Impact Evaluation of Development Programs: Challenges and Rationale
Introduction

Declared the International Year of Evaluation by EvalPartners, 2015 appears to be the right moment to reflect on the rationale and challenges for performing impact evaluations of development programs.

Every development program seeks to generate a positive impact on beneficiaries’ lives. One aspect that all these programs have in common is that they are always designed with the ultimate goal of improving the quality of life of a targeted population. However, in reality, there is no assurance that a given program is reaching its goals.

An important question to be asked after a program has been implemented is whether the underlying theory of change is aligned with reality. In other words, does that development program cause the impact it is said to have caused? Do the inputs and activities really generate the expected outputs and outcomes?

A rigorous impact evaluation (IE) is able to provide an answer to these questions. Impact evaluation is one of several types of evaluations. It is a particularly important one due to its ability to isolate the intervention’s impact. In other words, an impact evaluation is able to attribute impacts to a program and to that program alone. To do so, it contrasts the situation of the beneficiaries after the intervention with the counterfactual situation - what would have happened to them had they not participated.

There are several arguments behind the rationale for performing an impact evaluation but it remains a difficult task. On the one hand, it is a crucial tool to help policymakers decide whether to redesign a program, expand it, interrupt it or design similar interventions for other communities. On the other hand, it is time-consuming and requires careful thought, planning, and resources. The purpose of this paper is to discuss the reasons why policymakers should carry out an impact evaluation of their programs and the main challenges they may face in doing so.

Reasons for an Impact Evaluation

Identifying the Impact

The most important feature of an impact evaluation is its ability to determine the impact that is solely attributable to the program. An evaluation that compares the situation of beneficiaries before and after a given program may capture not only the impacts attributable to that specific intervention but also other impacts that would have occurred even in the absence of the program.

An impact evaluation defines impact as the difference between the situation of the beneficiaries after the program and the hypothetical situation they would be living in had they not participated in it. Therefore, to perform an impact evaluation, it is necessary to select a comparison group to represent the beneficiaries’ counterfactual situation in the absence of the program. By contrasting the beneficiary group with the comparison group, it is possible to properly identify the program’s impacts by separating them from those occurring independently of the program.

Therefore, it is often the case that a “naive before and after”2 evaluation finds an impact, while an impact evaluation finds no impact. This is valuable information for policymakers who may use it to redesign or discontinue the program and direct the resources to more productive projects.

1 This paper draws on authors’ previous work available on Barros and Lima, 2012 and INCT and SAE, 2014.
2 A “naive before and after” evaluation calculates impact through simply taking the difference between the situation of the beneficiaries before and after the program.
Validating the Theory of Change

An argument against allocating resources to the evaluation of a program is that, if program beneficiaries are volunteering to participate they must be benefiting from it or they would not participate. Theoretically, given that we believe that individuals are rational, beneficiaries would not spend their time and effort in something that would not provide some advantage to them. This line of thought implies that a program or policy must have a positive impact. If such is the case, why spend resources to perform an impact evaluation?

When there is no conditionality to enrol in a program, beneficiaries do not lose anything by participating although they may not gain anything either. Although the rationale of the program, its expected causality mechanisms and long-term goals are well defined and create expectations of a positive impact on beneficiaries, it is not certain that the program works as presumed. Therefore, evaluating a program is not about contradicting the rationality of individuals. It is, rather, a way to validate the theories that assisted in designing the program and to confirm whether or not the perception of beneficiaries and reality are aligned.

Beyond investigating whether or not there is any impact, an IE is important for enabling policymakers to “understand” the impact, that is, learning the dimensions over which the program has impact and the channels through which it emerges. Knowing this information is crucial for improving the program’s design, for its eventual adaptation to different communities, and for the identification of best practices in development.

Potential vs. Real Impact: the Role of Implementation

Another argument against allocating resources to an IE is that policymakers often confuse potential impact with real impact. Let’s take, for example, a program that was evaluated in its pilot phase with satisfactory impacts that served as evidence for the scale-up of the program in the same community. Since the evaluation of the pilot was already carried out, is it necessary to evaluate the expanded program?

If the beneficiary population in the pilot phase is similar enough to the beneficiary population in the expanded intervention, there are reasons to believe that the expanded program will have the same impacts. Its success in the pilot phase means it has a proven potential impact. Nevertheless, potential impact is different from real impact. The real impact of a program depends on its potential, of course, but it is also inextricably linked to its implementation conditions. A program may not reach its full potential impact due to flaws in the implementation process. Therefore, knowing the potential impact of a program is not a reason for not evaluating it. An impact evaluation is still necessary to understand the real impact on beneficiaries and useful to inform policymakers about the need to improve the process of implementation.

Let’s say that a program has been implemented and evaluated with successful results in a given community. Can the policymaker expand the same program to different communities assuming it will also have the same beneficial impacts? Not necessarily.

Population heterogeneity leads to different impacts. Benefits may vary depending on the targeted population and socioeconomic context. The returns of a given program in a wealthier community may differ completely from those in a poorer scenario and vice versa. Both preferences and returns are likely to vary with income levels, and with cultural and ethnic background. Therefore, if the beneficiary scenario changes, an IE is necessary to understand the impact dimensions and measure its magnitude on a population. In this situation, an impact evaluation is an important tool for the policymaker to decide whether the original design of the program should be adjusted in order to maximise its impacts for the new, targeted population.

Cost Analysis

An IE provides inputs for a program’s cost analysis. As resources are scarce, their effective and efficient allocation is a major concern for donors. Although the priority of development programs is to generate a positive impact on beneficiaries, the mere existence of an impact is not sufficient. If, in a given
society, different programs are pursuing similar goals, the choice of one program over another may be determined by a cost-effectiveness analysis that reveals which program delivers the largest impact at the lowest cost.

The allocation of resources to a given intervention often depends on a cost-benefit analysis, which takes into account all the costs related to the program and contrasts it with all the benefits – both those associated with the beneficiaries and with the externalities that might have been generated for the society. In order to have a favourable cost-benefit assessment, the net value of benefits (total benefits minus total costs) must be positive.

Choosing the Best Program Design
There are generally different types of interventions that can be designed to reach the same outcome. Let's assume a policymaker is trying to decide the best design for an anti-poverty program to be adopted nation-wide. Even if there is evidence of the impacts of each of the delivery modes on different countries, as previously discussed, there are reasons to believe that these impacts may differ according to the scenario. IE may be a useful tool for policymakers to test different policy designs in a pilot mode so as to make an informed decision on the best design for a given policy in a given context.

Challenges
Timing of an Evaluation
One of the important and challenging decisions related to an IE is when to execute it, given a budget restriction. Depending on the timing of the evaluation, different purposes and results are reached. An IE may occur during the implementation of the program or after it has finished.

An IE that is done during program implementation has the advantage of providing program managers with information on the program's performance, allowing them to correct potential implementation problems or modify the program design to better suit the original purposes. However, there is a trade-off pertaining to the decision for the best timing of an IE. On the one hand, the sooner it is done, the more modifications can be implemented on the project, thus improving its efficiency and benefits. On the other hand, if it is done too soon, because some program impacts may take some time to arise, impacts might be underestimated or unnoticed.

An ex-post IE is executed after the conclusion of the program. It is a very important accountability tool to show donors the returns for that particular investment. It is also important because it is more likely to capture the longer run impacts of the intervention compared to a mid-term evaluation. Finally, it contributes to the development community as it helps to build evidence of best development practices.

Coordination between Managers and Evaluators
An IE demands strong coordination between program managers and evaluators. A constant dialogue between them is crucial for a successful evaluation. The first step is to achieve a full understanding of every detail of the program by the evaluators, including the program activities, the expected causality mechanisms and the expected goals. Usually providing program documents alone to the evaluators is not enough – there are often important undocumented details related to a program or the documentation may not quite represent the reality of the program. A deep understanding of the program by evaluators is required for the development of appropriate instruments. If an evaluator is unsuccessful in fully understanding the program, he may fail to develop instruments to capture all the dimensions of expected outcomes and the IE may find that the program has no impact, whereas the program does have an impact but on a dimension that was simply not measured.

Ideally, an IE measures all aspects of a program, but this is often not feasible. An ideal survey would have as many questions as possible but large questionnaires are not only very costly and time consuming but they also come with difficulties in keeping the respondent's attention for a very long time. Therefore, choices have to be made about what to exclude and include in the questionnaire. These choices will determine the quality of the evaluation. It is crucial that they are made jointly by evaluators and program managers.
Counterfactual
It is decisive for the success of an evaluation to identify what would have happened in the community in the absence of the project, and what kind of services or goods beneficiaries would have access to instead of the ones offered by the intervention. Since the hypothetical situation of an intervention-free area does not exist, a comparison group must be selected to represent the counterfactual. This requires careful thought and is one of the greatest challenges of an evaluation, as an inappropriate comparison group may invalidate the evaluation results.

The Size of the Sample
Evaluators must also determine the size of the sample to be studied. The ideal sample size depends on the magnitude of the impact and its variability. The larger the impact variability and the smaller the magnitude, the larger the sample needed to detect the impact.

The challenge here is to foresee the impact magnitude before the program is implemented. Failure to foresee the impact magnitude may lead to overly large samples, which would be too costly and wasteful, or overly small samples, which would be insufficient for the evaluation to identify the impact.

Conclusion
There are several reasons for performing an IE of a development program. It is very important not only for the proper identification of program impacts, but also for measuring them and understanding their dimensions. An IE may help program managers to make informed decisions about the expansion, redesign, or interruption of an intervention. It is also useful for the international development community to build a knowledge base on best practices in development. It is, however, not an easy task. Carrying out an impact evaluation is costly and poses several managerial and technical challenges.

References

Institutos Nacionais de Ciência e Tecnologia and Secretaria de Assuntos Estratégicos. 2014. “Avaliação de Impacto – Conceitos, Utilidades e Desafios”

PROFILE OF THE AUTHORS
Lycia Lima is an Associate Researcher in the Applied Microeconomics Centre – FGV and Executive Coordinator of the Centre for Learning on Evaluation and Results for Brazil and Lusophone Africa. She has eight years of experience conducting impact evaluations of development programs. Previously Lycia held positions in the World Bank in Washington D.C., the UN in Geneva, the Secretariat of Strategic Affairs of the Presidency of Brazil and the Government of Minas Gerais State in Brazil. Lycia is a PhD candidate in Public Administration and Government at the Fundação Getúlio Vargas, Brazil. She holds a Masters in Development Economics from the University of London’s School of Oriental and African Studies, UK.

André Portela Souza is a Professor at the São Paulo Economics School – Fundação Getúlio Vargas. He is Coordinator of the Applied Microeconomics Centre and Executive Director of the Centre for Learning on Evaluation and Results for Brazil and Lusophone Africa. André holds a PhD in Economics from Cornell University and a Masters in Economics from the Universidade de São Paulo, Brazil. He also studied Economics at the Universidade Federal da Bahia, Brazil.

Dalila Figueiredo is Project Coordinator at the Centre for Learning on Evaluation and Results for Brazil and Lusophone Africa. She holds a Masters in Economics from the Nova School of Business and Economics in Lisbon, Portugal, and an Academic Masters in Economics from São Paulo School of Economics – Fundação Getúlio Vargas. She is a graduate in Economics from the Nova School of Business and Economics in Lisbon, Portugal.