

IMPROVING DEVELOPMENT POLICIES WITH IMPACT EVALUATIONS

The debate about the impact of development co-operation is as old as development co-operation itself. The debate about the relevance of impact evaluations is a more recent phenomenon. Wouldn't it be better to make use of the already scarce resources in projects instead of spending them on these costly evaluations? Our authors demonstrate when impact evaluations make sense – and when they don't – and why we can't really do without them.

By Bartłomiej Kudrzycki and Isabel Günther

Over the last 25 years, the share of the world population that lives in destitute poverty has dropped from 35 per cent to 10 per cent, and the share of people who are undernourished has fallen from 19 per cent to 11 per cent. These numbers hint at the progress we have made towards eradicating global poverty. Policy-makers, government officials and development practitioners can certainly take some credit for these improvements in people's lives – but how much? Have projects designed, financed and implemented by various organisations contributed to this success, and to what extent?

Most importantly, despite past collective success, what can be done to do even better going forward? About 700 million people still live on less than 1.90 US dollars purchasing power parity a day, and about 800 million are still undernourished. These are unacceptably high numbers.

By bringing rigorous analysis to empirical data, impact evaluations allow to measure the effect of development interventions and to generate knowledge about how a programme works and how its design and results can be improved. Impact evaluations are thus primarily a tool for learning and improving development interventions. They are an important part of a broader agenda of evidence-based policy-making – in contrast to ideologically, emotionally or politically driven policies. For example, impact evaluations have revealed that farmers in developing countries are less constrained by their access to credit than once thought. Instead, a lack of risk coverage or psychological biases appear to be more likely barriers to farmers investing in new crops and innovative technology. Such insights are of great value to policy-makers looking for effective measures to assist farmers in adopting new technologies.

Besides providing insights for policy-makers and development practitioners, impact evaluations first of all benefit the poor. While critics argue that experimentation on the poor is

unethical, it is also evidently unethical to intervene in the lives of the poor without understanding the changes, intended or unintended, that these interventions are likely to bring about.

IMPACT EVALUATIONS – A NEW BUZZWORD?

Thanks to technological progress in data collection and the ever-increasing availability of data as well as the creation of various institutions promoting impact evaluations, the number of them being conducted has risen rapidly over the last two decades (see Figure on page 8). While in 2000, the International Initiative for Impact Evaluation (3ie) recorded less than 40 new impact evaluations related to development and poverty, by 2012, the impact evaluation repository of 3ie was publishing over 400 impact evaluations a year. Whereas for many years, impact evaluations were focusing on health questions, the number of impact evaluations has been steadily increasing in other sectors since 2006, in particular in agriculture and nutrition.

“ *Impact evaluations first of all benefit the poor.* ”

At the same time, impact evaluation has become something of a buzzword in development co-operation. Major organisations are creating entire funds and policy priorities in their name, while many practitioners are left in the dark about what impact evaluations actually are and how they are used. Several large development agencies have therefore released primers and guidance documents to address this disconnect.

The lack of clarity around the concept, combined with the high cost of impact evaluations in terms of both time and money, have result-

ed in considerable backlash, even resentment, towards impact evaluations – especially in their most famous (or notorious) form, randomised controlled trials (RCTs).

In this issue, various authors hope to clarify and elucidate what impact evaluations are and when they are effective tools for learning: we believe they have become an indispensable tool for measuring and improving the impact of projects and policies on decreasing poverty and – what is equally important – setting precise and realistic aspirations for the future.

WHAT IS AN IMPACT EVALUATION?

For many years, the development community – including the Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD-DAC) in its Criteria for Evaluating Development Assistance – used the term “impact” to refer to the final level of the causal theory of change, or log frame. This definition has been replaced in recent years, and impact evaluations are now seen as “an objective assessment of the change that can be directly attributed to a project, programme or policy”. This could be the impact of an information campaign (about the importance of crop rotation) on farmer output, or it could be the effect of introducing rainfall insurance on a farmer's choice of crops. These changes are the impact: the difference in people's lives (farmers' output or choice of crops) with and without the intervention, measured after the intervention (information campaign or rainfall insurance) has taken place.

To assess the impact of a project or policy one needs to know what would have happened to the population in its absence. This is called the counterfactual, which is a crucial component of any rigorous impact evaluation, and which can be estimated using a variety of statistical methods. In contrast to monitoring, the use of the counterfactual methods lets policy-makers



In order to measure the impact of a crop insurance scheme, one needs to know how the farmer would have fared without it.

Photos: FAO/Luis Tato (left), FAO/Lou Dematteis (right)

and other stakeholders establish the causal effects of their programmes and policies.

For example, an impact evaluation might assess the impact of a programme that aims to improve farmer crop yields by offering farmers rainfall insurance. To estimate this impact, one needs to compare the outcomes of farmers who receive rainfall insurance to the hypothetical situation in which the same farmers were not insured. Studies have found that insured farmers grow riskier crops with higher yields. Thus, impact evaluations establish the direct connection between projects or policies and measurable, observable changes in people's lives.

While the term impact evaluation comprises a wide range of methodologies, one of them has garnered the lion's share of funding, attention, and criticism in the development community: the randomised controlled trial (RCT; see also articles on pages 9–11 and 12–14). RCTs are the most well-known form of impact evaluation, but it is very important to note that there are many other methods of constructing a counterfactual to estimate what would have happened to the target population in the absence of the project or policy without resorting to randomly allocating the target group to a control and treatment group.

MONITORING – USEFUL, IF INTERPRETED CAREFULLY

To better understand what impact evaluations are, it also makes sense to clarify what they are not. Monitoring is a common, yet non-rigorous method of estimating programme effects and is hence prone to errors. Only measuring the changes of outcomes for the population before and after a development programme, there is no way of knowing if the outcome would have remained the same in the absence of the programme. For instance, a monitoring system can observe that the nutrition of a village population improves after everyone in the village has received a crop storage container. However, unless all competing explanations can be eliminated – e.g. changes in agricultural productivity, construction of a new well, changes in income, or the presence of deadly diseases – we cannot be sure that the impact is indeed a result of the intervention.

Monitoring data is nevertheless often used in development work thanks to its ease and low cost for reporting and project evaluations. Monitoring is useful when the focus is on operation, implementation, or service delivery. However, when misinterpreted as evidence of a causal relationship between a development intervention of a programme and poverty re-

duction, conclusions drawn from studies solely using monitoring data can lead to ineffective or even harmful policies – and in most cases to a waste of public resources.

WHEN ARE IMPACT EVALUATIONS USEFUL – AND WHEN NOT?

Impact evaluations are a tool for policy-makers and development practitioners to improve development outcomes based on evidence. The findings of impact evaluations can help organisations to decide whether to scale up projects with proven positive impacts or discontinue interventions lacking in effectiveness. Impact evaluations can also identify the specific point – of the theory of change – at which policies don't work as planned. For instance, the Agricultural Technology Adoption Initiative (ATAI) shows that index-based weather insurance is very effective when taken up, but that at market premiums take-up is very low (6–18 per cent) – it is at the point of take-up, not after, that rainfall-index insurance programmes seem to run aground. Impact evaluations can help to design development programmes by comparing different interventions with regard to their effectiveness. For example, poor Kenyans were offered a variety of ways to encourage saving – text message reminders, match-

ing ten or twenty per cent of savings before or after the savings period, and a simple, fake gold coin with a number for each week of the experiment that served as a physical reminder of savings. The intervention that helped farmers save the most by far was, remarkably, the gold coin.

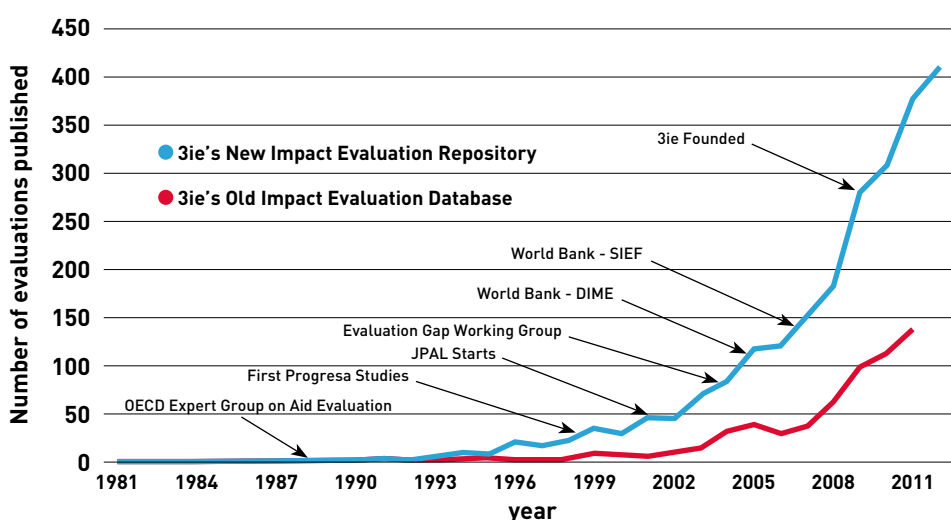
The two main drawbacks of impact evaluations are their high monetary costs and the time required for the results to come back. From the beginning of implementation to the results, an impact evaluation generally takes two years to complete, while many take longer. Average 3ie-supported studies cost 400,000 US dollars and last three years.

Hence, not all projects and programmes of an organisation should be evaluated with regard to their impact; only those where the learning potential is the highest. The project should be strategically or operationally relevant for the organisation and innovative in the sense that evidence on whether it works is needed because impact evaluations on the planned intervention are non-existent. For example, the impact of micro-credit, rainfall insurance and better price information on farmers' livelihoods have already been extensively studied. However, unlike clinical trials in medicine, the findings from impact evaluations (and RCTs) in agricultural development do not easily translate from one context to another. Rather than just providing estimates of the effects of cookie-cutter interventions, impact evaluations should hence be designed in a way to offer the opportunity to learn how context and intervention interact. For any individual study, there is little certainty that the findings will replicate in another context.

Once the number of studies run in different contexts reach a critical mass, however, impact evaluations can inform policy-makers and donor organisations whether they are following the best strategy for achieving a certain development goal; and be used for global policy-making and best practices. Systematic reviews use a structured approach to summarise the results of many impact evaluations from a particular sector or region, and give reliable indication about the success of a certain type of intervention – and are particularly useful for policy-makers and practitioners. For example, the recent evidence maps of 3ie on agricultural innovation and agricultural risks are great starting points for policy-makers working on rural development and agriculture.

Neither is it possible to analyse the impact of all types of development interventions. In oth-

Impact evaluations published by year, 1981–2012



SIEF = Strategic Impact Evaluation Fund; DIME = Development Impact Evaluation; JPAL = Jameel Poverty Action Lab
Source: 3ie

er words, not all the projects and programmes of an organisation can be evaluated with regard to their impact. For example, statistical approaches allow us to estimate the effects of specific interventions on precisely defined development outcomes, but are of little use when it comes to broad, long-term effects at aggregate levels, such as GDP growth. The time and effort required to track and measure the effect of, say, a microcredit loan on the livelihoods of individual farmers 20 years down the line far outweighs the usefulness of that information – to say nothing of country-level effects.

WHAT DOES THE FUTURE LOOK LIKE?

The debate about the relevance of impact evaluations fits into a larger discussion about various approaches towards poverty reduction. The economists Esther Duflo and Abhijit Banerjee view the role of development practitioners as analogous to “plumbers” – they argue that incremental fixes to incentive schemes and government service delivery systems add up to substantial improvements in the lives of the poor. From this perspective, impact evaluations are indispensable, as they are ideal for identifying small improvements and can be used to guide small adjustments in programme delivery.

Others argue that a focus on “plumbing” runs into the danger of missing the bigger picture, i.e. the root causes as to why some countries are able to escape poverty while others remain poor. Impact evaluations are poorly equipped to evaluate large, macro-sized programmes, structural change, regime changes, or large reforms such as trade liberalisation.

In sum, rigorous evaluations are always time-intensive and mostly costly. In the narrow context of the programme being evaluated, a rigorous impact evaluation is an imperfect instrument for accountability due to its high cost and long timeline. But impact evaluations do offer indispensable lessons on what works and what doesn't; it goes without saying that they are invaluable knowledge if we are to build more effective development programmes and spend the limited resources we have for poverty reduction more effectively and responsibly. Even if development organisations choose not to conduct impact evaluations themselves, they must know and make use of the ones existing in their field.

Bartłomiej Kudrzycki is a PhD student at the Center for Development and Cooperation (NADEL) at the Eidgenössische Technische Hochschule (ETH) Zurich, Switzerland. He is conducting research on youth labour markets in cities, focusing on the context of Benin.

Isabel Günther is a Professor of Development Economics and Director of NADEL at ETH Zurich. Her current research focuses on technologies for poverty reduction, analysis and measurement of inequality, urbanisation and population dynamics, sustainable consumption, and aid effectiveness. She has conducted most of her research in countries and cities of sub-Saharan Africa. Contact: isabel.guenther@nadel.ethz.ch

For references and more information on the projects, see the online version of this article at: www.rural21.com