reflecting the practices of impact funds, which we cross with the characteristics of these funds.

[11] The search was conducted using OpenAlex with the following keywords: Impact, Investing, Measure, Fund. It identified 2,401 works published between 2002 and 2025 (query run in February 2026). When narrowing the focus to Africa, the number of results drops to 255 (with many articles falling outside the scope of this study).

2.2. Methodology

2.2.1. Data

Sample Selection

Our analysis covers the 250 impact investment funds operating in Africa identified by Léon and Rabary (2025). This set was constructed according to two criteria. First, to be retained, a fund must have activities in Africa with at least one investee company, without necessarily being exclusive to the continent. Second, the activities of the structure must meet the definition of impact investment. Concretely, investors must explicitly seek a financial return as well as social and/or environmental impacts, have the intention to measure their impact or already conduct this practice, and finally have a team dedicated to the fund.

Constructing the Grid for Evaluating Practices

The literature and existing tools for assessing the impact evaluation process primarily propose frameworks aimed at supporting implementation or guaranteeing the quality of procedures, with a strong emphasis on operational aspects. Some works address common themes such as the theory of change, transparency or the diversity of results, but few offer a framework specifically designed to retrospectively analyse the impact measurement process as a whole. The question of method choice, their level of rigour and the use of a counterfactual is rarely addressed. Different instruments evaluating the impact measurement process are discussed in more detail in the Annex.

To fill this gap, we propose a grid structured around all the stages of impact evaluation. It rests on a binary scoring system, designed to enable statistical and econometric analyses, which limits subjectivity but does not allow for the consideration of the intensity or regularity of practices.

We define five categories, covering the different stages of impact evaluation:

1. The general analytical framework
2. The definition of indicators
3. The approaches used to attribute impact
4. The different types of results reported
5. The use made of results and transparency

For each, we define a set of practices, based on binary questions (yes/no) that allow them to be assessed. By summing these points, we obtain a total score specific to each structure as well as scores relative to each category. The scores of each category are normalised in order to be compared, since the categories do not all have the same number of questions. Thus, for each of them, all funds have a score between 0 and 1. From the perspective of the categories, we calculate normalised average scores for each one. Below, we list the questions comprising the evaluation grid according to the five dimensions.

Analytical Framework

For the analytical framework, four questions are defined:

(i) Is the Theory of Change (ToC) made explicit?
(ii) If yes to (i): Is the ToC detailed through a diagram or a more precise explanation of each stage of the ToC?
(iii) Are the fund's targets presented in terms of geographic area, sector, sub-population (women, youth, farmers, people from rural areas, etc.) or type of structure?
(iv) Is the intention to measure impact stated?

Definition of Indicators

For indicators, we list eight questions:

(i) Are indicators clearly defined?
(ii) If yes to (i): Is this the case for outcome indicators?
(iii) If yes to (i): Is this the case for impact indicators?
(iv) If yes to (i): Is this the case for indicators relating to social and environmental dimensions?
(v) If yes to (i): Is this the case for indicators relating to indirect effects?
(vi) If yes to (i): Is this the case for indicators capturing medium or long-term information?
(vii) Are the indicators used harmonised, i.e. do they come from existing standardisation initiatives in the impact investing sector?
(viii) Is the indicator monitoring framework made explicit in terms of format (interviews, survey, reading of reports, etc.), level of monitoring (company, household, individual, etc.) and frequency?

We return in greater detail to certain questions in this category. If the fund under review presents at least some indicators relating to inputs and/or outputs — such as capital, the number of employees within the fund structure (as opposed to the number of employees in investee companies), or the number of projects financed — we award one point for question (i). If this is also the case for outcome indicators, one point is awarded for question (ii). We define impact, in part, as a lasting effect — that is, a medium- to long-term effect. Accordingly, an outcome indicator that has been designed to be measured over several periods will be counted as an impact indicator, resulting in a positive response to questions (iii) and (vi). Indeed, a fund whose investments enable the creation or preservation of jobs over several years among its beneficiaries contributes in a sustained manner to local employment and growth. It is possible that a fund may define output indicators that are tracked over several periods; in such cases, points are awarded for questions (i) and (vi). Where multiple indicators are presented but only some are clearly defined, the point is nevertheless awarded for the corresponding questions.

Impact Attribution

The third category groups eight questions relating to the impact evaluation method:

(i) Has a quantitative approach been implemented?
(ii) If yes to (i), is this approach rigorous?

(iii) If yes to (ii), has a counterfactual been defined?
(iv) If yes to (iii), has the way in which it was constructed been made explicit?
(v) If yes to (iv), is this counterfactual of quality?
(vi) Has a qualitative approach been implemented?
(vii) If yes to (vi), is this approach rigorous?
(viii) Is a triangulation — combining rigorous quantitative and qualitative approaches — implemented?

The first five questions are designed to highlight the varying levels of rigour in the measurement of investment effects that we consider. If the fund under review presents simple before-after comparisons or observed-versus-target indicators, we respond positively to question (i) and negatively to questions (ii) through (v). If the indicators are embedded within methodologies — that is, if the fund treats them as inputs rather than as ends in themselves, without however applying methods that draw on the concept of a counterfactual — points are awarded for questions (i) and (ii) only. Where a fund explicitly references the use of a counterfactual without elaborating further on its definition, we respond "yes" to the first three questions. If, on the other hand, further details are provided on this point, points are awarded for questions (i) through (iv). All five questions yield points when the defined counterfactual is comparable to the treatment group and satisfies the assumptions underlying the approach employed. Questions (vi) and (vii) capture several levels of rigour from a qualitative standpoint. Question (vi) receives a positive response if descriptions of projects and activities are shared. If the fund goes further by conducting surveys or interviews with its clients, an additional point is awarded through question (vii).

Reported Results

The fourth section groups an analysis of communicated results:

(i) Are the results quantified?
(ii) Are the results monetised?
(iii) Are there results for the ToC outcomes?
(iv) Are there results for the impacts?
(v) Are there results relating to social and environmental dimensions?
(vi) Are there results for indirect effects?
(vii) Are there results for medium and/or long-term effects?
(viii) Does the fund present negative or nuanced results?

If a fund presents a description of non-quantified outcomes, such as in the following example — "The company has increased the income of its employees" — we respond no to question (i) and yes to question (iii). It should be noted that we incorporate information relating to ToC outcomes, impacts, medium- and long-term information, indirect effects, and social and environmental dimensions both under the Indicator category and under the Results category. Indeed, some funds may present their methodology by listing the indicators they use without disclosing the corresponding results; in such cases, points are awarded solely under the Indicator category. Others, conversely, report results without having previously defined the indicators used. One of the most commonly reported results across funds is the number of jobs created. A job may be created directly or indirectly, and may be full-time or part-

time. Where an organisation refers to employment-related results without further qualification, points are awarded under the Results category but not under the Indicator category.

Transparency and Communication

Finally, the last category also groups eight questions:

(i) Does impact-related information exist?
(ii) Is a document dedicated to impact publicly published?
(iii) Does the fund communicate on all its investments?
(iv) Is an expansion or intensification of activities planned?
(v) If yes to (iv), is this expansion/intensification plan detailed?
(vi) Does the fund implement a formative approach?
(vii) Does the structure share what it has learned from its activities with the sector, for example through participation in events or the establishment of a blog?
(viii) Are limitations regarding methods, results and/or indicators communicated?

We respond yes to question (i) if any of the following are disclosed: investment results, the expected consequences of investments (the Theory of Change), and/or information related to impact measurement. The same response applies to question (iii) if the results and/or the ToC pertain to the entire portfolio rather than to a few selected investment examples, and/or if the expected objectives and impacts are known for each investee company.

Data Collection

To collect the data, we apply the grid to the 250 funds conducting impact investing in Africa. Between May and August 2025, we visited their websites and read their annual reports if publicly communicated. Other resources such as the LinkedIn pages of the structures, press releases or, more rarely, interviews with fund managers are also used[12]. This documentary review enables responses to the thirty-six questions of the grid for each fund and the obtaining of associated information.

2.2.2. Empirical Analysis

To understand the results arising from the analysis of this database, we merge it with the FERDI impact investor mapping (Léon and Rabary, 2025)[13]. This merger enables an assessment of the existence of relationships between fund characteristics and their impact measurement practices. The characteristics used are the age of the fund, its size (measured by assets under management, AUM), its growth (measured by AUM growth between 2023 and 2024), the location of the headquarter and the African footprint. We exploit these data through a correlation analysis and an econometric approach.

[12] Press releases cover the period 2019–2025; interviews were conducted in 2024 and 2025.
[13] The database of impact investors operating in Africa is available at the following link: https://ferdi.fr/donnees/cartographie-des-investisseurs-d-impact-en-afrique

The first step consists of conducting correlation tests between the various fund characteristics and their impact measurement practices. This analysis allows a first exploration of possible relationships. To do so, we use various statistical tests depending on the nature of the variable considered.

Box 6: Statistical Tests for Correlation Analysis

To study possible associations, we use statistical tests: the Mann-Whitney test, the Spearman test and the Fisher's exact test. These tests are used in different contexts: the first when analysing a continuous variable and a binary variable; the second when considering two continuous variables; and the last when working with two binary variables. Several observations led us to choose these tests. First, the distributions of the variables do not follow a normal distribution. It is therefore not possible to use the Student's t-test (for a binary and a continuous variable) or the Pearson's test (for two continuous variables), which work on the assumption that variables follow this distribution. Second, for a certain number of variables relating to impact measurement, the groups are very unbalanced. In other words, a large share of the set can comprise one of the two categories, leaving the complementary group strongly in the minority. By way of illustration, one can cite the variable that takes the value 1 if the fund monetises its impacts and 0 otherwise: only 2.40% of the set comprises the group adopting this practice. The inverse situation can also be encountered with a variable such as the one indicating whether the fund specifies its target clientele: 97.20% of the set does so. Thus, repeatedly, when cross-tabulating two variables in a contingency table, we find a number of cells with fewer than five observations. According to the literature, below this threshold, the chi-squared test (for two binary variables) becomes less reliable (Campbell, 2007; Cochran, 1952).

The univariate analysis enables the highlighting of initial relationships, but does not account for the fact that certain fund characteristics may be collinear. To control for this, we resort to a multivariate analysis, enabling the identification of the specific effect of each characteristic while simultaneously controlling for other factors likely to influence the studied relationship. To do so, we conduct regressions whose independent variables are the fund characteristics and whose dependent variables are the scores and binary variables capturing the implementation of practices. For scores we apply a linear model (ordinary least squares, OLS). For practices we turn to a non-linear model (logit model, as the non-normality of the error terms does not allow us to use a probit approach). We work with the following model:

$$\begin{aligned}\text{Score ou P[Practice=1]} = {} & \beta_0 + \beta_1 * \text{Fund age} + \beta_2 * \text{African footprint} + \beta_3 * \text{Fund size} \\ & + \beta_4 * \text{Fund growth between 2023 and 2024} + \beta_5 * \text{HQ in Europe} \\ & + \beta_6 * \text{HQ in Americas} + \beta_7 * \text{HQ in Asia or Oceania} + \mu\end{aligned}$$

The location of the headquarters in Africa serves as the reference level for the characteristic of fund headquarters location. The headquarters locations' coefficients thus capture the difference in effect compared to a fund of African origin.

2.2.3. Limitations of the Analysis

We note two major limitations to our work.

First, we draw only on resources publicly communicated by the funds. Some structures reserve their impact reports for their investors and clients. When this is not the case, it is possible that only a portion of information is disseminated while the rest remains private for reasons of confidentiality. Strong and positive effects may be highlighted at the expense of weak or negative repercussions that remain known internally only. Thus, we are not able to count the corresponding points. Some zeros may therefore reflect not the absence of practices, but simply their non-communication to the public. Our analysis therefore reflects what funds accept to publish, with a portion of their activities potentially remaining opaque and unevaluable.

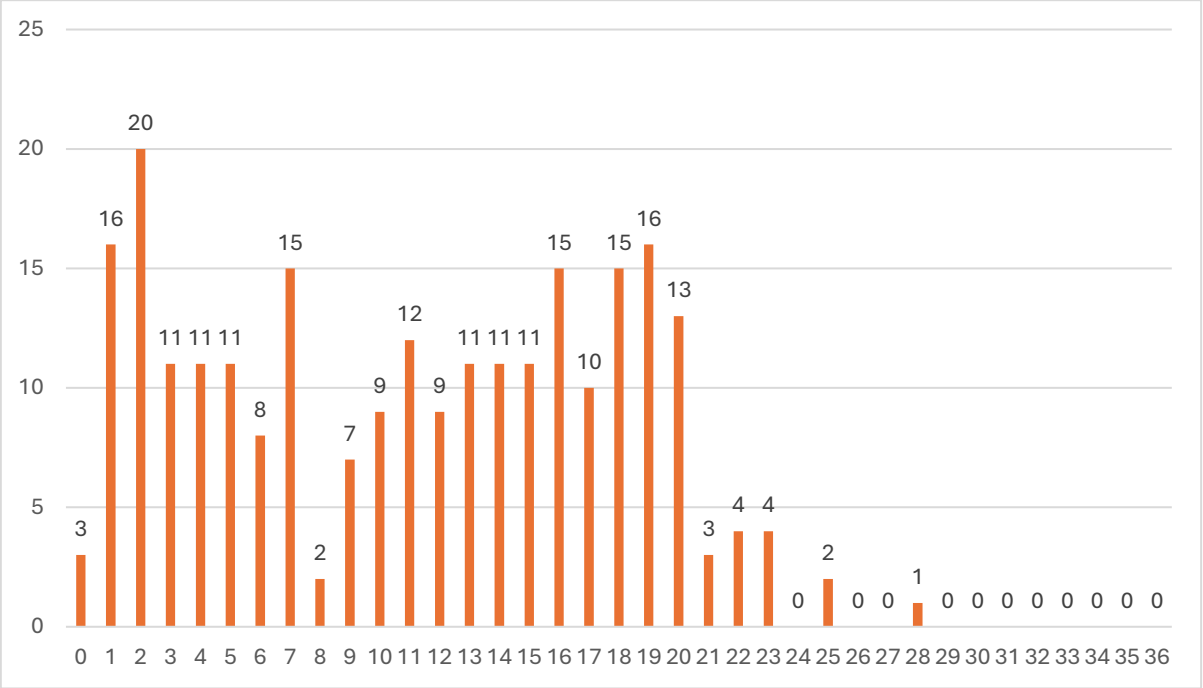
The second limitation relates to our binary scoring system — we assign either zero points or one point if we have evidence that a practice is carried out at least once. While it avoids subjectivity by not arbitrarily defining several levels reflecting what a more or less regular or satisfactory practice is, it nevertheless scores identically a fund that only occasionally implements a practice and an organisation for which it forms an integral part of its operating mode. The same is true for two funds implementing the same practice with different levels of rigour. Thus, our analysis does not distinguish practices according to the intensity or quality of their implementation.

2.3. Main Findings

2.3.1. General Findings

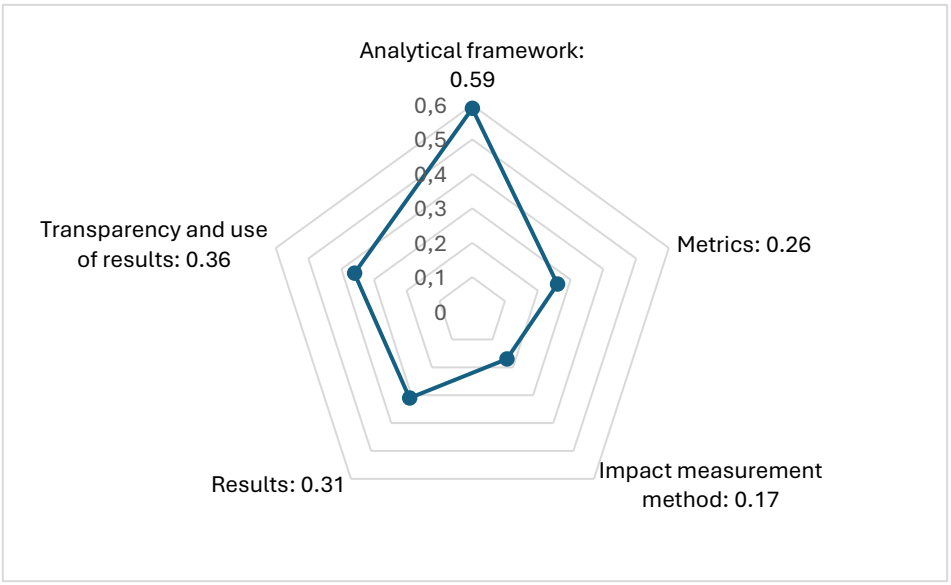
Before studying the correlations between fund characteristics and their practices, it is useful to have a general overview of the level of impact measurement practices. Figure 1 presents the distribution of total scores, without distinction between the five different dimensions. Scores are theoretically between 0 and 36 and increase as impact evaluation becomes more robust. We obtain an average score of 11.11 (the median is 11). Of 250 funds, three of them receive no points and nearly 30% of funds (72) have a score of 5 or below, reflecting a limited presence of impact measurement. Conversely, only 58 funds (23%) have a score above 18 (which is half the maximum score). Only three funds have a score above 25. These general results reflect practices of investment funds operating in Africa that are relatively distant from expected standards, as described in the previous section and reflected by the evaluation grid constructed. It should be noted that, in the absence of a comprehensive comparative analysis, it is difficult to know whether this result is specific to funds operated in Africa. Nevertheless, existing work, although partial, tends to indicate that practices in terms of evaluation are no better elsewhere.

Figure 1: Distribution of total scores



We then explore the five different dimensions to assess which practices are most widely adopted by funds. To do so, we normalise each dimension on a scale of 0 to 1 for each fund (to control for differences in the number of questions per dimension). The closer the score is to 1, the more the funds' practices approach the standards for the dimension considered. We then calculate the average of normalised scores for each category across all funds.

Figure 2: Average Normalised Score for Each Category



The results are presented in Figure 2. The first finding is the strong heterogeneity across the different dimensions. We observe a score of 0.59 for the "Analytical framework" category but a score of 0.17 for the "Impact measurement method" category. We explore each of the 5 dimensions in detail[14].

Analytical Framework (average score = 0.59): This dimension is the most commonly implemented by funds. The result of the first category is notably due to the communication of target clients, which is the most widely adopted practice by funds, concerning 243 organisations out of 250 (97.2%). The presentation of the ToC and the statement of the intention to measure impact are also applied by a majority of the set. Only the detail of the ToC concerns just 76 funds (30.4%).

Indicators (average score = 0.26): The most recurring outcome and impact indicators are those capturing the number of jobs created or maintained, the share of women among employees or clients of the structures, and environmental changes such as the quantity of greenhouse gas emissions avoided or the number of hectares reforested. However, metrics are not always clearly defined, particularly for employment. Only 33 funds (13.2%) use harmonised indicators, primarily from the IRIS+ metrics bank, with the HIPSO reference framework also sometimes mentioned.

Impact measurement method (average score = 0.17): The Impact measurement method category is the one least often implemented by funds. The most widely used approach by funds is the presentation and description of a certain number of projects in which they have invested. This majority seeks to prove its impacts in a narrative manner from a few specific cases of investments and companies. A large share of the set also resorts to the presentation of raw indicators and/or simple before-after investment or reality-target comparisons. This approach assumes that all observed changes are due to the investment. Sixteen funds (6.4%) go further by integrating these metrics into more elaborate methods to approach the real impact, using four main approaches (see Box 7). This indicates that some funds are aware that raw indicators are susceptible to capturing the effects of external factors and are acting accordingly.

[14] As a reminder, the normalised score is obtained for each fund by dividing the score per dimension by the maximum possible score. For example, for the "Transparency" dimension, a fund can earn a maximum of 8 points (8 binary questions). If its score is 2 points, its normalised score is 0.25 (2/8).
To obtain the average score, we calculate the mean of the normalised scores for each fund across each dimension. For example, if we have 5 funds with the following normalised scores for the "Transparency" dimension: 0.125 (1/8); 0.25 (2/8); 0.25 (2/8); 0.50 (4/8); and 0.875 (7/8), the average score will be 0.4 ([0.125+0.25+0.25+0.50+0.875] / 5).

Box 7: Impact Attribution Methods Used by Funds

In our set, sixteen funds seek to attribute impact using four different approaches.

First, some structures weight their indicators by an attribution rate. The latter is calculated as the ratio between the amount injected by the fund in the company and the total assets of the latter.

Second, some funds turn to the construction of an impact score. To do so, an analytical framework is constructed considering several dimensions. For example, the Nordic Impact Fund attributes a score out of twenty by evaluating four axes: depth, i.e. the magnitude of the observed variation and whether it constitutes a real transformation in the lives of beneficiaries; breadth, i.e. the number of people affected; inclusion, which takes into account the share of vulnerable people (women, youth, farmers...); and additionality, which measures whether the effect would have occurred without the fund's intervention. Depending on the fund, the dimensions analysed and the evaluation scale vary, with some attributing scores out of five.

Third, six funds using a rigorous quantitative approach use the Joint Impact Model (JIM)[15]. The latter uses economic matrices that describe average exchanges between sectors and countries. These are constructed from data collected by national and international statistical bodies such as the World Bank or the OECD. The model is completed with the structures' information to estimate the financial flows that an investment generates in the economy. The JIM then delivers direct, indirect and induced effects.

Fourth, three funds turn to the use of methods employing a counterfactual: Root Capital, Jasmine Social Investments and Gawa. The methods used are the randomised controlled trial, difference-in-differences and matching methods. Only Root Capital details how the counterfactual is defined for some of its investments. In Ghana, the fund collaborated with the Ghanaian research and evaluation organisation Participatory Development Associates to evaluate a financial literacy training programme for sorghum farmers. The evaluation is based on a randomised controlled trial with progressive introduction at the village level. In Mexico, in the State of Chiapas, Root Capital partnered with the Walmart Foundation to evaluate its effect on the living standards of coffee producers. The study used the propensity score matching method.

Reported Results (average score = 0.31): This category reveals that great importance is given to inputs and outputs as well as ToC outcomes. Indeed, the presentation of results relating to them concerns 186 funds and 160 funds respectively. Only 31 funds do the same for impact. The information communicated about outputs often refers to the number of beneficiaries or the amount of capital allocated to fund activities. Seventeen funds present nuanced or even negative results, and six organisations monetise their impacts.

[15] https://www.jointimpactmodel.org/documentation

Transparency (average score = 0.36): A majority of structures (145 funds, 58%) shares its knowledge with the impact investing community through participation in events, the publication of articles or the production of podcasts. Seventy-seven funds (30.8%) adopt a formative approach, while only eight structures discuss the limitations of their approaches.

Key Message: The main message of this analysis is that impact measurement practices by funds are far from expected standards. While most funds adopt a ToC, only a minority (6% of them) seek to establish a counterfactual analysis. Moreover, funds often use output and sometimes outcome indicators, but rarely impact indicators, and transparency is limited.

2.3.2. Explaining Differences Between Funds

The previous analysis provides a general picture of fund practices. Nevertheless, there is strong disparity between funds (Figure 1). The objective of this section is to analyse whether these differences can be linked to observable characteristics.

We first analyse the correlations between the total score and fund characteristics through a univariate analysis (simple correlation) and a multivariate analysis (econometric regression). The results are presented in the following table. To facilitate reading, we use a colour code depending on the direction of the relationship (green = positive; red = negative) and the magnitude of the relationship (the darker the shade, the stronger the relationship).

The results of Table 1 highlight that there are significant correlations in the univariate analysis, but these are rather weak. The association with fund size is moderate and positive, with a correlation coefficient of 0.37. This means that when larger structures are considered, high total scores are observed, and conversely, small sizes are associated with low total scores. Positive correlations, albeit less strong, also exist for age and the fact of having headquarters in Europe or Africa. In contrast, the results suggest that the higher the African footprint, the more likely it is to observe low total scores.

Nevertheless, accounting for relationships between characteristics (multivariate approach) makes most of the observed correlations disappear. The correlation with European origin is the only one to survive the passage to the multivariate approach. The score difference between a European fund and an African fund is +5 points, which is a rather strong impact insofar as most funds have a score below 20. On the other hand, the other dimensions are no longer significant. It is nevertheless useful to highlight that while African footprint is no longer significant at the usual 5% threshold, its economic effect remains relatively strong. Funds focused on the continent are less inclined to implement robust impact measurement practices.

Table 1: Results of Relationships Between Fund Characteristics and Total Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.25	<0.001	0.03	0.62
Size	0.37	<0.001	0.00	0.26
African footprint	-0.30	<0.001	-2.85	0.16
Growth	-0.04	0.60	-0.02	0.22
HQ in Africa	0.24	<0.001	/	/
HQ in Europe	0.28	<0.001	5.03	<0.001
HQ in Americas	0.05	0.44	1.49	0.37
HQ in Asia or Oceania	0.07	0.29	-3.39	0.35

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1 (no colour), weak if it falls between 0.1 and 0.3 (light green/red), moderate if it falls between 0.3 and 0.5 (medium green/red), and strong if it exceeds 0.5 (dark green/red). For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.36 (or non-significant), weak if it falls between 0.36 and 1.8, moderate if it falls between 1.8 and 3.6, and strong if it exceeds 3.6[16].
In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

We then explore in a second phase the relationships between characteristics and the different dimensions of impact measurement. We present the main results with the same colour code in Table 2 (the detail of each analysis is presented in Annex 2 at the end of this report).

The results reported in Table 2 are relatively consistent with the analysis of the link between characteristics and the total score. The univariate analysis highlights the positive correlation with the same four characteristics (age, size, African and European location) and the negative relationship with African footprint. The multivariate analysis confirms that only the presence of headquarters in Europe is robust. Thus, there does not appear to be a real difference depending on the dimension considered. This finding should however be nuanced in part. For example, the effect of having headquarters in Europe is particularly important for the implementation of the analytical framework, the measurement of indicators and the presentation of results. On the other hand, the economic effect is weaker for the adoption of attribution methods and especially for transparency.

[16] These thresholds are defined as follows: 0.36 represents the impact of a 1% increase in the total score (which ranges up to 36), 1.8 represents a 5% increase, and 3.6 represents a 10% increase. This ensures that the thresholds are consistent with the analysis of each dimension presented in the table below.

As previously noted, we underline the role of African footprint, which notably reduces the capacity to be transparent. The estimated coefficient is relatively strong and is significant at 10% (see Table A5 in the Annex).

Table 2: Relationships Between Fund Characteristics and the Five Dimensions of Impact Measurement

Characteristics	Analytical framework		Indicators		Attribution		Results		Transparency	
	Uni	Multi	Uni	Multi	Uni	Multi	Uni	Multi	Uni	Multi
Age	0.18*	0.000	0.19*	0.000	0.33*	0.000	0.26*	0.000	0.15*	0.000
Size	0.26*	-0.000	0.26*	0.000	0.28*	0.000	0.34*	0.000	0.31*	0.000
African footprint	-0.18*	-0.04	-0.28*	-0.0	-0.26*	-0.05	-0.23*	-0.06	-0.30*	-0.15
Growth	-0.08	-0.000	-0.01	0.000	-0.07	0.000	-0.02	0.000	-0.01	0.000
HQ in Africa	0.25*	-	0.23*	-	0.25*	-	0.18*	-25	0.23*	-
HQ in Europe	0.31*	0.19*	0.19*	0.11*	0.27*	0.08*	0.19*	0.14*	0.28*	0.02*
HQ in Americas	0.01	0.07	0.11	0.04	0.03	0.02	0.07	0.05	0.03	0.04
HQ in Asia or Oceania	0.03	-0.14	0.03	-0.07	0.10	-0.05	0.05	-0.05	0.09	-0.18

A relationship is considered statistically significant (indicated by an asterisk) if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1 (no colour), weak if between 0.1 and 0.3 (light green/red), moderate if between 0.3 and 0.5 (medium green/red), and strong if above 0.5 (dark green/red). For the multivariate analysis[17], the categories are as follows: null if the coefficient is below 0.01 (or non-significant), weak if between 0.01 and 0.05, moderate if between 0.05 and 0.1, and strong if above 0.1.
In the multivariate approach, no results are presented for African origin as it serves as the reference group for all other HQ locations. Accordingly, coefficients for other regions should be interpreted relative to Africa.

This analysis should be taken with caution. We specify that our results are only witnesses of associations and do not illustrate any causal link. Furthermore, a certain number of binary variables present very unbalanced categories: when we cross our variables in contingency tables, we obtain cells with very low numbers, even zero in some cases. Our results suggest that certain fund characteristics may have a real influence on some of our observations, but that other factors must also enter the equation.

Key Message: The main message of this exploration is that the observable characteristics of funds have little correlation with the adoption of impact measurement standards. Analysis of the explanatory power of the econometric models confirms that the inclusion of these variables is limited. At best, adding these different variables only enables 17% of the observed differences in scores between funds to be explained. In other words, more than four-fifths of the differences are unexplained by the observed characteristics. The adoption of impact measurement practices is therefore linked to factors independent of the structural characteristics of funds. In the last part of the report, we explore avenues to understand why funds do not adopt best practices and what measures could be taken to improve them.

[17] It should be noted that the thresholds used here differ from those in Table 1, owing to the nature of the dependent variable. In Table 1, the outcome is a score ranging from 0 to 36, whereas in Table 2 it is a relative score ranging from 0 to 1.

Part 3 – Improving Impact Evaluation Practices

The analysis of impact evaluation practices reveals a persistent disconnect between the social and environmental objectives of impact investment funds and their capacity to measure these impacts rigorously. To improve existing practices, it is useful to understand the obstacles explaining this situation.

3.1. Why Do Impact Investment Funds Struggle to Measure Their Impacts?

While our analysis did not consist of an investigation into the deeper reasons that may explain the differential between practices and expected standards in impact measurement, a review of existing literature enables the identification of several obstacles, both structural and behavioural, to explain this situation.

3.1.1. Lack of Knowledge of Principles and Methods

A first major obstacle is the lack of knowledge of impact evaluation principles and methods. Jalila (2022), through interviews with social entrepreneurs in Morocco, shows that impact evaluation is rarely adopted, partly because it is poorly known or poorly understood. Even when actors grasp the principle, they often ignore the rigorous methods to employ. Ormiston and Seymour (2011) highlight that the absence of clear formulation of missions and objectives prevents the definition of a coherent impact measurement strategy. Without this preliminary step, funds risk focusing on superficial indicators (such as growth or outputs) rather than actual impact.

Furthermore, the literature abounds with tools and analytical frameworks, but their usefulness and rigour are not always clarified, which generates confusion and limits their adoption (Gugerty and Karlan, 2018).

3.1.2. Difficulty of Implementation

Even when funds are aware of the importance of impact measurement, its implementation faces several challenges:

- Implementation Cost: Impact measurement requires continuous monitoring over the long term, specific skills (particularly for environmental aspects), and often the intervention of external experts (Koontz et al., 2020). This makes the process costly in terms of time, human resources and expertise (Jalila, 2022).
- Lack of Dedicated Resources: Impact investment funds are above all financial actors, whose skills and priorities remain centred on fundraising and profitability. Impact measurement mobilises different fields of knowledge, often unknown to managers (Ait Soudan and Kaffi, 2022).
- Difficulty Recruiting Qualified Profiles: Finding experts capable of conducting impact evaluations while understanding the specificities of the funds and their sectors of intervention is a major challenge.
- Data Access: Funds may have difficulty obtaining data, particularly secondary data, useful for constructing counterfactual-based analyses.

3.1.3. Lack of Incentives

A third obstacle is the absence of external pressure to measure impact rigorously. Stakeholders (investors, beneficiaries, regulators) often content themselves with monitoring basic indicators, without requiring an in-depth evaluation of social or environmental impacts (Ormiston and Seymour, 2011). Moreover, funds rarely perceive a clear and immediate return on investment from impact measurement. Indeed, the costs (recruitment, tools, time) are tangible, while the benefits (access to less expensive capital, competitive differentiation) are uncertain and deferred in time.

Incentives could notably come from capital providers who could require more robust analyses before agreeing to provide capital on favourable terms. Our analysis does not allow us to go further, due to data limitations, but it is possible that one of the variables that can explain differences between funds relates to their source of financing. Many funds are supported by development finance institutions (Léon, 2025) which are undoubtedly more sensitive to impact aspects and encourage their counterparts to invest in this domain.

Another important factor in incentivising funds to improve their practices is regulation. The European case illustrates this regulatory pressure with the adoption of the CSRD (Corporate Sustainability Reporting Directive), a European directive that entered into force in 2023 and considerably strengthens non-financial reporting obligations for companies. This regulatory pressure may be a key explanation for the fact that European funds have the most advanced practices in terms of impact evaluation.

3.1.4. Psychological and Cultural Biases

Finally, behavioural biases play a key role. The framing effect pushes actors to favour easily accessible and understandable information, such as financial performance, to the detriment of impact measurement, which is more complex (Ormiston and Seymour, 2011). The culture of funds often remains anchored in a financial logic, where growth and profitability indicators take precedence over social or environmental impacts.

These obstacles explain why impact measurement remains the “poor relation” of impact investing, despite its theoretical importance. Complementary analyses could (should) be conducted to refine the diagnosis on the relative importance of each of these factors and the possible existence of other obstacles.

3.2. What Solutions Can Embed Impact Evaluation at the Core of Funds?

The current situation is not inevitable. It is possible to make impact evaluation a central element of impact investment funds, while respecting their primary mission, which remains to invest. To achieve this, six action pathways can be envisaged, progressively and realistically, to respond to the previously identified obstacles.

3.2.1. Raising Awareness Across the Entire Ecosystem

The first step consists of raising awareness across the entire ecosystem about the importance of impact evaluation. For this to become a priority, it is essential that all stakeholders, from fund managers to funders, understand its usefulness. This implies recalling that claiming impact without proving it exposes funds to the risks of impact washing, which can harm their credibility in the long term. Funders, in particular, must be convinced that impact measurement is not a mere cost, but a lever for attracting capital at a lower cost, by demonstrating tangible social and environmental value creation. Studies, such as this report, can serve as support to illustrate that impact evaluation is not a constraint, but a tool for differentiation and transparency. This responsibility falls primarily on the umbrella actors of the ecosystem — professional associations, donors, sectoral platforms — who have the legitimacy and visibility to carry this message.

For impact measurement to become sustainable, it must be integrated into the very culture of the fund. This requires involving all staff, for example by organising internal workshops to explain its usefulness and its connection to the overall strategy. Integrating impact-related objectives into individual evaluations or designating an impact reference person in each team can also facilitate this buy-in. Impact measurement must become a routine, integrated into existing processes, such as portfolio reviews or investor reports. Thus, it will cease to be perceived as an external constraint and become a natural component of the fund's activity.

3.2.2. Developing a Methodological Framework to Guide Method Choice

The absence of a universal method in impact evaluation is often perceived as a major obstacle by funds. Faced with the diversity of existing approaches — theory of change, standardised indicators, quasi-experimental methods — teams can feel disoriented, lacking guidance on which method to prioritise in their specific context.

Rather than seeking to impose a single approach, which would necessarily be ill-adapted to the diversity of funds and territories of intervention, it would be more realistic and more useful to construct a practical framework. This would not say "here is the method to follow," but rather "here is how to choose your method," matching four key dimensions: the targeted impact, the data available in the intervention context, the budget available for the evaluation, and the most appropriate method given these constraints.

This type of tool would have a dual virtue. On the one hand, it would enable funds to navigate methodological complexity without requiring in-depth expertise internally. On the other hand, it would contribute to a certain harmonisation of practices, not by imposing a single method, but by structuring reflection around shared criteria. Its design could be the subject of a collective effort between ecosystem actors — funds, experts, donors — so that it reflects the reality of the African context and is appropriated by those who would most need it.

3.2.3. Applying Basic Methodological Principles

Even if no method is perfect, an impact evaluation must rest on credible foundations to be useful. This begins with a clearly defined theory of change, which identifies the intermediate steps between the

resources invested and the expected impacts. The chosen indicators must be aligned with these priority impacts and take into account adverse or indirect effects. The notion of counterfactual, even simplified, must be integrated into the analysis, for example by comparing the evolution of beneficiaries with that of a similar group of non-beneficiaries. Finally, transparency on the methodology is indispensable: explaining how data are collected, what the limitations of the measurement are, and why certain choices have been made enables the reinforcement of the evaluation's credibility. This responsibility rests above all on the funds themselves, but can be facilitated by practical guides or training provided by specialised actors.

3.2.4. Improving Access to Data

Access to data constitutes a prerequisite that is often underestimated. There are, however, many initiatives to make administrative data accessible, and even certain survey data. The World Bank, for example, makes rich and easily usable survey data available online. Similarly, administrative data are increasingly open, even if this dynamic remains limited in Africa. There are also new usable sources, such as satellite analyses, useful for measuring deforestation, habitat type, infrastructure or the level of economic activity (Donaldson and Storeygard, 2016; Burke et al., 2021).

Beyond access to data, it is sometimes information on their existence and availability that is lacking. Cataloguing available sources by impact theme would be a valuable complementary effort, which could be part of a broader approach to harmonising practices. This project could be led jointly by academic actors, open data platforms and donors active on the African continent.

3.2.5. Providing Financial and Technical Support to Funds

Many funds feel ill-equipped to face the complexity of impact measurement, for lack of sufficient resources or expertise. Financial support could take the form of dedicated funds or grants, covering part of the initial costs related to recruitment, tools or training. This responsibility falls primarily on donors and development finance institutions, which have both the means and the interest in raising the level of practice of the entire ecosystem.

In parallel, technical support is necessary to develop robust methods. This involves training teams, even those who will not directly implement measurement, so that they understand the issues. Partnerships with academic structures or experts, as observed in funds with the most advanced practices, could also offer valuable and tailored support.

3.2.6. Adopting a Progressive Approach to Build Competence

Impact measurement cannot be perfect from the start, and an iterative approach is more realistic. Funds could begin with a pilot project, selecting an investment where measurement is particularly feasible, for example in a sector where indicators are clear, such as employment or energy access. Rather than wanting to measure everything, it is preferable to focus on a precise objective, such as constructing a solid theory of change or developing a key indicator. The important thing is to engage in a process of continuous improvement, accepting that early evaluations may be imperfect. It is up to the funds themselves to commit to this approach, ideally drawing on peers who have already taken these first steps.

These six avenues are not independent of one another: their effectiveness rests on a collective dynamic, in which every actor in the ecosystem — funds, donors, experts, academic actors — has a role to play. It is on this condition that impact evaluation can become, in Africa as elsewhere, a genuine driver of transformation rather than a mere formal obligation.

References

Angrist, J.D. & Pischke, J-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of Economic Perspectives, 24(2), 3-30.

Burke, M., Driscoll, A., Lobell, D.B. & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. Science, 371(6535).

Campbell, I. (2007). Chi-squared and fisher-irwin tests of two-by-two tables with small sample recommendations. Statistics in Medicine, 26(19), 3661–3675.

Carter, P., Van de Sijpe, N. & Calel, R. (2021). The elusive quest for additionality. World Development, 141, 105393.

Cochran, W. G. (1952). The χ2 test of goodness of fit. The Annals of mathematical statistics, 23(3), 315–345.

Donaldson, D. & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. Journal of Economic Perspectives, 30(4), 171-198.

Fitoussi, J-P. (2013). Le théorème du lampadaire. Ed. Actes Sud, Coll. Babel, 256 pages.

Gugerty, M. K. & Karlan, D. (2018). Ten Reasons Not to Measure Impact – and What to Do Instead. Stanford Social Innovation Review, Summer 2018, 41-47.

Jalila, A. S. (2022). Evaluation de l'impact social de l'entrepreneuriat social le cas du Maroc.

Koontz, T. M., Jager, N. W. & Newig, J. (2020). Assessing Collaborative Conservation: A Case Survey of Output, Outcome, and Impact Measures Used in the Empirical Literature. Society & Natural Resources, 33(4), 442–461.

Léon, F. (2025). Development Finance Institutions in Developing Countries: A Literature Review. FERDI Working Paper, P350.

Léon, F. & Rabary, S. (2024). L'investissement d'impact en Afrique: Enseignements d'une cartographie. Rapport de la FERDI, 84 pages.

Léon, F. & Rabary, S. (2025). Actualisation de la cartographie de l'investissement d'impact en Afrique 2025: Quantifier l'empreinte africaine. FERDI Note brève, B288.

Ormiston, J. & Seymour, R. (2011). Understanding Value Creation in Social Entrepreneurship: The Importance of Aligning Mission, Strategy and Impact Measurement. Journal of Social Entrepreneurship, 2(2), 125–150.

Roor, A. & Maas, K. (2024). Do impact investors live up to their promise? A systematic literature review on (im)proving investments' impacts. Business Strategy and the Environment, 33, 3707-3732.

Annexes

Annex 1: Review of Evaluation Grids for the Impact Measurement Process

A number of documents address the evaluation process of an intervention, but they are most often used upstream and/or during implementation. International or national bodies, such as the Inter-American Development Bank (IADB) or the UK's Foreign, Commonwealth and Development Office (FCDO), thus offer guides aimed at guaranteeing the quality of evaluations. Conversely, our grid has as its objective the evaluation of what has already been carried out and, to our knowledge, there are few frameworks responding specifically to this objective.

The report "A Prospective Approach to Social Impact Measurement" by Avise, la Fonda et le Labo de l'ESS (AFL) presents an evaluation grid developed during a workshop bringing together experts, social and solidarity economy actors, researchers and funders, in order to analyse social impact measurement practices. This grid is structured around three broad categories: the design and implementation of the process, the impact of the evaluation process itself (and not that of the evaluated intervention) and the sought legitimacy. For their part, Rawhouser et al. (2017) propose, from a literature review in social entrepreneurship, a checklist for assessing the quality of an impact evaluation based on the coherence of the approach, the chosen measurement and the clarity of the time horizon and level of analysis (direct and indirect effects). Our grid adopts a different breakdown, structured around the stages of impact evaluation.

Despite these differences in structure and terminology, common themes emerge when examining these frameworks in detail. The theory of change, the adoption of a formative approach, and the communication of evaluation limitations are also present in the aforementioned documents or in the DAC-OECD's Quality Standards for Development Evaluation. The FCDO's EQUALS service, responsible for quality assurance of evaluations, particularly emphasises methodological rigour through the coherence between the chosen method and the evaluative question. Among all the grids and frameworks analysed, the AFL's is however the only one to explicitly mention the notion of counterfactual and to distinguish two levels of methodological rigour: a first relying on before-after comparisons or on the observation of post-intervention results, and a second based on the use of a counterfactual. Our grid introduces an intermediate level, integrating methods that resort to indicators to approach the real impact without constructing a counterfactual.

Beyond the themes addressed, a fundamental difference lies in the framing of the analysis. Some tools focus on specific aspects of the evaluation process. For example, the World Food Programme's checklists focus primarily on transparency and communication. Others adopt a broader framing: the MOPAN, which evaluates the performance of multilateral organisations, has developed a grid based on 12 key performance indicators, five of which relate to evaluation and results, but also covering dimensions such as stakeholder relations or the conceptualisation of interventions. Insofar as our report focuses specifically on how impact measurement is approached in practice, our grid covers all stages of the process, with particular emphasis on the measurement method itself. We thus favour a methodological rather than operational or organisational approach.

Another difference concerns the nature of the questions comprising the analytical frameworks. Our grid rests exclusively on closed questions, like those of EQUALS or the DAC-OECD Standards, while the AFL grid combines open and closed questions. This choice depends on the expected use of the responses: the AFL grid aimed to feed into a collective debate, while ours has the objective of producing scores usable in statistical tests and econometric models, which requires readily codable responses to conduct this type of analysis.

Finally, few tools are associated with an explicit scoring system. The MOPAN, for example, draws qualitative conclusions from averages of scores attributed to micro-indicators. Our system is similar, albeit simpler, with two levels of indicators (categories and questions) and binary scoring (0 or 1). This choice limits subjectivity, but leads to scoring identically practices implemented with different degrees of intensity or rigour.

In conclusion, while many tools address the evaluation process and share certain themes, few explicitly target the retrospective analysis of this process. When they do, the emphasis is often on operationalisation, without a clear scoring system or genuine consideration of methodological rigour or counterfactual. Furthermore, distinctions between outputs, outcomes and impacts are rarely made explicit, with the term "result" generally covering all of these components of the theory of change. This review rests solely on public documents; it is therefore possible that some organisations use internally tools closer to our grid.

Annex 2: Detailed Statistical Analysis for Each Dimension

Table A1 – Results of Empirical Approaches for the Analytical Framework Category Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.18	0.004057	1.59e-03	0.609
Size	0.26	0.0002824	-3.69e-12	0.721
African footprint	-0.18	0.01379	-0.04	0.676
Growth	-0.08	0.3173	-1.42e-04	0.837
HQ in Africa	0.25	0.0006864	/	/
HQ in Europe	0.31	1e-06	0.19	0.006
HQ in Americas	0.01	0.9048	0.07	0.392
HQ in Asia or Oceania	0.03	0.6173	-0.14	0.403

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1, weak if it falls between 0.1 and 0.3, moderate if it falls between 0.3 and 0.5, and strong if it exceeds 0.5. For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.01 (or non-significant), weak if it falls between 0.01 and 0.05, moderate if it falls between 0.05 and 0.1, and strong if it exceeds 0.1.
In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

Table A2 – Results of Empirical Approaches for the Indicators Category Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.19	0.002418	1.46e-03	0.53572
Size	0.26	0.0001823	1.13e-11	0.15104
African footprint	-0.28	7.118e-05	-0.07	0.31221
Growth	-0.01	0.9299	-5.21e-04	0.32222
HQ in Africa	0.23	0.0002515	/	/
HQ in Europe	0.19	0.003169	0.11	0.03314
HQ in Americas	0.11	0.07933	0.04	0.45305
HQ in Asia or Oceania	0.03	0.6191	-0.07	0.59527

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1, weak if it falls between 0.1 and 0.3, moderate if it falls between 0.3 and 0.5, and strong if it exceeds 0.5. For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.01 (or non-significant), weak if it falls between 0.01 and 0.05, moderate if it falls between 0.05 and 0.1, and strong if it exceeds 0.1.
In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

Table A3 – Results of Empirical Approaches for the Impact Measurement Category Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.33	1.841e-07	9.59e-04	0.44510
Size	0.28	7.626e-05	1.04e-12	0.80233
African footprint	-0.26	0.000229	-0.05	0.18838
Growth	-0.07	0.3868	-4.50e-04	0.10861
HQ in Africa	0.25	9.755e-05	/	/
HQ in Europe	0.27	2.629e-05	0.08	0.00613
HQ in Americas	0.03	0.5924	0.02	0.48736
HQ in Asia or Oceania	0.10	0.1024	-0.05	0.50343

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1, weak if it falls between 0.1 and 0.3, moderate if it falls between 0.3 and 0.5, and strong if it exceeds 0.5. For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.01 (or non-significant), weak if it falls between 0.01 and 0.05, moderate if it falls between 0.05 and 0.1, and strong if it exceeds 0.1.

In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

Table A4 – Results of Empirical Approaches for the Results Category Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.26	4.12e-05	1.20e-03	0.588118
Size	0.34	1.277e-06	1.15e-11	0.119919
African footprint	-0.23	0.001278	-0.06	0.400868
Growth	-0.02	0.7554	-7.39e-04	0.135371
HQ in Africa	0.18	0.0002164	/	/
HQ in Europe	0.19	0.002689	0.14	0.006286
HQ in Americas	0.07	0.2536	0.05	0.372732
HQ in Asia or Oceania	0.05	0.3972	-0.05	0.661216

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1, weak if it falls between 0.1 and 0.3, moderate if it falls between 0.3 and 0.5, and strong if it exceeds 0.5. For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.01 (or non-significant), weak if it falls between 0.01 and 0.05, moderate if it falls between 0.05 and 0.1, and strong if it exceeds 0.1.
In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

Table A5 – Results of Empirical Approaches for the Transparency and Use of Results Category Score

Characteristics	Univariate approach		Multivariate approach	
	Correlation coefficient	P-value	Correlation coefficient	P-value
Age	0.15	0.01557	-3.50e-04	0.905946
Size	0.31	2.469e-05	8.68e-12	0.378839
African footprint	-0.30	1.096e-05	-0.15	0.097456
Growth	-0.01	0.8681	-4.91e-04	0.456916
HQ in Africa	0.23	0.0002164	/	/
HQ in Europe	0.28	9.614e-06	0.02	0.002230
HQ in Americas	0.03	0.6488	0.04	0.622363
HQ in Asia or Oceania	0.09	0.1321	-0.18	0.254211

A relationship is considered statistically significant if the associated p-value is less than or equal to 0.05. Where significance exists, a correlation is considered non-existent if the absolute value of the correlation coefficient is below 0.1, weak if it falls between 0.1 and 0.3, moderate if it falls between 0.3 and 0.5, and strong if it exceeds 0.5. For the multivariate analysis, the categories are as follows: negligible if the coefficient is below 0.01 (or non-significant), weak if it falls between 0.01 and 0.05, moderate if it falls between 0.05 and 0.1, and strong if it exceeds 0.1.
In the multivariate approach, no results are presented for the African origin variable, as it serves as the reference group for the other headquarters locations. Accordingly, the coefficients for the other regions should be interpreted relative to Africa.

“Sur quoi la fondera-t-il l’économie du monde qu’il veut gouverner ? Sera-ce sur le caprice de chaque particulier ? Quelle confusion ! Sera-ce sur la justice ? Il l’ignore.”

Pascal

FERDI

Créée en 2003, la **Fondation pour les études et recherches sur le développement international** vise à favoriser la compréhension du développement économique international et des politiques qui l’influencent.

Contact
www.ferdi.fr
contact@ferdi.fr
+33 (0)4 43 97 64 60