Chapter 2

Designing and conducting EIB-GDN deep dives[1]

Authors: François Bourguignon (Emeritus Professor at the Paris School of Economics and École des hautes études en sciences sociales, and former Chief Economist and Senior Vice President at the World Bank), **Shahrokh Fardoust** (Research Professor, Global Research Institute, College of William and Mary, United States), **Pierre Jacquet** (President of the Global Development Network, member of the French *Cercle des Economistes*), **Emmanuel Jimenez** (Senior Fellow, International Initiative for Impact Evaluation 3ie), **Arianna Legovini** (Head of Development Impact Evaluation at the World Bank), **Alexandros Sarris** (Emeritus Professor of Economics, National and Kapodistrian University of Athens, Greece), **Claudio Cali** (Associate Economist, EIB), and **Nina Fenton** (Senior Economist, EIB)

JEL classification: C10, C19, C80, O1, O2

Keywords: deep dives conception and implementation, social and environmental impacts, capacity building, mentoring, technical rigour, research design and implementation, data collection, surveys, impact evaluation, theory of change, causality, value of research to private sector

Email: francois.bourguignon@psemail.eu (F. Bourguignon), sfardoust@WM.edu (S. Fardoust), pjacquet@gdn.int (P. Jacquet), ejimenez@3ieimpact.org (E. Jimenez), alegovini@worldbank.org (A. Legovini), alekosar@otenet.gr (A. Sarris), c.cali@eib.org (C. Cali), n.fenton@eib.org (N. Fenton)

[1] Abhay Gupta provided input for this chapter. Research assistance from Alexis Atlani and Brendan Harnoys-Vannier is gratefully acknowledged.

Summary and key messages

The previous chapter introduced deep dive impact studies as a way to complement other evaluation efforts and combine methodological rigour with the time requirements of stakeholders in an impact investment. This chapter discusses the experience and lessons from a recent pilot programme, developed and funded by the European Investment Bank (EIB) in collaboration with the Global Development Network (GDN), which conducted 16 deep dives of investments in ten African, Caribbean and Pacific (ACP) states. It details the design and principal components of the deep dives, how they were implemented, their main outputs/outcomes (and the value added for stakeholders), and the lessons learnt about design, implementation, utility and use.

The deep dives were conducted by 30 research fellows from ACP states. The researchers were well-trained but relatively junior, as an important subsidiary objective was to build evaluative and research capacity. Expert advisors and EIB and GDN staff supported the researchers in completing studies that passed academic review in a limited time period, adding to general development knowledge, with some expected to be published in peer-reviewed journals. The deep dive exercise provided lessons on how research can be relevant for operational and policy concerns (Chapter 3).

Lessons have been learnt about conducting these programmes. First, deep dive research must take into account a lack of existing data. While trying to make the best use of data, researchers needed to plan and budget for primary data collection. Second, engagement with the company implementing the investment is key. Company buy-in is essential for research and to ensure successful cooperation and data sharing. Third, researchers must be flexible about research approach and methodology. Time and resource constraints, and the nature of the project being assessed, meant that randomised control trials planned in advance were not generally feasible. Other techniques to address attribution were needed. Fourth, conclusions need to be tailored to the limitations of the analysis. It is a challenge to draw firm conclusions that are not generic, that add value to the global community and that are operationally relevant. Fifth, reporting in the right presentational form is crucial; expert advisors, GDN and EIB staff and consultants supported the researchers in their communications. Sixth, research capacity can be built, but it is a slow process; technical capacity building was limited by the short time period available. Seventh, the costs of the deep dive, although significant, were not prohibitive and judged by the reactions of all stakeholders were outweighed by the benefits.

Increasingly, investors, financial intermediaries and private suppliers/providers need to consider, in addition to their financial returns, the impact (positive or negative) of their projects on people, the environment and governance of their economies. Most multilateral development banks[2], bilateral donors and the foundations that support them have established practices to understand the development impact of their investment programmes, including private-sector investments. Almost all have frameworks to track some of the direct "results" of investments, often measuring these impacts during implementation as well as after project completion. Most carry out evaluations of all or selected investments (some institutions evaluate each operation, while others take mainly a thematic approach, with detailed case studies of a sample of operations).

[2] See, for example, the Evaluation Cooperation Group (ECG)'s Good Practice Standards for the Evaluation of Private Sector Investment Operations. The ECG consists of the independent evaluation units of the major international financial institutions.

However, as noted in Chapter 1, results measurement frameworks generally only capture the direct impacts of projects and focus on indicators that can be aggregated over different operations. Evaluations can study particular investments or themes in more detail, but are generally only implemented after project completion, and look at impact as one of several criteria. Stakeholders often want a deeper understanding of the impact of particular investments, including during the implementation period. Furthermore, results measurement frameworks do not generally address the issue of causal attribution, which would need to be tackled project by project, even though some multilateral development banks and bilateral donors, such as the World Bank, the Inter-American Development Bank and the Millennium Challenge Corporation, undertake numerous rigorous impact evaluations. Few institutions have made extensive use of impact evaluations that meet the highest academic standards. This is partly because of cost considerations and partly because traditional impact evaluation studies yield results after the investment project has ended. Assessing the evolution of impact during implementation, rather than only after the investment has ended, can help institutions to adapt practices and strategies to maximise impact.

The previous chapter introduced deep dive assessments as a way to complement other evaluation efforts and combine methodological rigour with the time requirements of stakeholders in an impact investment. What has been the experience with EIB-GDN deep dives and what lessons can be drawn from them? Deep dives are relatively new in the evaluation field and vary in scope and character (Chapter 1). This chapter discusses the experience and lessons from a recently completed pilot programme, funded by the EIB and implemented by the GDN. It details the design and principal components of these deep dives, how they were implemented, their main outputs/outcomes (and the value added for stakeholders) and the lessons learnt in their design, implementation, utility and use.

1. What went into the EIB-GDN deep dive programme?

The EIB-GDN project was started in 2017 and ran for three years (organised in three one-year cycles). It was conceived through a series of discussions starting in 2014 between the GDN and the EIB. The EIB was interested in complementing its impact measurement framework by developing the research component of its operations and involving developing country researchers. It identified the GDN as a partner to build research capacity in developing countries, and initiated a partnership focused on understanding the development impact of the Impact Financing Envelope (IFE) for Africa, the Caribbean and the Pacific regions.

The project aimed to strengthen the capacity of young researchers in developing countries while producing research-based insights into the development impact of investments funded by the EIB under the Impact Financing Envelope. Specifically, the project designers wanted to:

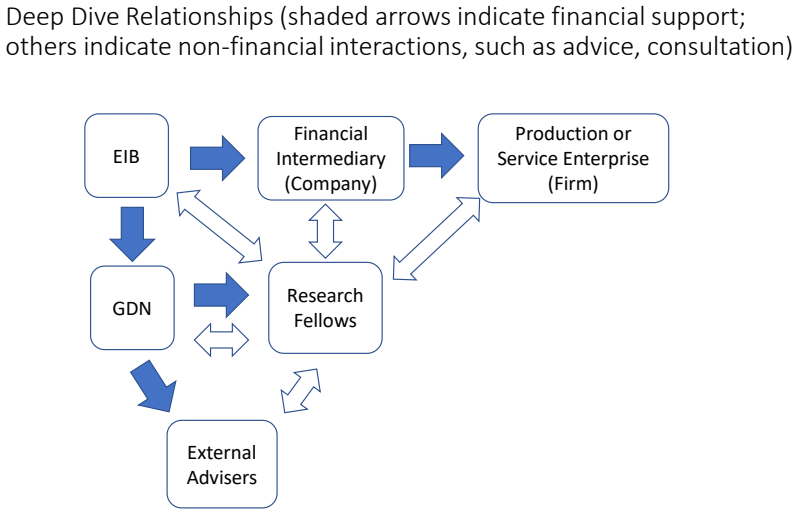
- Answer questions relevant to EIB clients and other stakeholders.
- Be rigorous in methods.
- Contribute to a new way of thinking about impact measurement that combines the needs of academic research and investor interests.

The programme aimed to match projects (in selected countries of Africa and the Caribbean) with researchers based in, or with a background in, those countries. The aim was to bring contextual knowledge and legitimacy by mobilising research fellows from beneficiary countries. This provided a unique opportunity for the researchers to acquire insights into projects financed by one of the biggest multilateral development banks in the world, while strengthening the pool of local experts.

To ensure that the studies met high standards of technical rigour, the GDN matched the researchers with a team of expert advisors. Chosen by the GDN for their experience, competence and high profile in the field, they ensured research met appropriate academic standards. This was particularly important given the innovative nature of the programme and the need to produce quality work in the fast-paced environment of the private sector in developing countries. Meeting quality objectives under tight time constraints and looking at complex investments requires skills, expertise and experience.

EIB staff were heavily involved in supporting the programme. The EIB had been expected to play a role in identifying the operations for analysis, making introductions to the relevant EIB clients, and ensuring strategic alignment between the studies and EIB objectives. In reality, the involvement of the EIB was more extensive (Figure 1 illustrates the key role of the EIB and the GDN among the stakeholders). Because the research fellows were relatively inexperienced, particularly in a private-sector context, the EIB team played an important role in explaining the nature of the investments and the underlying business models. They also acted as facilitators between the researchers and the firms, provided technical reviews of documents and edited reports, drawing on the technical expertise and subject knowledge of staff in the Bank's Economics Department and the Independent Evaluation Division.

Figure 1: Deep dive stakeholders



NB: This is an application of Chapter 1, Figure 2 to the case of deep dives.

The research outputs of the programme were called "deep dives." They were expected to answer questions relevant to EIB clients and other stakeholders; to be rigorous in methods; and to contribute to a new way of thinking about impact measurement that combines the needs of academic research and investor interests. As shown in Figure 1, the deep dives were written by researchers, and benefited from interactions from various stakeholders. They provided in-depth insights into specific research questions, shaped through interactions between the researchers, expert advisors, EIB and GDN staff and the companies that benefited from EIB funding. They were developed under the practical constraints of data availability and a timescale of one year per study. The following subsections detail how the EIB-GDN deep dive programme was implemented.

1.1. Recruitment of research fellows

The first task for each cycle was to recruit the research fellows through a competitive application process in three successive cycles. The process was geared geographically to facilitate matching between the selected researchers and a preliminary list of projects to be evaluated, proposed by the EIB, attempting to target francophone as well as anglophone researchers to make it easier to carry out studies across most countries of EIB operations in the relevant regions. The call focused on candidates from low- and middle-income countries in Africa, the Caribbean and the Pacific, since the source of funding was related to the Cotonou Agreement, which covers those regions[3]. Given the location of the projects selected for deep dives, most of the research fellows (except one from Haiti in the Caribbean) were from countries in sub-Saharan Africa, and lived and worked in the region (Table 1). A total of 30 researchers were recruited during the programme (as listed in the "Biographies" section of the book).

[3] Cotonou Agreement – Consilium (europa.eu).

The research fellows were aged under 40, with a median age just under 35. Though research fellows from a variety of disciplines were considered, there was a focus on quantitative skills that would be useful in programme and project evaluations. Most applicants were economists, with a strong grounding in the application of statistics and econometrics to the social sciences. There was a preference for those with or about to complete a PhD. Applicants had to take a test at the application stage, focusing on the understanding of, and ability to use, quantitative techniques. The test also checked data handling experience, strategic approach and communications. Quantitative analysis, while necessary, did not rule out the use of other techniques in deep dives – research fellows were encouraged to supplement their work with qualitative and institutional analysis.

Although female applicants were in a minority, the deep dive programme tried to ensure as much diversity as possible in the selection process. Among applicants with similar experience and training, women were preferred. Although gender parity was not achieved, the figure of 25% women among the research fellows is not insignificant, given the rate of female representation in the economics profession generally[4] and the fact that recruitment was targeted at those from low- and middle-income countries. Both factors limited the female applicant pool.

Table 1: Characteristics of research fellows

	Number recruited/number applied	Number of recruits with a PhD in economics (% of total)	Number with foreign degrees	Median years of age/work experience	% women	% sub-Saharan African nationals	% living/working in sub-Saharan Africa
Cycle 1	10/128	8 (80%)	8	Under 35/ 3.5	30	100	70
Cycle 2	10/158	8 (80%)	6	Under 35/ 3.5	20	100	60
Cycle 3	10/189	10 (100%)	7	Under 35/ 4	30	90	70

The final recruits were chosen by a committee consisting of expert advisors, the EIB and the GDN. The candidates all had strong academic credentials. There appears to be a strong pool of researchers in sub-Saharan Africa with an academic background in quantitative microeconomics, who could be suitable to carry out deep dive studies. It was sometimes difficult to choose among the finalists. However, the researchers were relatively inexperienced in carrying out independent research, few had experience working in or with the private sector, or of analysing private-sector investments from an academic viewpoint, and few had experience of independently managing research projects.

Given the relative inexperience of the research fellows, they were allocated into teams to relieve the pressure and enable them to provide support to each other, especially because of the intensity of engagement and data work. Most of the research fellows formed teams of two (one worked alone) and there were 16 deep dive projects. This fitted with the number of projects funded under the Impact Financing Envelope that were at a suitable operational stage for analysis. The pairing provided a good

[4] See https://blocnotesdeleco.banque-france.fr/en/blog-entry/economics-where-are-women.

geographical match, as pairs of research fellows were more likely to include at least one researcher originating in countries where the projects analysed were located.

1.2. Research support

Through EIB funding, the researchers were given a one-year fellowship that included a two-week induction course, financial resources to conduct the analysis and access to research support from the expert advisors. They received a stipend of €15 000 and a research allowance (for surveys and other research-related expenses) of up to €10 000, amounting to €20 000 per project where the researchers worked in pairs.

The principal research capacity support given to the researchers was through five expert advisors, who were chosen because of their background in development research, impact assessment and evaluation[5]. Their principal role was to mentor the research fellows during the design, conduct and write-up of the deep dives. Each expert advisor was given a notional research fellow allocation, although some assignments were shifted after project selection if a particular expert advisor had a comparative advantage in the project to be implemented.

The support from the expert advisors was supplemented by the logistical services provided by the GDN and a dedicated team from the EIB Economics Department, with input from the Evaluation Department and other EIB services. The EIB team managed the partnership for the Bank, especially in making links to EIB clients and acting as facilitators between the researchers and the private-sector counterparts. They reviewed documents from a technical perspective and to ensure that EIB operational practices and the business models of the EIB clients were appropriately reflected, drawing on their professional expertise and roles in the Economics Department and Independent Evaluation Division. They provided significant support in editing the studies to a standard acceptable for publication as an EIB report, and worked with the researchers to provide readable summaries that were circulated as blogs.

2. How were the deep dives implemented?

The steps in conducting the deep dives were as follows: project selection; identification of the evaluation questions by all stakeholders, taking into account the context of the projects, and development of an analytical framework to analyse these questions; write-up and review of the research plan including a literature review; data collection and analysis; write-up and review of the finished draft deep dive. This section describes each and summarises some of the issues confronted.

To facilitate the integration of stakeholder experiences and opinions concerning the various aspects of the interventions, a short survey was given at the time of writing this book to all the stakeholders of the programme, namely the EIB staff involved, expert advisors, research fellows, managers of the investment companies who acted as financial intermediaries, and managers of the firms who were the

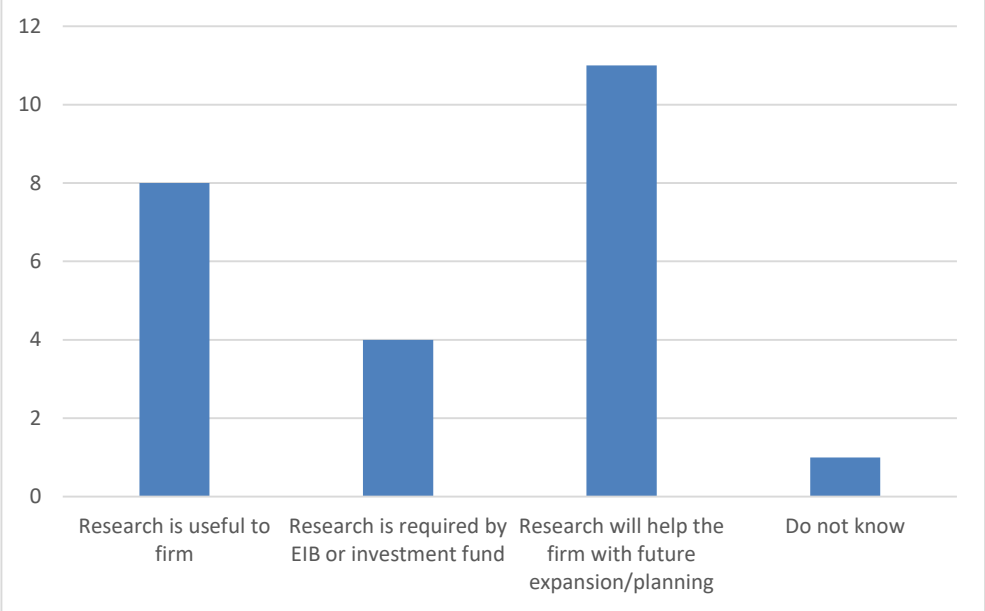
[5] The five advisors, François Bourguignon of the Paris School of Economics (PSE), Shahrokh Fardoust of the College of William and Mary, Emmanuel Jimenez of the International Initiative for Impact Evaluation (3ie), Arianna Legovini of the World Bank, and Alexandros Sarris of the University of Athens, saw the project as an opportunity to build evaluative research capacity and to promote richer interaction between academia and the operations of a development finance institution such as the EIB.

ultimate recipient of the investment funds. These short surveys aimed at eliciting the opinions of all stakeholders involved with respect to the key aspects of the project. These included questions on the following: how the ultimate beneficiary firms were selected; who among the beneficiaries experienced change as a consequence of the financing and how; types of economic and social value created; the thoroughness of the theory of change produced; impressions about whether the various changes would have happened in the absence of the intervention; the extent of useful knowledge learnt by the firm managers; the research design of the deep dive; the implementation of the research including survey design and conduct; the clarity and robustness of the research results obtained; and the type of research capacity built. The answers to these short surveys were incorporated into this and other chapters of this book.

2.1. Project selection

The EIB proposed a preliminary list of projects before recruiting the fellows so that some geographical and capacity matching could occur. However, the list was intended to be amended and modified depending on the implementation stage of a project, the appetite of funds and companies receiving EIB support and implementing the projects to participate in the deep dive exercise, and the research potential of each project – that is, the possibility to address relevant and useful research questions. The research fellows reported that firms were willing to participate because it would help with future expansion and planning (Figure 2). This is roughly aligned with how firms viewed the deep dives, as discussed in the next chapter.

Figure 2: Number of deep dives on the elements that convinced firms to participate, as reported by the research fellows



Source: Survey of research fellows

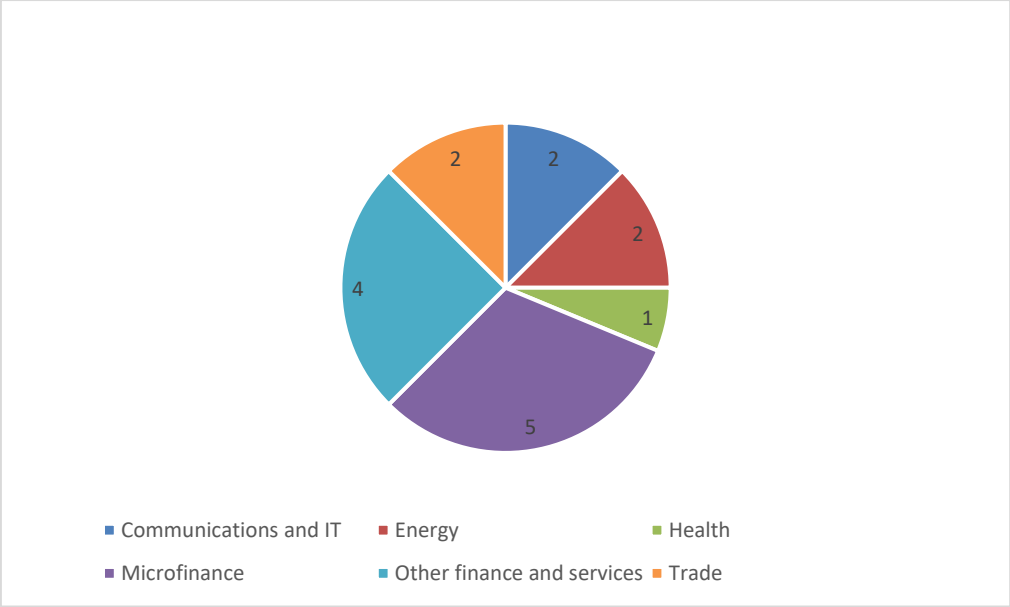
Initial projects were assigned by the GDN to the research fellows before the two-week introductory induction. They were required to study them on the basis of available information and to submit a provisional research work plan that would be discussed during the induction. These work plans were subjected to comments (by the expert advisors, other researchers, the EIB and the GDN). The

information provided by that discussion enabled the list of projects to be fine-tuned and allocated to research fellows. In the first cycle, the matching problem (based on capacity and geography) was dealt with by grouping research fellows in pairs. Given the range of projects under consideration, mentoring facilities were enhanced by grouping the expert advisors in pairs for each project, on the basis of their own interests and expertise.

During the first cycle, it took up to three months (Table 3) to select projects and finalise the allocations (2.5 months on average). Thus, most research fellows had less than ten months to identify the research questions and complete the analysis.

Most EIB investments analysed under the programme were undertaken via financial intermediaries (microfinance institutions, microfinance funds, and private equity and venture capital funds), which in turn funded firms (typically small and medium enterprises, or SMEs) in various sectors. The deep dive process initially chose the financial intermediaries that were willing to participate. In most cases, these financial intermediaries were not directly the subject of the deep dives. Instead, those intermediaries, together with the EIB and the researchers, selected firms that they had financed that could be interesting targets of the studies. These firms were then asked if they would be willing to participate in the deep dive research. They operated in a wide range of sectors, including microfinance, information and communications technology (ICT), education and health (Figure 3).

Figure 3: Sectoral breakdown of deep dives (number of deep dives)



The programme aimed to examine a variety of investments rather than providing a representative picture of the portfolio under the Impact Financing Envelope[6]. In most cases, the focus on the ultimate recipient of the funds also meant that the operation of the financial intermediaries was not a focus of assessment. This was a deliberate choice since various stakeholders, including the EIB, wanted to know the effect on environmental and social outcomes of actual investments.

[6] The EIB has an independent evaluation group that examines individual projects, sector-wide policies, programmes, partnerships and financing instruments. External stakeholders, such as the European Commission, may also undertake evaluations of the mandates or investments that they support.

The choice of sectors for the deep dives (Table 2) was guided by practical criteria (willingness to participate) and feasibility concerns (for studies that could be completed under the resources allocated to the programme). This differed from a sector-based strategy, which could include a series of studies in a given sector and deepen learning about what works in different contexts.

The opportunistic approach also partly explains the wide geographic distribution of the deep dives. Twelve of the deep dives were in eight African countries: Benin (one), Cameroon (one), Côte d'Ivoire (one), Ethiopia (two), Ghana (one), Kenya (three), Nigeria (two) and Senegal (one). Two were in multiple African countries. Two were in the Caribbean (Haiti and Jamaica). None were in the Pacific. This precluded any strategic focus on certain sub-regions, but had the advantage of covering a diverse group of ACP countries.

Table 2: Characteristics of projects selected and research questions

Country	Financial intermediary	Company to be evaluated	Sector of economic activity	Key research questions	Principal method of assessment	Mixed methods used?	Original data generated?
Kenya	TLCom	Ajua CRM	ICT	- What are the determinants of CRM software adoption? - What is the impact of CRM software on customer volume and sales performance?	- Theory of change (TOC) - Econometrics (with the use of two instrumental variables)	No (except the use of theory of change)	Yes (surveys)
Multiple countries	Direct	Currency Exchange Fund (TCX)	Finance and other related services	- In a given African country, how has infrastructure been financed (local or hard currency) and why? - How has foreign exchange volatility been? Is it costly? - What is the willingness to pay for hedging?	- Qualitative and quantitative analysis - Construction of a financial model	Yes	Yes (qualitative interviews)
Ethiopia	Direct	M-BIRR	Finance and other related services	- What is the average stated willingness to pay for mobile money use in utility bill payments? - How sensitive to price is the demand for mobile money use in utility bill payments? - How effective are non-price factors in affecting the demand for mobile money use in utility bill payments?	- Estimation of the willingness to pay (using stated preferences on random given attributes, the contingent valuation methodology and a probit model estimation) - Qualitative research (interviews)	Yes	Yes (surveys and semi-structured interviews)
Kenya	Novastar East Africa Ventures	SOKO Project	Trade	- For artisans, what are the main impacts of company affiliation on income, skills, access to healthcare and education, and poverty status? - Is there a differentiated impact of SOKO by gender?	- Theory of change - Econometrics (t-test, regressions, propensity score matching - Qualitative research	Yes	Yes (surveys, interviews)
Cameroon		Nachtigal	Energy	- The potential impacts of Nachtigal Hydro Power Company investments on vocational education and training. - Does the renovation or building of school infrastructure and the provision of quality training equipment lead to increased student enrolment and attendance in vocational schools?	- Theory of change - Econometric estimation	Yes	Yes (surveys, interviews)

Country	Financial intermediary	Company to be evaluated	Sector of economic activity	Key research questions	Principal method of assessment	Mixed methods used?	Original data generated?
Nigeria	Energy Access Ventures	RenSource	Energy	- What is the profile of Solar_Co's customer base? - What drives some merchants to subscribe to Solar_Co solar energy? - How do Solar_Co customers benefit from the service?	- Theory of change - Descriptive statistics - Econometrics (including using propensity score matching and endogenous switching regression (ESR) approaches)	Yes	Yes (surveys, interviews)
Jamaica		Development Bank of Jamaica (DBJ)	Finance and other related services	- What is the impact of DBJ loans on microfinance institutions? - What is the impact of DBJ loans on the performance of SMEs?	-Econometrics (instrumental variable and two-stage least squares (2SLS)) - Qualitative analysis and review of public quantitative documentation	Yes	Yes (surveys, interviews)
Ghana		Baobab Microfinance Company (BMC)	Finance and other related services	- Has BMC been reaching the poorest women in its areas of operation? - Has microcredit impacted women's economic empowerment in the five districts of BMC operations? - Has there been a reduction in the poverty level of BMC's female clients?	- Theory of change - Econometrics (average treatment effect, including using propensity score matching and inverse probability weighting (IPW)) - Qualitative research	Yes	Yes (surveys, interviews)
Haiti		Microfinance (ACME)	Finance and other related services	- Did the receipt of an ACME loan stimulate the expansion of Haitian microenterprises? - Did it help Haitian microenterprises acquire more fixed business assets? - Did it have an impact on the profits of microenterprises?	- Econometrics (using matching)	Yes	Yes (surveys, interviews)
Côte d'Ivoire	LMDF	Première Agence de Microfinance - PAMF-CI	Finance and other related services	- Understand the profile of beneficiaries and former beneficiaries. - Assess whether they have benefited, and if so how, from borrowing (in terms of business success, incomes, wealth and subjective well-being). - Measure customer satisfaction and draw strategic and operational lessons to inform the design and implementation of operations.	- Econometrics	No	Yes (surveys)

Country	Financial intermediary	Company to be evaluated	Sector of economic activity	Key research questions	Principal method of assessment	Mixed methods used?	Original data generated?
Nigeria	Partech	TradeDepot	ICT	- To what extent has the company's digital retail distribution platform lowered the supply price of a select basket of goods for small retailers? - To what extent do retailers who use the platform pass on the lower supply price of goods to consumers? - Which groups of small retailers benefit from the company's digital retail distribution platform (specifically women and young individuals)?	- Theory of change - Literature review - Quantitative methods	No	Yes (surveys)
Benin	PAMIGA	RENACA/ PAMIGA	Finance and other related services	- Which financial and output market and non-market mechanisms are associated with smallholder farmers' decisions to invest in agriculture in general, and in irrigation technology in particular? - How do these mechanisms explain differences in irrigation technology adoption among farmers?	- Theory of change - Descriptive statistics - Econometrics - Qualitative analysis	Yes	Yes (surveys, focus groups)
Multiple		Africa Energy Guarantee Facility (AEGF)	Finance and other related services	- Review the case studies of the use of risk mitigation instruments for energy projects in Africa. - Demonstrate how such instruments improve the viability of energy projects by reducing the financing costs.	- Theory of change - Literature review - Reviews of former case studies (impact, costs and so on) - Realisation of a case study: cost-benefit analysis	Yes	Yes (interviews)
Senegal	I&P	NEST	Health, education and related	- Who are the NEST patients? - How willing are patients to pay for and to use midwife-led monitoring and delivery care services? Why? - Does the provision of information via short stories (that is, visualisation) influence patient demand for midwife-led care services better than basic information?	- Theory of change - Descriptive statistics - Auction methods - Field experiment - Qualitative analysis	Yes	Yes (surveys, focus groups and field experiments)

Country	Financial intermediary	Company to be evaluated	Sector of economic activity	Key research questions	Principal method of assessment	Mixed methods used?	Original data generated?
Ethiopia		Cepheus Growth Capital/ Novastar	Finance and other related services	- What are the main barriers to financial access for firms in Ethiopia? How do they influence the different types of growth-oriented investments? - Does a lack of access to finance constrain firm growth and investment decisions?	- Descriptive statistics - Econometrics (including using instrumental variable and propensity score matching	No	Yes (online and field surveys)
Kenya	Novastar East Africa Ventures	poa! internet	ICT	- How did providing free internet access to schools in Kenya affect student outcomes (including internet skills, ICT skills, school attendance and access to e-education)? - Determination of customer satisfaction and related challenges.	- Non-experimental econometrics (including using inverse probability weighting and propensity score matching - Qualitative research	Yes	Yes (surveys and interviews)

Note: ICT refers to Information and Communications Technology
Source: EIB-GDN (2019, 2020, 2021)

2.2. Defining the research/evaluation questions

The programme demonstrated that deep dive studies could be valuable to companies and the EIB while appealing to the academic interests of researchers and expert advisors with technical rigour. This meant that the initial identification of a question by the research fellow was generally followed by iterative discussions with the EIB team, the project promoters and the expert advisors.

Based on these discussions (and in line with the objectives of the deep dive programme presented in Box 5 in Chapter 1), the research fellows adjusted the research focus to meet the expectations of the EIB and its clients and to fit with the particularities of the business models. In many cases these were not initially fully understood by the researchers, who were not experienced with private-sector operations. The iterations between the companies and the research fellows helped to fit the research ambitions with what would be practically feasible in terms of timeline and data availability. Several iterations were often needed, with many of the discussions facilitated by the EIB team. Few of the research fellows had experience communicating with private-sector companies, and often presented their ideas in a highly technical and academic manner. It was sometimes hard for the private-sector companies to see how the research could be of interest to them. The EIB team and, in some cases the expert advisors, acted in many cases as "translators" between the young researchers and the private-sector counterparts, particularly at the initial stage.

Table 2 shows the research questions that were the focus of the assessments. Many of the questions were of clear interest to the company as well as to the EIB, such as more information about the customer base. This spanned a wide range of services, including take-up of microfinance loans, midwifery services or use of telecommunications for mobile payments. Collecting this detailed profile information, which went beyond the information the project promoters could reasonably collect on a regular basis, gave the EIB deeper insights into its social impact, particularly whether the customer base served relatively poor and disadvantaged groups. The profile information was also often useful for commercial purposes, and in some cases more detailed commercially relevant information was collected, including the willingness to pay for services that would affect the companies' bottom line materially, or the degree of satisfaction with the service the company has provided (see Chapter 3 for further information about the use of deep dives).

In many cases it was relatively easy to identify overlapping interests. In most cases, however, the research ambitions of the research fellows had to be narrowed to fit with the availability of data, time and financial resources. As noted in the earlier chapters, the other stakeholders sometimes had objectives that differed markedly from research fellows' original proposals[7]. While some shared the research priorities initially identified by the research fellows, most had more immediate operational questions in mind, which required different research methods. For example, in one of the studies, the research fellows initially proposed using satellite data to study the effect on economic activity in the area over time. Since this aspect was already being assessed by another investor, the company asked the researchers to focus on a relatively small sub-component – the upgrade of training programmes in the affected communities – to see if their social investments were helping to improve the lives of people in the affected communities. This required a reassessment of the research proposal.

[7] For more than ten deep dives, firms reported that they had accepted to take part in the deep dive exercise because they were convinced that it would help later expansion. In four of the deep dives, research fellows had to use the argument that the EIB required an evaluation.

This process of interaction to identify suitable research questions proved essential to ensure both operational and academic relevance for the studies. The average time from the choice of a project to the start of relevant research activities was two to four weeks.

2.3. Development and review of the research plan

Once the research questions were defined, the research fellows wrote up their approach, defining what data and methods were going to be used to address them. This document was then reviewed by the expert advisors and the EIB, who typically raised questions about whether the methodology selected would enable the research questions to be answered, the likely availability of the data to implement the methodology, and the realism of the implementation plan. The research plans were typically revised at least once after the initial round of comments. The research fellows were responsible for keeping the private-sector companies updated on the research plans.

A key part of the review centred on recreating the theory of change of the intervention, which was key to understanding the pathways through which the companies' inputs would affect the outcomes. A theory of change presents which assumptions are key to ensuring that project inputs will lead to stated outcomes. It discusses how inputs, such as materials or staff bought with loan proceeds, are combined through activities (production processes) that lead to outputs such as products. Assumptions include the availability of factors of production and efficiency of implementation. A theory of change outlines how such products are marketed and sold, thus generating outcomes such as increased revenues, and impacts on the incomes and well-being of those who benefited from the loans. Key assumptions are the demand for the products, the marketing skills of the sellers and how they manage the revenues generated. Such a theory of change can be used to identify the key assumptions that connect inputs to outcomes and can thus be the subject of the most relevant research questions. A theory of change is also important for understanding the reasons behind quantitative research findings regarding impact.

The managers of the companies that benefited from the loans envisaged a logic chain when interviewed, and the research fellows reconstructed the theory of change in conjunction with EIB staff and the expert advisors. About one-third of the survey responses of the expert advisors expressed the view that the project did not have a fully articulated theory of change. A similar percentage felt that economic theory could have been used more by research fellows in crafting the theory of change. This was an area where the expert advisors showed their value added to the process and made suggestions to clarify key assumptions critical to determining outcomes.

Box 4: How do interactions between research fellows and expert advisors affect the preparation of research work? Assessing the impact of SOKO

The EIB supports investment companies and funds in African, Caribbean and Pacific countries, which in turn have invested in local small and medium enterprises. SOKO is a Nairobi-registered firm that has been financed by Novastar, an investment fund company based in Kenya and Nigeria. It designs ethically sourced fast-fashion jewellery and accessories, which are produced by a growing network of low-income artisans in Kenya, and sells the products online to consumers in the West through hundreds of fashion boutiques, large established retail chains and direct online channels. SOKO's central innovation is its variable cost "virtual factory" of artisans managed by a proprietary

mobile technology platform. It uses smartphones to connect independent artisans and global customers directly.

The project undertaken by two research fellows of the GDN and supervised by two expert advisors investigated the social impact of SOKO membership in Kibera, a suburb of Nairobi. The initial research questions formulated by the research fellows asked whether SOKO membership enhanced income and sustainable skills, as well as investments, children's education, family health and labour market participation. The methodology proposed was to compare SOKO members (lead artisans who are workshop owners and workers) with a counterfactual group of potential SOKO members through a propensity score matching (PSM) technique.

After a first round of review and suggestions by EIB staff and the expert advisors, the proposal was amended. First, with the help of expert advisors, a sound theory of change between the SOKO inputs and the impacts was formulated based on human capital theory. In SOKO's case, the human capital investment consisted of on-the-job training for lead artisans and workers. Second, on the suggestion of expert advisors, the research questions were modified to be more specific and to include the following: the impact of SOKO affiliation on lead artisan and worker income, access to health and education, skills, poverty status, issues of income diversification and whether women were motivated to participate.

A survey was designed with the support of the expert advisors in framing the questions and administered to the treatment and control groups. Problems arose in achieving the required number of interviews for the treatment and control artisans, because the sample frame was incomplete. Various techniques were employed to amend this, particularly "snowballing," whereby an artisan was asked for contacts with other artisans.

The initial method employed for the analysis was comparison of means for diverse variables among SOKO- and non-SOKO-affiliated artisans. This method cannot provide attribution for the differences, so the expert advisors recommended using econometric techniques. This was done, and a propensity score matching methodology was used to select the most similar non-SOKO artisans to those who were SOKO members. The result was a sounder analysis of impact on economic and non-economic aspects of artisans' well-being, and a more detailed analysis of the social impact of SOKO.

As a measure of welfare and living conditions, a Poverty Probability Index was used, despite recognised weaknesses. Several other variables thought to be correlated with welfare were tested. The end result was that workers with SOKO affiliation seem to enjoy higher welfare levels, and SOKO seems to have had a positive social impact on lead artisans and workers. With some caveats, these conclusions appear to be valid, although more robustness checks are needed.

There were other areas where expert advisors contributed, such as the feasibility of methods, the adequacy of data and the write-up of the results (Boxes 1 and 2). Some initial drafts overreached in writing up conclusions when the methods to establish causality were imperfect. In several projects, expert advisors provided guidance on the proper wording of what constituted programme "impacts."

An important issue was what method to use. The programme aimed to employ rigorous methods using counterfactual analysis. In most cases, however, the so-called "gold standard" of randomised controlled trials could not be employed. Only one deep dive implemented a field experiment (Table 3). The short time frame, and the fact that many of the supported businesses had been operating for

several years and some time had passed since the receipt of support from the EIB or the funds, made it difficult to implement a randomised controlled trial. Most deep dives used quasi-experimental causal analysis and four relied on correlations to make their conclusions. A key lesson is that deep dives need to identify a protocol for what would constitute a "satisfactory" causal analysis under practical constraints.

Table 3: Distribution of methods and cost of research across cycles

Research cycle	Average duration of project selection (months)	Average duration of defining research questions (months)	Number using randomised controlled trials	Number using other quasi-experimental techniques	Number unable to do causal analysis	Average cost of research including researcher's stipends (in thousands of euro)
Cycle 1	3	3	0	2	3	25.0
Cycle 2	3	2	1	4	0	26.5
Cycle 3	2	1	0	4	1	28.5
Total	2.5	2	1	11	4	27.0

Another issue was the availability of data to conduct the analysis. Expert advisors often warned about the difficulty that researchers would have, and they could not assume that data would be readily available. Primary data collection had to be worked into the budget and the available timeline. Researchers had to explore all possibilities to exploit secondary data. This is detailed in the next section. Box 2 provides examples of how criteria were used to adjust research designs.

Box 2: Alternatives to complete counterfactual analysis when experimental or quasi-experimental methods are unfeasible

In the deep dive of the TCX investment, the original research design, besides providing information about major energy projects involving foreign providers, was an elaborate economic modelling of the social willingness to pay for hedging against exchange rate volatility in a power plant operated by a foreign provider in Tanzania. Application to real data was disappointing, because Tanzania's exchange rate had been relatively stationary, with little short-run volatility over the recent past, whereas domestic prices showed limited fluctuations around a stable inflation trend. Hedging proved of little interest based on past exchange rate behaviour. The expert advisors guided the focus away from present conditions so that the deep dive evolved towards a full simulation analysis of the social willingness to pay for hedging against varying levels of exchange rate volatility. This analysis provided the hedging company with an estimation of the full demand function for its services and the public energy supplier with an instrument that would help evaluate

the contracts offered by the hedging company, depending on key policy parameters like the extent of the pass-through of exchange risk to end users, risk aversion or time discount rate.

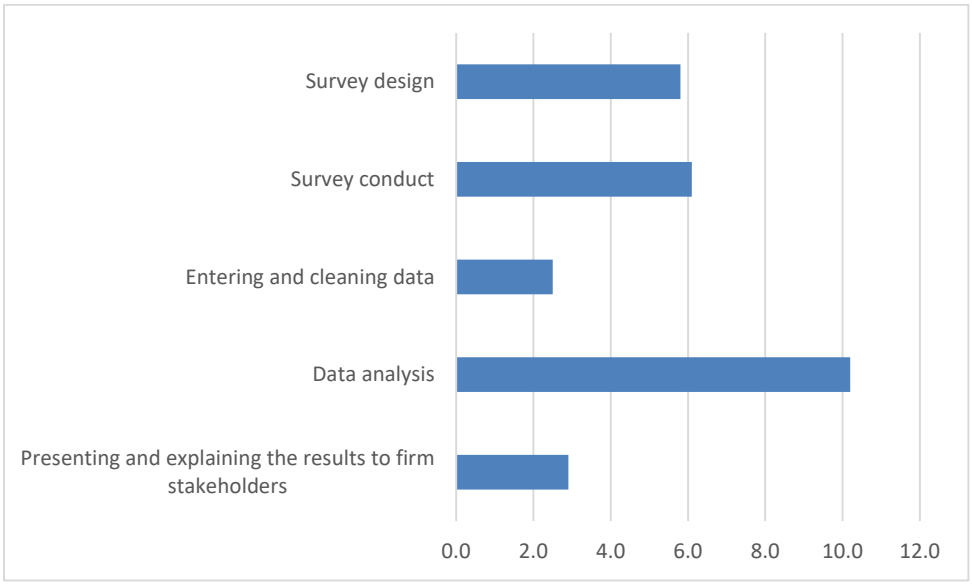
Source: EIB-GDN (2021)

Researchers also had to be aware of the limited budget, which allowed for a maximum of two field trips. Research fellows were invariably overoptimistic in their plans and had to adjust them to fit reality.

2.4. Data collection and analysis

All the deep dives had to supplement the data provided by the companies with primary data gathered from surveys and interviews. Survey design, data collection and data entry collectively took over 14 weeks on average (Figure 4). The survey questionnaires took an average of 25-35 minutes to conduct. In some cases, administrative data were helpful to identify basic client characteristics (such as data on gender from the client bases of telecoms providers), but did not have the requisite information on outcomes needed to address the research question. Deep dives invariably needed to build in a significant time period for surveys, interviews and focus groups to amass a sufficient database for analysis. The research fellows were generally able to achieve this within their time and resource constraints, despite the further restrictions imposed by the COVID-19 pandemic during 2020-2021.

Figure 4: Average number of weeks by phase



Source: Survey of research fellows

Expert advisors and the EIB were able to comment on most of the questionnaires and survey guides. Some research fellows conducted interviews, but most employed local survey companies. Some conducted in-person interviews while others carried out interviews by phone or online, particularly as a result of the COVID-19 crisis. Research fellows reported generally being satisfied with the data collection. Pilot surveys were done in some cases. Expert advisors were not usually able to check on the quality of the data, except when the research fellows flagged anomalies. Expert advisors assumed that the research fellows had done the requisite due diligence to assure reliability.

Data analysis, which lasted just over ten weeks on average, was undertaken by research fellows. They often sent work in progress to be reviewed by expert advisors. An initial report describing the data was usually circulated prior to a draft final report that tested the research hypotheses. The research fellows were generally responsive to addressing the issues raised during the reviews, according to the expert advisors. The research fellows reported that they received useful feedback from the expert advisors.

The EIB and GDN recognised the importance of active engagement with the companies that were being assessed. At the recommendation of the EIB, a "data note" was produced after the data collection phase and before the final reports. This kept momentum, ensured companies were properly informed about the status of the research, and was an opportunity to share preliminary insights of the research.

2.5. Review

The draft reports were reviewed by the expert advisors and the EIB. This progressed faster when the initial data descriptions had been shared earlier. The review process generally took a relatively short time, usually four to six weeks, thanks to the responsiveness of the research fellows.

Most of the issues raised by the expert advisors and the EIB during review were technical, many dealing with the robustness of the causal claims in the report. Data limitations (such as the lack of an adequate baseline or control) meant some final reports had to include substantial caveats. Many comments concerned the logical structure and readability of the report. The reviews required an average of two to three iterations to produce a draft final report.

After the technical review, companies were asked to comment on the final report: many provided suggestions or factual corrections. The research fellows reported these comments to focus on clarifying and simplifying the write-up to make it understandable to a general audience. The firms' reactions are detailed in Chapter 3.

3. What was produced by the deep dives?

Table 4 lists the final reports that have been shared and discussed with the companies and other stakeholders. These are the outputs of the deep dives. Chapter 3 reports on the results of that sharing. Here we consider the products for each research cycle in terms of reports and databases and discuss whether and how they met initial expectations.

16 research studies were completed: ten of these reports are publicly available and six are expected to be shared publicly by April 2021. Most of the reports reached a conclusion of potential commercial value to the relevant businesses. These included information on the nature of their client base and the demand for their services.

Many of the reports addressed the social returns. For example, there was much interest in findings related to the client base, particularly the income groups that the company reached. This sometimes led to considerations of how to reach less wealthy market segments (see Chapter 3 on the uses of the deep dives).

Table 4: Deep dive summary

Report title and author names	Research cycle	Databases generated (availability and storage of data sets)	Journal article prepared	Main conclusion
Potential Impacts of the Africa Energy Guarantee Facility (AEGF) Oluwasola Omoju	1	No	No	- Risk mitigation instruments have been used to mitigate energy investment risks in Africa with success. - These instruments need to be provided alongside other forms of support such as technical assistance. - Major demand-side factors limit the adoption of these instruments. - Evidence from similar instruments for small and medium enterprises points to significant positive effects. - The AEGF has good potential to reduce financing costs, which would have a major impact on development indicators.
M-BIRR: Mobile Banking in Ethiopia Yesuf Awel & Eleni Yitbarek	1	Yes	Submitted	- Determination of the social value of providing utility bill payments by mobile money. - Proof that the willingness to pay is higher than the marginal cost and that there is potential business demand. - Determination of technical, legal and regulatory needs.
Première Agence de Microfinance (PAMF): Promoting Financial Inclusion in Côte d'Ivoire Uwingabiye Gloria & Guylaine Nouwoue	1	Yes	Submitted	- PAMF is actively contributing to financial inclusion in some of the poorest areas of Côte d'Ivoire. - Clients are highly satisfied with the products and services. - Most loans invest in business activities. - Successive borrowing has a positive impact on the economic and financial performance of microenterprises. - No evidence could be found of an impact on longer-term fixed asset accumulation. - Women make up over half of the client sample, although the impact on business outcomes is less pronounced for female entrepreneurs.
poa! internet: Community Internet in Low-income Areas of Kenya Charles Yaw Okyere	1	Yes	Submitted	- Improved access to unlimited data has changed the way students and teachers use the internet. - Students in poa! internet's partner schools are more likely to benefit from ICT training during school hours than their peers in comparable schools and make more use of the internet for private purposes, including accessing educational content. - There was no evidence that the internet was being used to enhance learning outside of ICT training. - Improvements in school ICT infrastructure may be needed to translate better internet access into enhanced student achievement.

Report title and author names	Research cycle	Databases generated (availability and storage of data sets)	Journal article prepared	Main conclusion
Irrigation Technology Adoption and Microfinance in Rural Benin: RENACA's Agricultural Loan Programme Mariam Sangaré & Bassirou Sarr	1	Yes	Submitted	- Farmers in rural Benin find it difficult to access finance to improve agricultural productivity (although most respondents are financially included, only a minority have accessed a formal loan). - RENACA is contributing to financial inclusion, in particular by providing investment assets and working capital (fertilisers, seeds and so on), which appears equally important to farmers. - Lack of collateral reduces the ability of farmers to access finance and to invest. - Farmers who rent or borrow land indicated a lower willingness to invest in their activities and adopt advanced technologies, including irrigation equipment (they invested in transport, which allowed them to get perishable products to markets in time but did not improve productivity in the long run).
Impact of Customer Relationship Management Software on Firm Performance in Kenya Abdi Yuya Ahmad & Laura Nelima Lalampaa	2	Yes	Submitted	- Information technology infrastructure, firm size and informal competition are positively associated with the adoption of CRM software. - CRM software has a positive and significant impact on two measures of firm performance (customer volume and sales). - Government investment in ICT infrastructure could play a critical role in fostering firm performance and overall economic growth in a sound investment climate.
Provision of Quality and Affordable Maternal Health Care in Developing Countries: A Case Study of NEST in Senegal Hamidou Jawara & Gilles Quentin Kane	2	Yes	Submitted	- Clients are middle-class women (although few are among the extremely poor), including those without health insurance coverage and those working in the informal sector. - Providing information via visualisation raises both willingness to pay and use of the service more effectively than providing basic information. - Overall, participants are satisfied with the services offered but highlight some possible improvements.
Finance Access and Growth-oriented Investments by Manufacturing Enterprises in Ethiopia Muhammed Abdella Usman & Seid Hassen Mohamed	2	Yes	No	- About 27% of the surveyed enterprises obtained a bank loan, whereas 10% relied on informal sources. - Bank collateral requirements are one of the main causes of loan applications being rejected. - A large share of debt money is directed to investments aimed at improving products and industrial processes. - Enterprises rarely obtain finance from private equity or venture capital firms - Overall, firms that use debt capital show a capital growth rate at least 7.6% higher than that of firms with no debt capital.

Report title and author names	Research cycle	Databases generated (availability and storage of data sets)	Journal article prepared	Main conclusion
Microfinance Loans, Women's Economic Empowerment, and Poverty: A Case Study of Baobab Microfinance Company Franklin Amuakwa-Mensah & Edgar Cooke	2	Yes	Submitted	- Data suggest that BMC clients experience lower poverty. - This is consistent with focus group discussions with control respondents and BMC clients. - Clients are not always among the extremely poor, and female clients are more economically empowered.
Understanding the Sustained Impact of SOKO on Artisans in Kenya: Empirical Evidence from Kibera Timothy Kinoti & Soazic Elise Wang Sonne	2	Yes	Submitted	- On average and across different socioeconomic outcome indicators, worker artisans, including female artisans, affiliated with SOKO do better than their non-SOKO counterparts. - In SOKO and non-SOKO workshops, female worker artisans are more vulnerable to poverty than male worker artisans.
The Impact of Microfinance on Business Investment and Outcomes in Haiti Boaz Anglade	3	Yes	No	- Receiving subsequent loans from ACME had some positive impact on Haitian businesses. - There was a positive impact on business expansion through a different set of outcomes. - The firms that had received more than three cycles of loans had a higher probability of hiring new workers in 2018, a higher probability of opening new branches, and a higher likelihood of increasing merchandise stock as a way of expanding their business.
Development Bank of Jamaica Jacob Novignon	3	Yes	No	- The credit amount received had a positive impact on the performance (particularly profits) of microenterprises and small and medium enterprises. - The impact of the loan amount on performance was significant for firms that received the loan at least one year before, those that were older (over 40 years) and those owned by males.
Impact of Rensource Solar Energy on Subscribers in Selected Markets in Nigeria Esther Leah Achandi & Enoch Owusu-Sekyere	3	Yes	No	- Subscription to Rensource solar energy has a positive impact on monthly earnings, sales volumes and merchant profits.
Cost of Funding Infrastructure in Hard Currency Nnamani Alexander Uchenna & Kebba Jammeh	3	Yes	No	- TANESCO can record positive net benefit from hedging starting from the effective date of an off-take agreement and this benefit increases with the lifespan of the power purchasing contract. - The results reveal TANESCO's low willingness to pay for currency hedging.

Report title and author names	Research cycle	Databases generated (availability and storage of data sets)	Journal article prepared	Main conclusion
The Impact of Digital Consumer Goods Distribution on the Small Retail Sector in Nigeria: The Case of TradeDepot Matthew Townshend & Jaison Chireshe	3	Yes	No	- ShopTopUp users internalise lower supply prices to reduce consumer prices for selected products. - The usage of ShopTopUp is associated with a reduction in both the gross profit margin and the duration of product stockouts.
Deep Dive Report on the Impacts of Nachtigal Hydro Power Company Investments on Vocational Education and Training in Cameroon Nantongo Mary Gorret & Mireille Ntsama	3	Yes	No	- The results show that the activities of Nachtigal Hydro Power Company were unable to improve enrolment but increased school attendance among students.

Source: EIB-GDN (2019, 2020, 2021)

It is difficult to say whether the conclusions that had commercial and social ramifications can be generalised more broadly for development. All of the studies are of interest to those who have a specific stake (investors, companies and firms whose operations are being evaluated). Moreover, some of the conclusions will be of interest to other investors who may wish to consider the social impact of investments in the same sectors. The expert advisors felt that six of the ten studies completed in cycles one and two could have results that would be of interest to the wider development community because of their generalisable findings or methods. The deep dives were implemented in a varied set of contexts and were not case studies purposefully selected to provide generalisable conclusions. Applying the conclusions to other contexts would need to be done carefully to ensure that the comparisons are appropriate.

Expert advisors felt that the final outputs were the best that could be achieved under the circumstances – the relative youth and inexperience of the research teams, the fact that they were "parachuted" into ongoing operations, and the time and budget constraints of the process. The work has been thoroughly reviewed. Eight of the papers from the first two cycles have been scrutinised by blinded external referees, and a special issue of the *Journal of Development Effectiveness* will publish about half of the deep dives. As of March 2021, most of these papers are at the "revise and resubmit" stage. If enough papers are accepted for a special issue of the journal, this is a good indicator that the overall quality of scholarship meets academic standards.

Another principal outcome of interest was the building of research capacity. A survey of the research fellows at the conclusion of the project showed that the following were most valued:

- Methodological skills: many research fellows felt more confident and have been using the methodology in their projects or teaching.
- Soft skills and communication: some research fellows feel more confident in their relationships with colleagues and other stakeholders (including private-sector firms and companies).
- Completing a research project from beginning to end: some research fellows learnt a lot about the different steps, and now feel more confident and better able to guide students.

Most of the research fellows are continuing their careers as researchers.

The EIB, expert advisors and GDN largely agreed with the research fellows' conclusions. If properly supported – in terms of budget and mentoring – relatively young researchers at the beginning of their careers can produce reports that are relevant to many impact investing stakeholders and are rigorous enough to meet academic standards. Implementing such deep dives is a powerful way to build the capacity of talented young researchers. Substantial support was required to make these studies a success: a programme of deep dives structured in this way would likely be attractive only to implementers with an explicit mandate or objective of capacity building.

4. Main lessons

What were the main lessons learnt about how to conduct these types of programme? This section draws on the findings of this chapter.

First, deep dive research must consider, from the outset, the general lack of existing data. The research should make the best possible use of data available. Deep dives may be easier if monitoring

data collected and stored by project promoters are available (respecting relevant confidentiality concerns); if the projects assessed are operationally advanced enough to generate data on outcomes and impact; and if the opportunities offered by new data sources and data analysis techniques (for example, machine learning) can be exploited. However, in most cases researchers will need to plan and budget for primary data collection.

Second, diverse perspectives and needs must be considered and addressed. Company stakeholder/investee engagement is key. The buy-in of the small business, for example, that has secured financing is essential to ensure successful cooperation and data sharing. Early and frequent interactions between companies and researchers are needed to build trust.

Researchers, on the other hand, have careers to build and manage and most are aiming for academic positions. When studying development impact, they want to mobilise their academic knowledge and instruments to investigate economic and social returns to a given investment and to produce academic studies recognised by their peers. Researchers and expert advisors tend to focus on this and on building local capacity. EIB operational staff want to determine the likelihood that their project will be a success, and what the implications might be for similar future investments. Investment funds and firms do not have time to spend on studies that do not bring them useful business-related insights.

An ongoing issue is how to formulate research questions that will be most relevant to the investment funds and the companies in which they invest. Defining the right research questions implies a productive tension between research objectives driven by academic ambition and feasibility, and firms' operational concerns driven by practical relevance and importance in terms of business and financial objectives. As discussed in Chapter 1 (aspirations 1 and 2 in Box 5), deep dives should align stakeholders' interests. That tension is at the core of the connection between research and operations, and the deep dive programme has explored the connection.

Getting stakeholders to allow and actively participate in projects was a priority, especially given the innovative nature of this first deep dive exercise, so the choice of projects was relatively opportunistic. To broaden their contribution to impact measurement, it will be useful to focus several deep dives on the same sector and to choose the sectors strategically. Another limitation of ensuring meaningful stakeholder buy-in is that the choice of projects may be liable to self-selection bias (volunteering projects that are more likely to be successful for a deep dive). This was not an issue in our programme given the objective of demonstrating the potential of a research-based approach to assessing impact, but should be considered when scaling up the approach.

Third, researchers must be flexible about research approach and methodology (aspiration 3 in Box 5). Researchers, especially those with less experience in operational matters, can allow the research methodology and related empirical techniques to drive the research question, indicating a need to develop capacity in defining issues of interest and shaping policy or operationally relevant research questions. Acquiring such flexibility was an important aspect of the capacity that research fellows built throughout this programme.

One way to ensure that technique does not overshadow relevance is to develop a clear analytical framework to clarify a theory of change that identifies the assumptions used in linking investments to outcomes. Such an initial framework was available for the majority of the projects and would have helped focus the work in all others. Another approach is to be flexible on technique. It was challenging

to choose an appropriate methodology for the evaluation work, knowing there is no "gold standard" and that conditions for a methodological approach are not always met. There may be a trade-off between technical "purity" and feasibility when studying an operationally relevant research question in a time- and process-bound context.

Fourth, conclusions must be tailored to the limitations of the analysis. While the research fellows could conduct technical analysis of the data, it was more of a challenge to draw firm conclusions that are not generic and add value to the global community, as well as being operationally relevant. Some conclusions that are useful operationally may not seem notable in terms of changing the broader development discourse. Other conclusions may not be deemed sufficiently operational by the company clients, financial intermediaries or the EIB. Researchers must be realistic about the conclusions they can draw and what is justified by the analysis and data.

Fifth, reports must be communicated in the correct form for accountability, learning and promotion (see the three functions of evaluations in Chapters 1 and 3). The importance of this dimension can be underestimated. EIB staff had to invest a significant amount of time to make the studies fit for publication and to produce comprehensible summaries that could be shared with stakeholders or as blogs. Although they received some communication training, the research fellows struggled to connect with non-academic audiences. This area of capacity building needs greater weight in future programmes.

Sixth, building research capacity is a slow process (objective 3 in Box 5). A limited amount of technical capacity could be built in the short time period available. The preparatory seminar on methods and research design helped but could not prepare research fellows for deep technical embedding. The deep dives exposed the research fellows to new areas or engaged them further in familiar areas, exposing them to operational constraints in the "real" world. The research fellows and expert advisors noted that capacity had been successfully built in terms of discussing research plans with non-researchers (companies) and negotiating appropriate research questions (that met academic standards and other stakeholders' needs and interests). The induction seminar included two days of communication training, which exposed research fellows to some of the challenges of negotiating a work plan with third parties, and eventually framing and reporting their results.

Seventh, the costs of the deep dive were neither insignificant nor prohibitive and were lower than those of typical impact evaluation work. Beyond the production of deep dives, researchers from low- and middle-income countries had the opportunity to apply their technical expertise to real-world problems and gain experience in dealing with practitioners and policymakers, especially in the private sector. In many cases, the EIB, the firms and the intermediaries judged many of the findings useful. Despite the budget and time limitations, this exercise showed that, with proper management, research can productively combine operational concerns and technical rigour.