

Evaluating collaboration between the private sector and development cooperation – how to prove the added value?

To achieve the Sustainable Development Goals (SDGs) by 2030, the international community increasingly relies on the private sector, both for mobilising additional finance and for project implementation. However, research by the German Institute for Development Evaluation has revealed that the quality of evaluations in the field is currently not sufficient to accurately assess the added value of private sector engagement.

By Valerie Habbel

Increasingly, private sector companies are being involved in public sector development projects. Many of the projects and the instruments used seek to lower the risks that companies face when investing in developing or emerging countries, for example by providing them with a guarantee in case of project failure or with a grant that matches the company's investment. Development actors expect that this cooperation contributes to channelling more finance into the Global South while taking advantage of the innovation and creativity that is associated with the private sector. While this cooperation seems like a win-win situation, a closer look reveals potential downsides. Crucially, how do you ensure that private companies looking to invest in developing countries actually require the additional push provided by development finance? Some companies may have made the same investment either way, in which case public resources would have been more useful elsewhere.

Development actors regularly commission or carry out evaluations of their projects to understand what worked and what didn't. However, is the quality of these evaluations sufficient to justify the increasing use of public resources to engage the private sector in developing countries?

An evaluation synthesis by the German Institute for Development Evaluation (DEval) aimed at identifying the (intended and unintended) effects of private sector engagement and the conditions that are conducive or unfavourable to them, as well as assessing the quality of evaluations in the field. The quality assessment was also used to exclude evaluations and studies deemed inadequate to ensure that the findings of the synthesis were based on reliable evidence. A total of 75 evaluations and studies from international and national actors in development cooperation were assessed using nine indicators that are based on recognised evaluation standards by the Organisation for Economic Cooperation and Development (OECD) and the German Evaluation Society



(DeGEval). The indicators ranged from descriptive aspects, such as whether the evaluations describe their data sources and evaluation questions, to more complex issues such as whether they justify their choice of methods. Only evaluations and studies that exceeded the threshold of 60 per cent of the best possible rating were included in the synthesis.

The quality of evaluations differs widely between actors

Out of 75 assessed publications, only 51 were deemed reliable enough to be included in the synthesis. While almost all evaluations by international donors/ agencies and academic studies were included, only around half of evaluations by GIZ and KfW made the cut (*see Figure*).

Generally, both evaluations and studies score well reporting on descriptive aspects: most contain a description of the subject of the evaluation, its area of inquiry, sources of information and the procedural steps specified for carrying out the evaluation or study. However, only few evaluations discuss the appropriateness and limitations of the methods they used or relate their findings and conclusions to the underlying data and data analysis. This limits the transparency of their findings.

Outcomes and impacts are usually estimated rather than measured

Most projects in international development use a Theory of Change, a methodology that can be applied both to plan a new project and to evaluate it later on. It explains the process of change by outlining the causal pathways between inputs and expected effects on different levels: outputs, outcomes and impact. While outputs are the products, goods and services that result from an intervention, outcomes are the short- and medium-term effects, such as participants' increased knowledge in a certain area. The impact refers to the long-term effects, such as better employment opportunities.

To measure the success of a project, the expected effects are generally operationalised through indicators at the planning stage (e. g. 50 trainings carried out by end of 2020), and the completion of these targets is then used in the evaluation to assess the effectiveness of the project. The evaluation synthesis found that in most evaluations, outputs are operationalised well by indicators which are then reported on in the evaluations. However, almost half of the evaluations do not include any indicators at the outcome or impact level, which means that they have no measurable basis for statements regarding medium- and long-term effects.

Worse, in a few cases, indicators are designated as outcome indicators, but more properly relate to the output level, which could lead to unfounded claims of wider-reaching effects.

Instead of measurable indicators, claims regarding outcomes and impacts are often based on estimates. While estimation models can be useful tools, for example when precise measurements of the impact of an intervention on the economic development of a country or sector are not possible, the underlying indicators and assumptions of these models are not always made transparent to the reader. In addition, both estimates and measurements are prone to “over-reporting” of effects: for instance, an estimate of new jobs created as a result of a project usually does not consider whether these job gains were at the expense of competing companies that might have lost jobs.

Another difficulty with measuring impact is the attribution to the specific intervention – even if new jobs were created at a company, for instance, how do you know whether these are due to the intervention or external factors such as an economic upswing? To answer this question with certainty, a rigorous impact evaluation (RIE) is necessary, where participants (in this case companies) are randomly assigned to an intervention and a control group. Evidence on various metrics such as the turnover of the companies would be collected before and after the intervention. While the metrics might change over time in both groups, changes should be greater in the intervention than in the control group if the intervention had an impact. Although RIEs provide the highest degree of certainty for attributing changes to an intervention, they are often not feasible. One challenge is the random allocation of participants: companies usually apply for funding and are selected based on specific criteria rather than randomly. RIEs also tend to be rather cost- and time-intensive.

Additionality often not considered

Let’s say a private company receives funding from the German government to provide trainings to its agricultural partners in a partner country to improve the quality of its imports. An evaluation using a rigorous design might then conclude that the farmers’ standard of living has significantly improved as a result of the intervention and the project would be considered a success. However, what if it later emerged that the company was planning to do the same training even without the public funding, using only its own resources? In that

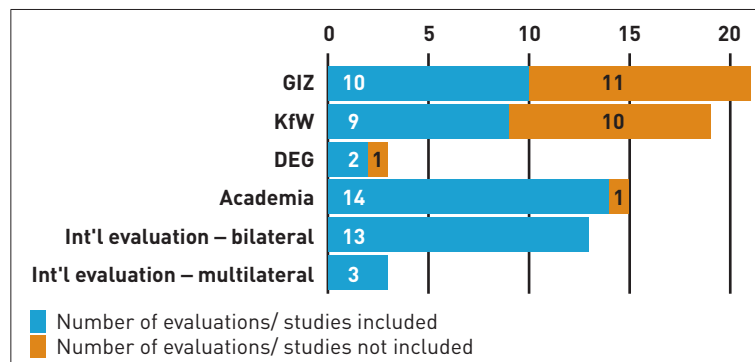
case, even though the project had an impact, the public funding would have not been used effectively, and financial additionality (see Box) would be lacking.

This example highlights the importance of examining additionality, particularly in projects that engage the private sector, where there is the possibility of deadweight effects. It is also clear that additionality needs to be considered at different stages of the project cycle: at the very least at inception, to ensure that only projects that are additional are funded in the first place, and during the evaluation, to understand whether the assumptions made regarding additionality at the beginning actually materialised.

Despite the importance of additionality, the evaluation synthesis by DEval found that most evaluations or studies (35 out of 51) do not even discuss additionality, let alone use a systematic approach to assess it. Where additionality is discussed at all, it is mainly from an ex-post perspective. This makes it problematic to assess the degree of additionality that existed at the beginning of the project or instrument, and how this might have changed during project implementation. Some evaluations and studies also report having no suitable evidence base on which to assess the additionality of the respective project or instrument.

The way forward for evaluating projects that involve the private sector

The scarcity of public resources that are available to close the financing gap is often cited as a reason for increasingly engaging the private sector to achieve development objectives. However, this argument only holds if the private sector contributes additional resources to stretch public resources further, and if these ac-



tually contribute to the objectives set. But how do we know whether this is actually the case? At the moment, evaluations and – in some cases – academic studies are the only evidence policy-makers have to assess these claims. To justify the use of public resources for private sector engagements, evaluations therefore need to provide convincing evidence on the additionality and impact of interventions.

As this article has outlined, evaluations do a good job in some regards, usually providing the required descriptive information on the project as well as quantitative information on output indicators. However, the evidence on outcomes and impact is much more vague, often based on non-transparent estimates and assumption. In addition, additionality is not even discussed in most evaluations. We recommend that development actors such as GIZ and KfW in Germany improve the evidence basis on private sector engagement by, among other things, systematically examining additionality both when new instruments or projects are created and when they are evaluated, as well as by improving the assessment of impact by explicitly measuring it for projects of high relevance; other evaluations may use theory-based approaches or estimation models if they set out from transparent assumptions and indicators.

Valerie Habel is a former evaluator and team leader at the German Institute for Development Evaluation (DEval) in Bonn, Germany. Contact: amelie.eulenburg@deval.org

Additionality

The Organisation for Economic Co-operation and Development (OECD) makes a distinction between financial and development additionality. An official investment is defined as financially additional when it supports a company that is unable, without public support, to obtain financing of a similar amount or on similar terms from local or international private capital markets, or when it mobilises investments from the private sector which would otherwise not have been invested. Development additionality, on the other hand, is defined as the development impact resulting from the investments which would otherwise not have occurred. Especially in relation to private sector engagement, the review of additionality is of key importance when drawing conclusions about the efficiency or cost-effectiveness of projects and instruments, since there is a risk that public funding might finance activities that the private sector would have financed anyway, even without the subsidy component.