Impact Evaluation for Policy Making: A close look at Latin American countries with weaker research capacities

Maria Laura Alzua, CONICET, CEDLAS at Universidad Nacional de La Plata

Habiba Djebbari, Laval University, CIRPÉE and IZA

Martin Valdivia, Grupo de Análisis para el Desarrollo (GRADE)

SIG WORKING Paper 2013/2

Contents

Foreword

IDRC's Supporting Inclusive Growth (SIG) program supports research that goes beyond descriptive assessment and offers viable policy options. This may involve policy analysis, and impact evaluation (IE) of specific policies or programs. SIG does not favour any particular method. The starting point for the assessment of SIG research projects is the relevance of the research questions, and selection of the most appropriate methods ought to follow.

There have been important methodological advances on the field of IE and the use of experimental and non-experimental approaches. The use of IE (especially Randomized Control Trials - RCTs) in the developing world has increased considerable. IE studies offer a channel to engage in policy debate with program managers, policy makers and other relevant stakeholders and to respond to some relevant policy questions. Recent trends in the field of IE and how IE is being applied in practice make it relevant to have a closer look at how this field is evolving.

Despite the promise of IE to generate useful evidence for program and policy debates, design and reform, in many cases the actual use of IE to inform program and policies is unclear. In some cases, impact evaluations are a pre-condition for loans to finance a particular intervention. If IE are not reflecting the real demand of policymakers and program managers, the prospects of its real use for policy debates are reduced. And does IE answer the most relevant policy questions? IE have clear limitations to the types of questions that can be answered: while it is relevant to know whether an intervention has an impact or not, IE say little about the reasons behind the findings. We also learn little about who: RCTs estimate the mean impact on those participating in an intervention but what about the distribution of impacts from standard RCTs? Furthermore, randomization and isolation from interventions might not be possible in many contexts and for many policies and programs: some argue that RCTs are only applicable to 5 to 10 per cent of development interventions. Heterogeneous and changing contexts of interventions impose many limitations to external validity of RCTs. There have also been growing ethic concerns regarding RCTs. The quick rise of IE is in part a response to some donors and academics preferences, which over-emphasize favored methods for IE. Does this trend marginalized or crowded out other ways of knowledge, learning and accountability? Does it marginalize some researchers and research communities?

These questions are examined in this paper that IDRC recently commissioned on the demand and supply of impact evaluation in Latin America and the Caribbean. The review includes more than 315 impact evaluations developed in 21 countries in the region since 1995. It examined the policy issues covered and methodologies used; the research actors and implementing agencies involved; the degree of involvement by the government *vis a vis* the cases in which the key implementing actor is a development NGO; and the role of donors and other funding agencies. Three case studies in El Salvador, Dominican Republic and Peru show the diversity of institutional arrangements and their implications in terms of the role of local researchers and the use of the studies to effectively shape policy design and implementation. The paper reveals some of the

limitations associated to how IE is being currently applied to development. It shows that in many of the IE being conducted in the region local researchers are underrepresented in the design and development of the studies; many of them being involved in field work but not on the actual analysis of the data.

IE, and RCTs in particular, are relevant tools for policy research and can make a difference to development. But the way IE is being implemented has implications for research and policy. This paper is a contribution and invitation to rethink the way the IE field is being shaped, and suggests priorities for renewed directions.

Carolina Robino

IDRC

Arjan de Haan

IDRC

1. Introduction

With the transition to market-based systems, many countries are designing and implementing social policies targeted to specific populations, e.g. social protection to poor people, job training programs to the youth and the unemployed, agricultural development programs to farmers. Decision-makers, donors and taxpayers are interested in knowing whether the programs have the expected benefits, hence demanding rigorous assessments of the impacts of social policies and programs.[1] This could further foster accountability in public expenditures and may lead to improvements in program design and implementation, if installed within the right institutional framework (Briceño and Gaarder, 2010) and in combination with other evaluation tools such as process evaluations, monitoring mechanisms, qualitative information, etc.

Interest in Impact Evaluation (IE) has grown rapidly in Latin America and the Caribbean. However, rigorous impact evaluations are still very much concentrated in a few countries (Mexico, Colombia, Chile, Brazil, Argentina, Uruguay, and Peru to a certain extent). This study looks at the way impact evaluation studies are being produced and used for policy making in a sample of countries in the region that are a priori considered less capable to absorb the current trend observed in other more developed countries in the region. This study is threefold: (i) we performed a systematic search for the studies that evaluate the impacts of programs and policies with sound identification strategies. Then we analyzed time trends and the key actors in the demand, production and funding of the studies. We also (ii) carried out three case studies (of the Dominican Republic, El Salvador and Peru) to explore the institutional factors that work in favor and against the demand and use of rigorous impact evaluations for policy making, and (iii) we searched and identified training activities for the promotion of IEs, and their role in shaping policies and programs.

We systematically reviewed IE studies in a selected set of countries where the use of IE approaches for program evaluation is more limited, building on the previous effort by Bouillon and Tejerina (2007)[2]. Besides learning about different IE experiences for countries where IE is less common, we compared the dynamics vis-à-vis IE for this group of countries to those in the rest of LAC: Are they following the same trends in terms of production of IE studies and institutionalization of IE as a basis for decision-making, only with some delays? Are there challenges specific to this group of countries? In order to better understand their specificities, we also designed and conducted a series of case studies to deepen our understanding of the processes underlying IE studies. Our objective is to analyze not only the magnitude and characteristics of the production of IE studies in our sample of countries, but see how IE studies inform policy-making and program management, and how, these, in turn, influence the choices of methods for IE studies. Thus, we complement the study with three case studies that help us delve deeper into the factors that favor and limit the production of IE studies and the systematic use of such studies for policy making. Finally, we take a preliminary look at the supply of training in the region, considering that the level of research capacity is a factor influencing this two-way relationship between the policy sphere and IE efforts. We collected information on training courses on modern methods of impact evaluation of social programs to local researchers and policy makers in the countries under study.

[1] Heckman, Lalonde and Smith (1999) stress this aspect in the following excerpt: "An emphasis on objective publicly accessible evaluations is a distinctive feature of the modern welfare state, especially in an era of limited funds and public demands for accountability."

[2] See the full list in Appendix B.

We divide this report in 7 sections including this introduction. Section 2 presents the background, rationale and objectives of the study. Section 3 describes the conceptual framework and methodological choices made for the systematic review and the case studies. Section 4 presents the trends in the production of IE studies in our sample of countries *vis-à-vis* the rest of Latin America and the Caribbean. Section 5, in turn summarizes the findings and lessons learned from the three case studies, while section 6 presents the key features of the training performed in the region by key IE study institutions. Finally, we close with a section that summarizes the findings and draws some conclusions.

2. Background, rationale and objective of the study

The interest in Impact Evaluation (IE) has grown rapidly in Latin America. There are two regional-specific research networks working on IE: PEP-PIERI Latin America node (housed at GRADE) and LACEA´s Impact Evaluation Network (housed at CEDLAS) and a multiplicity of international initiatives that support and implement IE studies in the region such as the Abdul Latif Jameel Poverty Action Lab (J-PAL), with a one year old regional office in Latin America based at the Pontificia Universidad Católica de Chile, Innovation for Poverty Action (IPA), the Development Impact Evaluation (DIME) initiative from the World Bank and the International Initiative for Impact Evaluation (3ie), OVE (Evaluation Office at the IADB) among others.

However, rigorous impact evaluations are still very much concentrated in a few countries (Mexico, Colombia, Peru, Chile, Brazil, and Argentina to a certain extent). There are sub-regions for which IE studies are relatively scant, such as the Andes, Central America and the Caribbean. So, one rationale for taking a closer look at countries where IE is less common is to evaluate if these countries are experiencing the same trends as the rest of LAC countries, but only delayed, possibly leap-frogging to the IE state-of-the-art technology and learning from the lessons of LAC's leading IE actors, and if they are experiencing specific challenges that may reflect the initial delay in promoting evidence-based policy-making.

Thematic or sectorial coverage is also concentrated in a few areas, namely social protection, and especially conditional cash transfers programs. In addition, some donors are prioritizing Impact Evaluation in their agendas and project approval processes. While this means that knowledge will soon be available for evidence-based policy-making, there is a risk that the studies will primarily be in concordance with the donors' agenda, especially in the countries that depend more on donors for funding, designing or implementing development programs. Thus, our second rationale for focusing on countries where IE is less common is to highlight the source of the demand for IE.

When donors drive the demand for impact evaluation, not only thematic or sectorial coverage are more likely to reflect their interest, but researchers from the region may also play a more marginal role in facilitating data access and field work. In general, researchers from most countries in the region (with exceptions from some of the ones mentioned above) are underrepresented in the design and development of the IE studies. While this opens the opportunity for local researchers to collaborate with northern-based experts, developing southern researchers' capacities to conduct IE studies remains a critical challenge, especially in the Andes, Central America and the Caribbean.

Lower participation of researchers from the region may, in turn, undermine the capacity of the country to institutionalize the use of IEs for policy-making. In contrast, an emerging growing technical specialization in governmental spheres in many countries of the region may offer new windows of opportunity for research to feed into the policymaking process and a growing appetite for IE studies from the policymaking world. In theory, IE studies offer a channel to engage in policy debate with program managers, policy makers and other relevant stakeholders. Can IE studies live up to these expectations and provide a basis for healthier policy debate and reform, especially in the countries with lower local IE capacities? It is true that the production of IE studies cannot stand alone, but a favorable institutional framework may be required for improved accountability and for policy making to be guided by the conclusions of these studies (Briceño and Gaarder, 2010). Also, such accountability requires complementary information from ex-ante evaluations, process evaluations, monitoring systems, qualitative information, etc. Still, the focus on the production and use of impact evaluations for policy making could provide important insights in the process towards the institutionalization of the use of IEs for policy making.

The objectives of this study are:

- To conduct a diagnostic of IE research (what is being done in and with Impact Evaluation research), in Latin American and Caribbean countries with weaker local research capacities. We focus our analysis on a selected sample of countries, including in Latin American (Bolivia, Paraguay, Peru, Ecuador, Nicaragua, Honduras, El Salvador, Guatemala) and Caribbean countries (Dominican Republic, Jamaica, Trinidad and Tobago, Guyana).
- Compiling and reviewing carried-out IEs as well as analyzing the experience of feeding the IE results into the policymaking process.

We are particularly interested in the following research questions:

1. How many studies have been done and on what topics or policy issues and with what methodologies (RCT, q-experimental, IV, etc.)? What are the emerging fields or sectors of interest being evaluated?
2. Who is doing it? The role of local universities and research centers *vis-à-vis* the work by northern-based researchers, institutions and initiatives such as J-PAL, IPA, as well as the donors and the multilaterals, etc.
3. Who is funding it (3IE, Gates, MCC, the multilateral banks)? How have their project approval processes been modified to give primary importance to impact evaluation designs?
4. How are they being used to shape policy? What are the institutional arrangements that can better insert impact evaluations into policy design?

3. Conceptual framework

Both experimental and non-experimental approaches to impact evaluation can produce reliable estimates of the impact of a program, or fail to do so (see Appendix A for a critical review of the various IE strategies). Experiments face many challenges—at the design stage, at the implementation stage—because they rely on the goodwill of donors (this is quite an expensive endeavour), and on the support of local politicians and the monitoring of activities in the field by

program managers and sector specialists. Non-experimental methods require assumptions, and more importantly, support for assumptions for which we usually do not have a statistical test to rely upon. A careful inspection of the leading scientific journals shows that what matters is the quality of the finished work: both experimental and non-experimental studies get published in these reviews. However difficult to obtain, scientific rigor is merely a first step towards policy influence. For this study, we also rely on a case studies approach in order to analyze how impact evaluation studies is absorbed into policy-making and program management, and how, these, in turn, may influence the choices of methods for impact evaluation studies.[3]

In the next two sections, we present the methodological frameworks for the systematic review and the case studies.

Methodology

We focus our review on sub-regions in LAC where impact evaluation studies are relatively scarce. It therefore includes a number of countries in the Andes (Bolivia, Peru, Paraguay), Central America (Ecuador, Nicaragua, Honduras, El Salvador, Guatemala) and the Caribbean (Dominican Republic, Jamaica, Trinidad and Tobago, Guyana).

We review all impact evaluation studies, starting from 1995, and including on-going work. We build on the review by Bouillon and Tejerina (2007) that collected impact evaluation work up until 2007. The review clearly shows the scarcity of work in this area for the sub-regions of interest before 1995. In this sense, our choice of time span allows us to obtain a nearly complete review of all impact evaluation work in the selected countries. As we will show in the results section, many new impact evaluations are produced now, so we take special care in documenting current evaluation efforts, including on-going and uncompleted work.

In the quantitative analysis, our aim is to document the supply (who conducts the studies, on which themes, based on which methodologies) and demand (who finances IE and who funds each type of methodology) in the selected countries. In order to build on the previous systematic review for the region (Bouillon and Tejerina, 2007), we collect the following information to be used as classification criteria at the analysis stage: *country* in which the program takes place,[4] *year* of publication of the impact evaluation results, *thematic focus*, *name of the program/project* evaluated, *type of data* used in the impact evaluation study (general survey/evaluation survey/project or administrative data, whether a *baseline* is available), and the type of *methodology* for identification of the impact.

In terms of thematic and sectorial focus,[5] we review impact evaluations of interventions in the following areas: active labour market (ALM);

- Agriculture and rural development, which includes: Agriculture (AGR), Environment (ENV), Transport and communication (TC)
- Education (EDU)

[3] We explain in section D the methodology for case studies.
[4] Each study is assigned a code. A study code starts with the first three letters of the country name.
[5] From here on, we use the three letters for country name and an abbreviation for thematic focus to identify the studies we review.

- Entrepreneurship, which includes Microfinance (MIC) and SMEs (SME)
- Local Governance, which includes Governance (GOV), Social Investment Fund (SIF)
- Other human capital, which includes: Early child development (ECD), Health (HEA), Nutrition (NUT)
- Social protection (SP)
- Urban development, which includes: Public services (PS), Housing (HOU)
- Others, which includes Crime (CR)

As explained in the previous section, we consider impact evaluations that provide a rigorous framework for identifying impact. This includes two broad categories: randomized experiments and non-experimental evaluation (instrumental variables, difference-in-difference and other longitudinal methods, matching, regression discontinuity, and structural estimation).

In addition to the criteria selected from the review by Bouillon and Tejerina (2007), we collect information to help us address key research questions/ investigate the hypotheses outlined in section A. In addition to thematic focus, we collect data on the intervention target group (*e.g.*, women, youth, children, teenagers, entrepreneurs, farmers, the poor, unemployed), the intervention target group size (total number of beneficiaries), the budget for the intervention, the main outcomes of interest (*e.g.*, on education: enrollment, attendance, test scores) and findings. We gather data on whether the intervention is funded by the government/a multilateral agency or an NGO. We also document the identity of the principal investigator for the IE research: name of the author(s) of the publication; whether the evaluation is done in-house or by an independent institution; the identity of the employer/donor or granting agency (research grant/university-funded *vs.* research contract from implementing agency *vs.* research contract from other sources). We determine whether the IE study is completed or still on-going, and if completed, whether it is published as a peer-reviewed article/book or a non peer-reviewed document (working paper, report). Finally, we document if local researchers are involved in the IE study, and if so, the type of involvement (in data collection only/ at the research and analysis stage).

As discussed previously, we include all IE studies reviewed in Bouillon and Tejerina (2007) conducted in our sample of countries. They based their analysis on existing systematic reviews in specific areas (e.g., Rawlings and Rubio 2003 on CCTs), as well as on the available databases of IE studies. We also search these databases and more recent ones for studies published since 2007. We base our search on the most common databases for academic papers (IDEAS/RePEC,[6] EconLit[7] and JSTOR[8], and SSRN Randomized Social Experiments[9]). We also identify the main organization funding IE research and searched their databases. They include databases from World Bank Development Impact Evaluation Initiative (DIME),[10] the International Food Policy Research Center,[11] Innovation for Poverty Action (IPA),[12] from the Abdul Latif Jameel Poverty Action Lab

[6] http://ideas.repec.org/search.html
[7] http://www.aeaweb.org/econlit/index.php
[8] www.jstor.org
[9] http://www.ssrn.com/link/Randomized-Social-Experiments.html
[10] http://web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDEVIMPEVAINI/0,,contentMDK:21553788~pagePK:64168445~piPK:64168309~theSitePK:3998212,00.html
[11] www.ifpri.org
[12] http://www.poverty-action.org/project-evaluations/search

(J-PAL),[13] the International Initiative for Impact Evaluation (3ie),[14] and the Bill and Melinda Gates Foundation.[15] In order to complement these searches, we look for the most recent and on-going IEs based on past LACEA and IEN programs, the World Bank Impact Evaluation webpage,[16] the Inter-American Bank of Development Operational Office of Evaluation and Oversight and Development Effectiveness program webpages.[17]

We make use of as well as augment the database of Bouillon and Tejerina (2007) gathered for all LAC countries except the ones we focus on here. This database is used as a benchmark for comparison.

Methodology for the case studies

We develop three case studies in selected LAC countries (Dominican Republic, Peru and El Salvador). We want to learn if good practices impact evaluations are feasible and can help improve policy making, as well as cases in which good impact evaluations have not been able to reach the policy sphere. In addition, the case studies are a key input for a discussion of the institutional arrangements in countries for their demand, implementation and use of IE studies.

One first step is to define whom to interview in each of the countries. We start with a sample of researchers that run some of the evaluations to learn about their experiences from the process, how their interaction with the implementing agency was and how they succeeded or not in disseminating their results. Next, we define a sample of public officials in charge of implementing government programs, with and without IEs. In the process, we also identify if there is a specific public office that is in charge of supervising the quality of social expenditures and programs, and if they demand and use IEs.

The interviews are conducted using the methodology outlined below. We structure the analysis around three different, but closely related, questions. The first question relates to the existence of a context for institutionalizing IE, since this is an important step in making IE a useful tool for policy makers. We follow closely the material developed by Briceño and Gaarder (2010) in order to assess the extent to which the many facilitating factors for institutionalizing IE are present or not. Here, it is crucial to find out whether the following elements are present in the three selected countries:

1. Is there an agency with a mandate to conduct/commission impact evaluation of different government programs? If such an agency exists,
a. Which is its level of independence?
b. How is it financed?
c. How long ago was it created?
d. Does it have policy influence? Is there a systematic process of checks and balances in order to feed evaluation results into program innovations/expansions?

[13] http://www.povertyactionlab.org/search/apachesolr_search?filters=type:evaluation
[14] http://www.3ieimpact.org/database_of_impact_evaluations.html
[15] http://www.gatesfoundation.org/grants/Pages/search.aspx
[16] http://go.worldbank.org/169GZ6W820
[17] Respectively, http://www.iadb.org/en/about-us/departments/about,1342.html?dept_id=OVE and http://www.iadb.org/en/topics/development-effectiveness/development-effectiveness,1222.html

2. How important is the presence of foreign donors and how does their demand for sound IEs spur its institutionalization?

3. Which are the main obstacles for institutionalizing IE?

We then address the question of whether the different factors which facilitate such institutionalization are present or not. Following Briceño and Gaarder's "wish list", we look for the existence of a democratic system with vibrant opposition, as well as the existence of influential sound previously carried-out IEs to lead the process (for example: the initial evaluation of PROGRESA In México and the posterior inception of CONEVAL). Additionally, we search for the presence of a powerful stakeholder – Congress, Ministries, Presidency- which may facilitate the triggering of the institutionalization process. Finally, we also assess the presence of technical assistance in the country to conduct or commission IE.

The second category is related to a number of IEs which provide information about specific programs and their contexts, which may or may not facilitate such studies. While most LAC countries lack an institutionalized agency for IE, there are IE studies conducted in response to different demands (mainly from external donors). In the context of specific IEs we can also establish specific issues affecting their very existence. Moreno, Campuzano and Levy (2009) point to barriers and facilitators for conducting rigorous IE. We analyze both barriers and facilitators for existing and planned IE in each of the analyzed countries. We also examine the quality (in terms of being methodologically sound) of specific conducted IE.

We pay particular attention to the following barriers:
1. Lack of support for rigorous IE.
2. How unrealistic plans for program implementation may endanger the evaluation design?
3. Are pilot case experiences (with or without IE) used for learning before the implementation of a program?
4. Is there good secondary data available? Can this be useful with the purposes of the evaluation (for eligibility criteria, eligible population power calculations, etc)?

We are particularly interested in the following facilitators:
1. Degree of involvement of government officials implementing the projects to be evaluated.
2. Capacity for independent data collection.
3. Existence of a legal framework for conducting the evaluation (for example, some IDB/WB loan conditions the approval of such loan to the design of an IE and set up a specific budget for that).

In terms of the quality of IE's available:
1. Are they sound?
2. Which programs were evaluated? Was there a significant change in the methodology once the evaluation was designed?
3. Who demanded the evaluations?
4. Who funded them?

We also seek to understand if the existence of isolated IEs may pave the way for institutionalizing IE in the future.

Finally, we investigate whether IE have influenced policy-making (both institutionalized or not). Here we base our case study methodology in Lindquist (2001) and Weirauch and Díaz Langeau (2011). We investigate this aspect by asking the following questions:

1. Have IE improved the knowledge of the actors involved?
2. Have IE modified existing programs or policies or caused fundamental re-design of programs or policies?
3. Have IE helped develop technical capacities at the local level (either within government bodies, think tanks, universities, etc.) in order to promote future IE?
4. Have IE results provided learning/networking opportunities for sharing the knowledge, internally or with colleagues elsewhere?
5. Have IE introduced new concepts for framing debates, putting ideas on the agenda, or stimulating debate?

The evaluation of all these dimensions provides us with a basis to complete a diagnosis about the main strengths and weaknesses of IE as a tool for policy makers in the selected countries.

4. Quantitative analysis

Before presenting descriptive statistics on the IE studies database that we collected, we alert the reader on three points.

First, in this work, we distinguish between the attributes of the IE studies under review from the interventions that these studies are assessing. For instance, CCTs typically generate more than one IE study. Some of the results we present are best framed in terms of studies produced (type of methodology pursued, whether it led to a publication, type of funding for the research, etc). Others are best framed in terms of the intervention being evaluated (e.g., whether it is government-run, who is funding the program, etc.).

Secondly, we acknowledge the limitations due to missing observations on some of the collected variables of interest. Table 1 reports the number of studies/projects with non-missing information for each of the variables of interest. Most of the missing information concerns the involvement of local researchers and the funding for the research (about a third of the cases are missing for these variables).

Table 1: Impact evaluation studies (# completed and non-missing cases).

	Our Study Area	Rest of LAC
Studies		
Total # of studies	129	188
# of completed studies	91	153
with info on local researchers involvement	87	136
with info on identification method	89	149
with info on publication status	90	143
with info on who is conducting the research	90	136
with info on who is funding the research	87	126
Projects		
# of projects evaluated	102	138
with info on who is funding the project	94	105
with info on who is implementing the project	93	125

A final and related point is that we need to distinguish between on-going and completed studies. In our study period (16 years, from 1995-2011), we find 244 completed studies and also consider 73 studies that are currently on-going (Table 1). Many of the missing cases that are documented in Table 1 are due to the fact that some of the studies we review are still on-going. Note that there are relatively more on-going studies in the study area (30%) than in the rest of LAC (19%). In the first part of this section, we provide a description of the geographical coverage, time trends and type of research produced. We then describe the attributes of the assessed programs.

Our review concerns 317 IE studies in 21 countries. We distinguish between two groups of countries (Table 2):

Table 2: Geographical coverage.

Our Study Area			Rest of LAC		
Country	**# of studies**	**%**	**Country**	**# of studies**	**%**
The Andes					
Peru	37	28.7	Mexico	61	32.5
Bolivia	15	11.6	Colombia	38	20.2
Ecuador	13	10.1	Chile	26	13.8
Paraguay	2	1.6	Brazil	25	13.3
Central America			Argentina	18	9.6
Nicaragua	15	11.6	Uruguay	9	4.8
El Salvador	13	10.1	Costa Rica	5	2.7
Honduras	11	8.5	Panama	4	2.1
Guatemala	5	3.9	Haiti	1	0.5
Caribbean Countries			St. Lucia	1	0.5
Dom. Republic	10	7.8			
Jamaica	7	5.4			
Trinidad & Tobago	1	0.8			
Total	129	100.0	**Total**	188	100.0

- Those located in our study area, which includes 12 countries, including Guyana,[18]
- The rest of LAC, based on Bouillon and Tejerina's review study, and including 10 countries.

Our study area counts about 12.9 IE studies on average per country, compared to 18.8 for the rest of LAC. Clearly, the countries are heterogeneous on a number of dimensions and the two groups differ. For example, the fact that Brazil (situated in the rest of LAC) produces a large number of IE studies can be simply explained by the stock of programs available for evaluation in the country.

More interestingly, we find that in our study area, Peru accounts for 29% of all IE studies. A similar pattern is found in the rest of LAC: Mexico produced about 32% of all studies for the region.

[18] No IE study is found for Guyana.

The countries in our area of focus are composed of two main subgroups:

- Those with less than 10 IE studies (all Caribbean countries + Paraguay), and
- Those with 11-15 IE studies (all Central American countries + Bolivia).

In this sense, Peru clearly stands out with 38 studies.

As a comparison, the countries in the rest of LAC can also be grouped into two categories:

- Those with less than 10 IE studies (Caribbean, Panama, Costa Rica and Uruguay), and
- Those with 18-38 studies (Brazil, Chile, Colombia and Argentina).

In the same time frame, Mexico produced 61 studies.

Then, both Peru and Mexico stand out for each group of countries. Yet, the gap between Peru and the next subgroup of countries is larger than the gap between Mexico and the second subgroup in the rest of LAC.

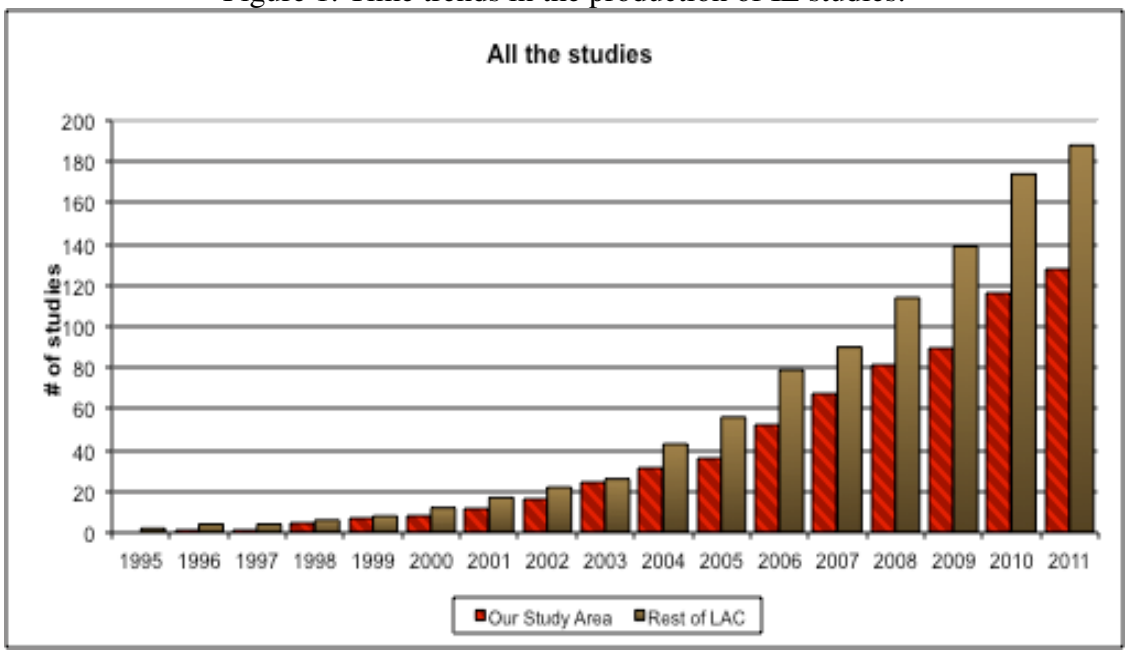
Going back to our area of focus, we find that the categories described above correspond to geographic clusters:

- Caribbean countries are globally falling behind when it comes to doing impact evaluation.
- Central American countries are at a median position.
- The two closest Andean countries (Peru and Bolivia) constitute a leading group with more than 40% of the total production of IE studies in this group of countries.

This geographical clustering is consistent with two hypotheses: local knowledge spillovers (learning from others), and/or the effect of a common factor (e.g., specific geographic interest in IE by external funders). The geographic clustering that we observe in the study area is not observed in the rest of LAC. The subcategories described above seem to be related to the size of the countries and their level of income.

Most IE studies have been produced starting in 2006 (see Figure 1). This trend is the same in the two groups of countries (71% in the area of focus and 70% in the rest of LAC). However, there is more heterogeneity among the countries in the rest of LAC, with Mexico and Colombia having a steady stream of IE produced since 2000.

Figure 1: Time trends in the production of IE studies.



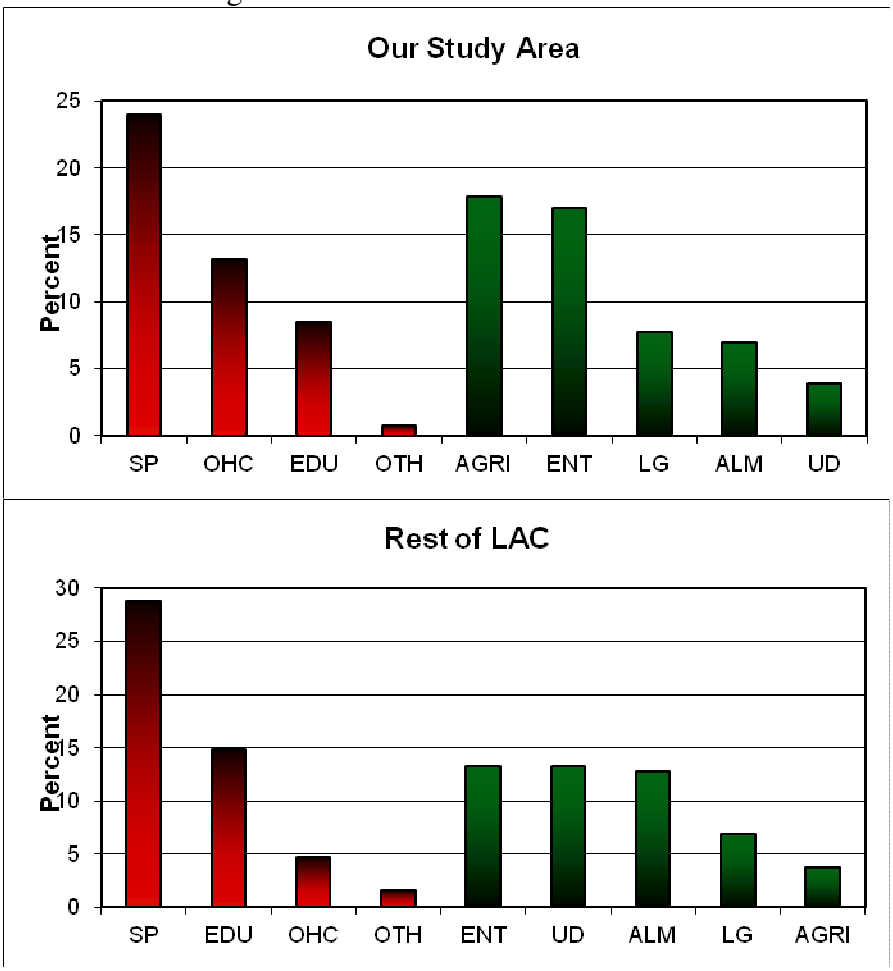
Given the time frame for IE studies, it is likely that we are observing the first wave of IE studies in the area of focus. Whether the interest in IE studies will remain (as in Mexico and Colombia) is left to see.

Programs from two types of policies are typically evaluated: social development policies and growth investment policies (see Table 3 and Figure 2).

Table 3: Distribution of IE studies across themes (as % of all completed studies)

		Our Study Area (%)			Rest of LAC (%)
	Social Development			**Social Development**	
SP	Social Protection	24.0	SP	Social Protection	28.7
OHC	Other human capital	13.2	EDU	Education	14.9
EDU	Education	8.5	OHC	Other human capital	4.8
OTH	Other (Crime)	0.8	OTH	Other (Crime)	1.6
	Growth Investments			**Growth Investments**	
AGRI	AGRI & Rural Dev.	17.8	ENT	Entrepreneurship	13.3
ENT	Entrepreneurship	17.1	UD	Urban development	13.3
LG	Local Governance	7.8	ALM	Active labor market	12.8
ALM	Active labor market	7.0	LG	Local Governance	6.9
UD	Urban development	3.9	AGRI	AGRI & Rural Dev.	3.7
	Total	100		**Total**	100

Figure 2: Thematic focus in IE studies



In both areas, social protection programs are the most frequently assessed. They correspond to 24% (respectively 28%) of all IE studies in the area of focus (and the rest of LAC).

The two areas then differ in terms of the type of programs evaluated. In the study area, the focus of evaluation is on agriculture and rural development programs, followed by programs facilitating entrepreneurship and those that help improve health, nutrition and early childhood development. In the rest of LAC, the focus is on evaluating education reforms and programs and active labour market strategies, urban development projects and entrepreneurship programs.

These priorities likely reflect the types of policies that are in place in the two areas. The countries with weaker research capacities are also countries where health, nutrition and ECD are still lacking, agriculture and the rural sector represents a sizeable part of the economy and formal labour markets are not yet well developed.[19]

Table 4 presents the distribution of studies across topics. We find that education is actually the second most studied topic within social development policies in the two areas, before health,

[19] Yet, it is also important to keep in mind that this breakdown by sector or program only pertains to those interventions that are actually assessed. In order to provide a fuller picture, one would need to compile information on all the programs in these sectors for all the countries, but this is beyond the scope of this work.

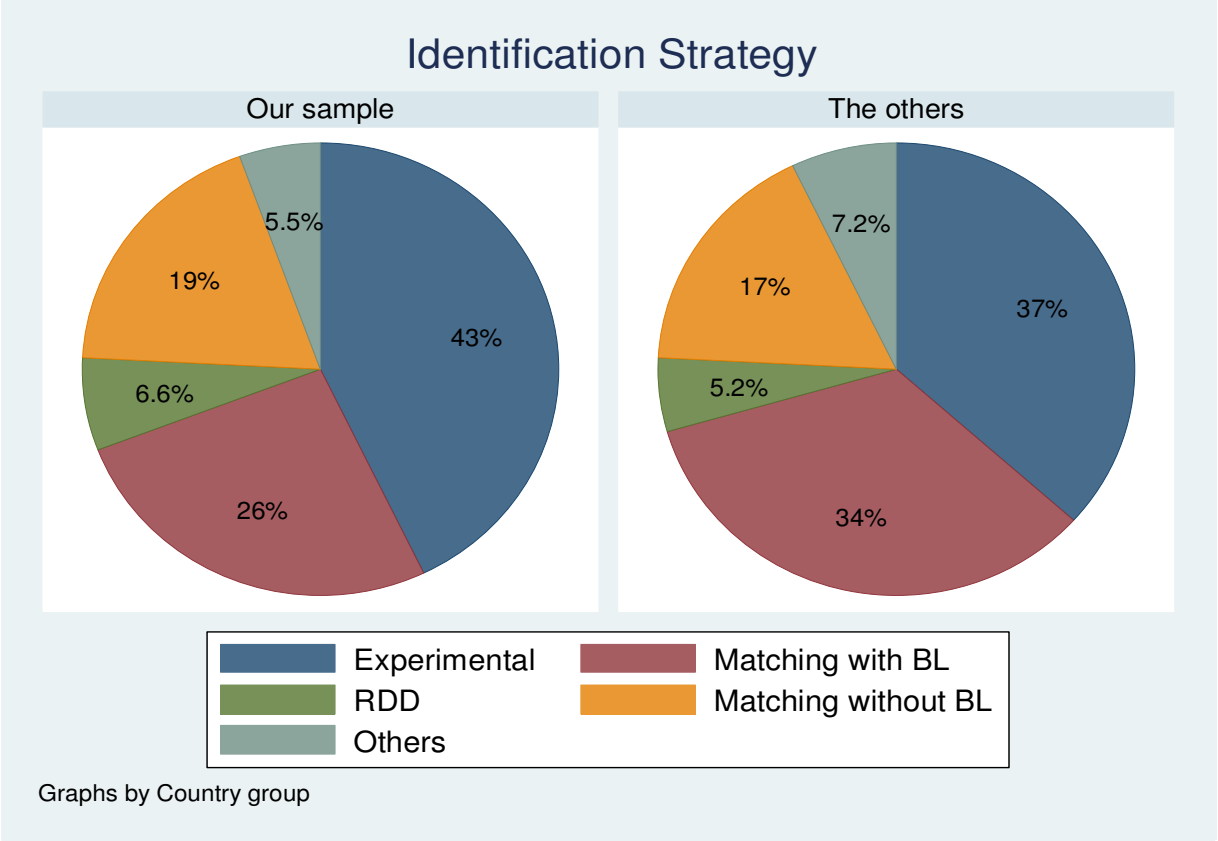
nutrition and ECD. Microfinance topics generate most of the IE studies focusing on entrepreneurship in the two areas. Social investment funds and transport & communication are specific to the study area.

Table 4: Topics of interest in IE studies.

Our Study Area (%)		Rest of LAC (%)	
Social Development		**Social Development**	
Social Protection	24.0	Social Protection	30.3
Education	8.5	Education	14.9
Nutrition	5.4	Health	3.7
Early child development	3.9	Crime	1.6
Health	3.9	Early child development	0.5
Crime	0.8	Nutrition	0.5
Growth Investments		**Growth Investments**	
Microfinance	13.2	Active labor market	12.8
Agriculture	7.8	Microfinance	9.0
Active labor market	7.0	Urban development	8.0
Transport & Communication	6.2	Governance	5.9
Public Services	4.7	SMEs	4.3
Social Investment Fund	4.7	Public Services	3.2
SMEs	3.9	Agriculture	2.7
Governance	3.1	Environment	2.7
Environment	1.6		
Urban development	1.6		
Total	100	**Total**	100

Most of the IE studies in the area of focus are based on an experimental design. This is also true in the rest of LAC (Figure 3). Experimental IE actually represent a larger share of all completed IE work in the study area than in the rest of LAC.

Figure 3: Identification strategies for IE studies.



Other identification strategies are also used at the same relative intensity in each area. For those studies that are not based on an experimental design, the most popular approach is matching in a difference-in-difference framework, followed by simple matching, regression discontinuity designs and other non-experimental. The composition is similar in the two groups of countries.

Table 5 indicates that the number of IE studies based on an experimental design has tripled between 2005 and 2006 and continues to grow. This trend has actually taken off more vigorously in the study area (where the number of experimental studies increased ten-fold between 2005 and 2006) than in the rest of LAC. However, this growth seems to be more stable in the rest of LAC than in the study area.

Table 5: Time trends in the choice of identification strategy.

	Our Study Area (%)					Rest of LAC (%)				
Year	RCT	Match w/ BL	RDD	Match w/out BL	Others	RCT	Match w/ BL	RDD	Match w/out BL	Others
						0.0	1.9	0.0	0.0	0.0
1996	2.6	0.0	0.0	0.0	0.0	0.0	3.9	0.0	0.0	0.0
1998	7.9	0.0	0.0	0.0	20.0	1.9	1.9	0.0	0.0	0.0
1999	2.6	4.4	0.0	0.0	0.0	0.0	1.9	0.0	0.0	9.1
2000	0.0	0.0	0.0	5.9	0.0	3.7	3.9	0.0	0.0	0.0
2001	2.6	8.7	0.0	5.9	0.0	5.6	1.9	0.0	4.0	0.0
2002	0.0	4.4	0.0	17.7	0.0	1.9	1.9	0.0	4.0	0.0
2003	2.6	26.1	0.0	0.0	40.0	3.7	1.9	14.3	0.0	0.0
2004	2.6	8.7	0.0	5.9	0.0	3.7	17.3	14.3	12.0	18.2
2005	2.6	4.4	0.0	5.9	20.0	7.4	7.7	0.0	4.0	27.3
2006	10.5	8.7	16.7	5.9	0.0	5.6	17.3	0.0	24.0	18.2
2007	13.2	0.0	0.0	11.8	20.0	5.6	3.9	14.3	8.0	0.0
2008	15.8	21.7	16.7	0.0	0.0	14.8	9.6	28.6	8.0	18.2
2009	10.5	4.4	16.7	11.8	0.0	13.0	13.5	0.0	8.0	9.1
2010	18.4	4.4	0.0	29.4	0.0	18.5	7.7	28.6	20.0	0.0
2011	7.9	4.4	50.0	0.0	0.0	14.8	3.9	0.0	8.0	0.0
Total	100	100	100	100	100	100	100	100	100	100

According to Table 6, publication in academic journals is lower for IE studies from the area of study (12%) than for those from the rest of LAC (16.5%). This is not due to the fact that there are more on-going studies in the area of focus than in the rest of LAC (the proportions are very similar). This is either related to lower motivation to publish the results or facing higher difficulty in meeting publishing requirements (credibility of the results). But the number of published works is too low to pursue the analysis further (a total of 35 studies are published).

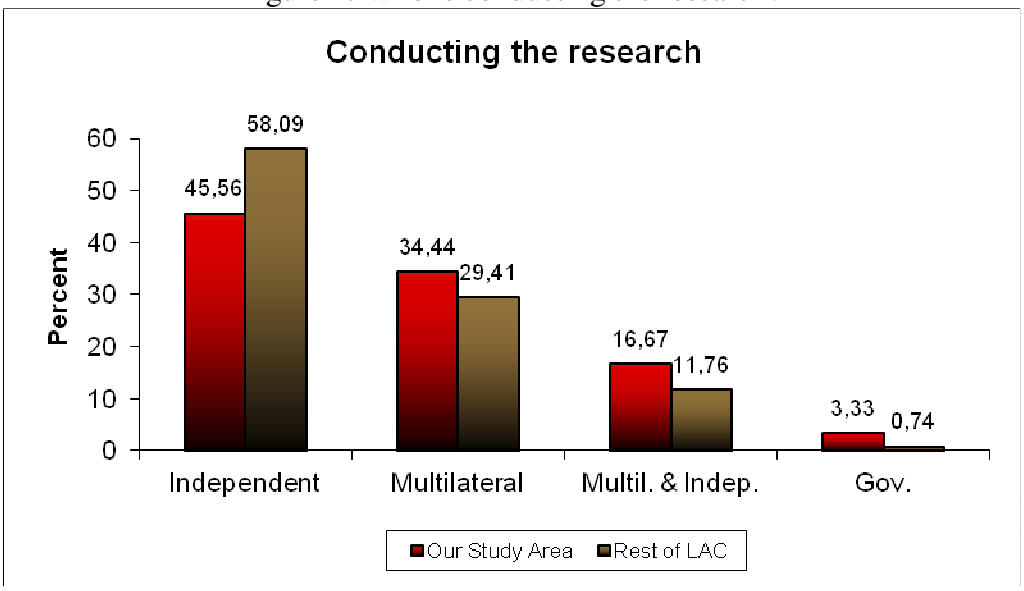
Table 6: Time trends in publication.

Year	Our Study Area (%)				Rest of LAC (%)			
	Published artIcle	W.P.	D.P.	Report	Published artIcle	W.P.	D.P.	Report
1995	0.0	2.4	0.0	0.0	4.2	0.0	0.0	0.0
1996	0.0	2.4	0.0	0.0	0.0	2.4	0.0	0.0
1998	18.2	4.9	0.0	0.0	0.0	1.2	0.0	0.0
1999	0.0	4.9	0.0	0.0	4.2	0.0	0.0	0.0
2000	0.0	0.0	0.0	4.8	4.2	0.0	0.0	8.7
2001	18.2	0.0	0.0	9.5	0.0	3.5	9.1	4.4
2002	9.1	4.9	5.9	0.0	4.2	2.4	0.0	4.4
2003	9.1	7.3	0.0	19.1	4.2	3.5	0.0	0.0
2004	9.1	2.4	0.0	14.3	4.2	12.9	0.0	13.0
2005	0.0	7.3	0.0	4.8	12.5	3.5	0.0	21.7
2006	9.1	9.8	11.8	9.5	8.3	14.1	9.1	17.4
2007	0.0	12.2	17.7	0.0	4.2	5.9	0.0	4.4
2008	18.2	12.2	17.7	9.5	12.5	15.3	9.1	8.7
2009	0.0	9.8	11.8	9.5	12.5	14.1	18.2	0.0
2010	9.1	19.5	11.8	9.5	20.8	11.8	45.5	8.7
2011	0.0	2.4	23.5	9.5	4.2	9.4	9.1	8.7
Total	100	100	100	100	100	100	100	100

We can nevertheless look at the evolution through time in the number of publications. Most of the studies were published beginning in 2006 in the rest of LAC, and later (2008) in the area of focus.

The majority of IE research is typically conducted by independent researchers and organizations, followed by multilaterals, a mix of both multilateral and independent researchers/organizations, and government agencies. The pattern is globally similar in both groups of countries (Figure 4).

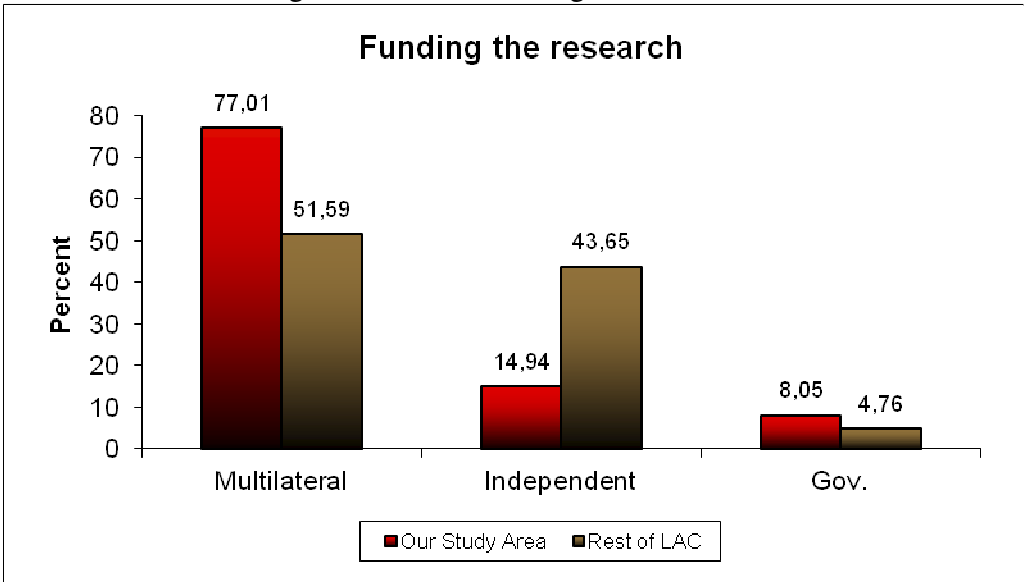
Figure 4: Who is conducting the research?



However, there are some notable differences: independent research constitutes a smaller fraction of completed IE research in the study area than in the rest of LAC, while research led by multilateral agencies is relatively more prevalent in the study area than in the rest of LAC.

Furthermore, the pattern for funding is even more clearly differentiated in the two groups of countries (Figure 5). In the study area, 77% of completed IE studies were funded by multilateral agencies. In the rest of LAC, funding is balanced between multilaterals and independent research.

Figure 5: Who is funding the research?



Independent organizations conduct and fund the most scientifically rigorous studies. (Figure 6). When multilateral agencies conduct or fund these studies jointly with independent organizations, a higher scientific rigor is also more likely to be obtained than otherwise.

Figure 6: Choice of identification strategy for the impact assessment depends on who runs/funds the IE study

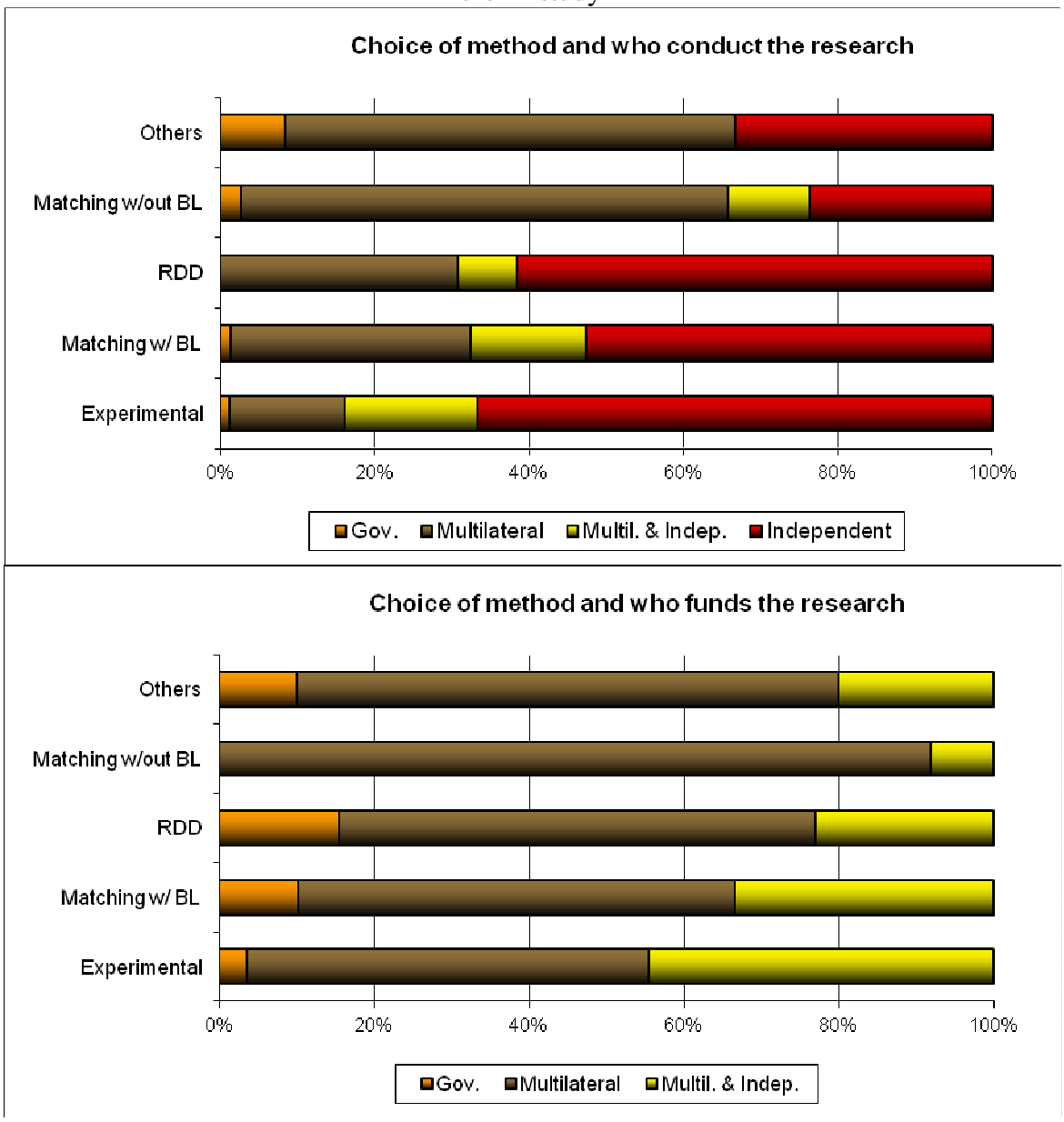
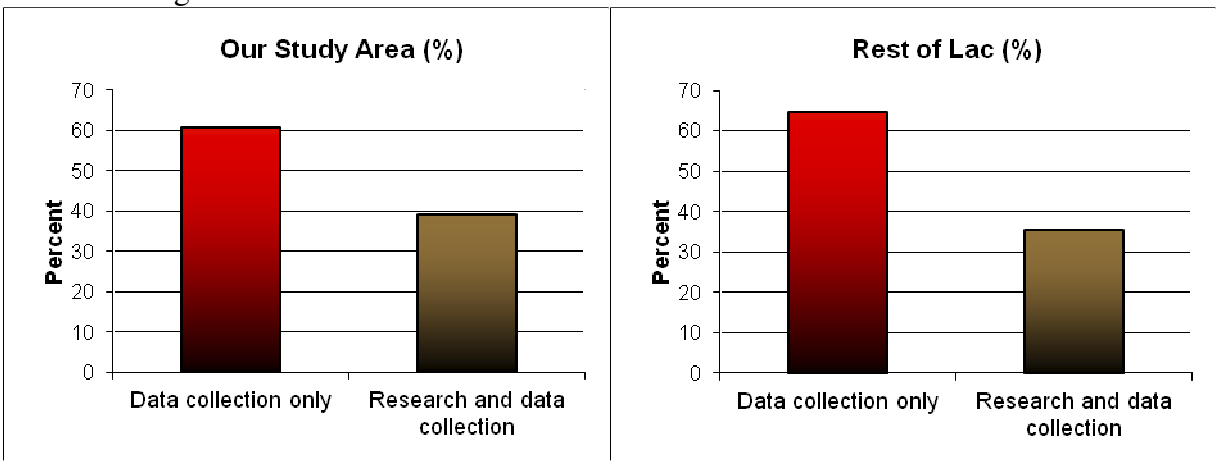


Figure 7 looks at whether local researchers are involved in research and/or data collection. We find that most IE studies involve local researchers in the data collection. Yet, very few of them also involve them in other types of research work. This pattern is similar for the group of countries in our study area and for those in the rest of LAC (resp. 39% and 35%).

Figure 7: Local researchers involvement in research and data collection.



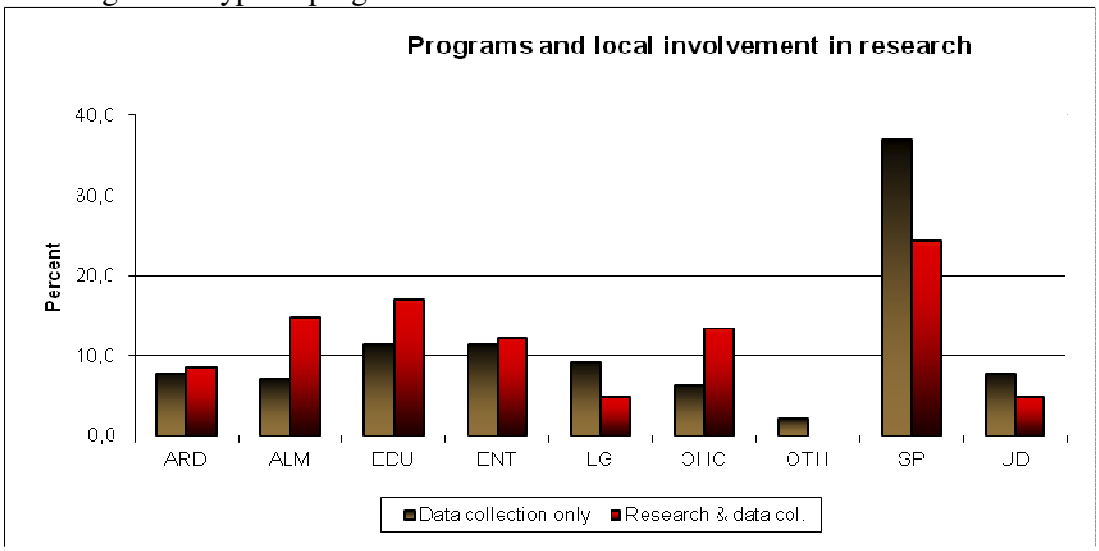
When looking at the evolution through time (Table 7), we find that involving local researchers in other research work started in 2006 and has been maintained since.

Table 7: Time trends in the involvement of local researchers in IE studies.

	Our Study Area (%)		**Rest of LAC (%)**	
Year	Data collec.only	Involvement in research	Data collec.only	Involvement in research
1996	0.0	2.9	2.3	0.0
1998	5.7	2.9	1.1	0.0
1999	3.8	0.0	0.0	2.1
2000	1.9	0.0	3.4	0.0
2001	3.8	5.9	3.4	4.2
2002	1.9	5.9	1.1	6.3
2003	7.6	14.7	4.6	0.0
2004	5.7	2.9	8.0	16.7
2005	3.8	2.9	5.7	8.3
2006	11.3	5.9	10.2	16.7
2007	11.3	5.9	5.7	6.3
2008	11.3	17.7	13.6	12.5
2009	11.3	5.9	12.5	10.4
2010	9.4	23.5	17.1	12.5
2011	11.3	2.9	11.4	4.2
Total	100	100	100	100

Overall, we find that the share of completed studies with an involvement of local researchers, beyond simple data collection, depends on the domain of intervention (Figure 8). Social protection, education, active labour market programs and other human capital investment programs are domains in which local researchers are active. There are some differences between the two groups of countries that we study. In our study area, local researchers are more active in social protection, nutrition, health and ECD, entrepreneurship. In the rest of LAC, it is social protection, education, active labour market interventions and projects that facilitate entrepreneurship that draw most of the attention from local researchers, or are more likely interested including local researchers in the assessment.

Figure 8: Type of program assessed and involvement of local researchers.



Note: ARD (Agriculture & Rural Development), ALM (Active labour market), EDU (Education), ENT (Entrepreneurship), LG (Local Governance), OHC (Other human capital), OTH (Others: Crime), SP (Social Protection), UD (Urban development)

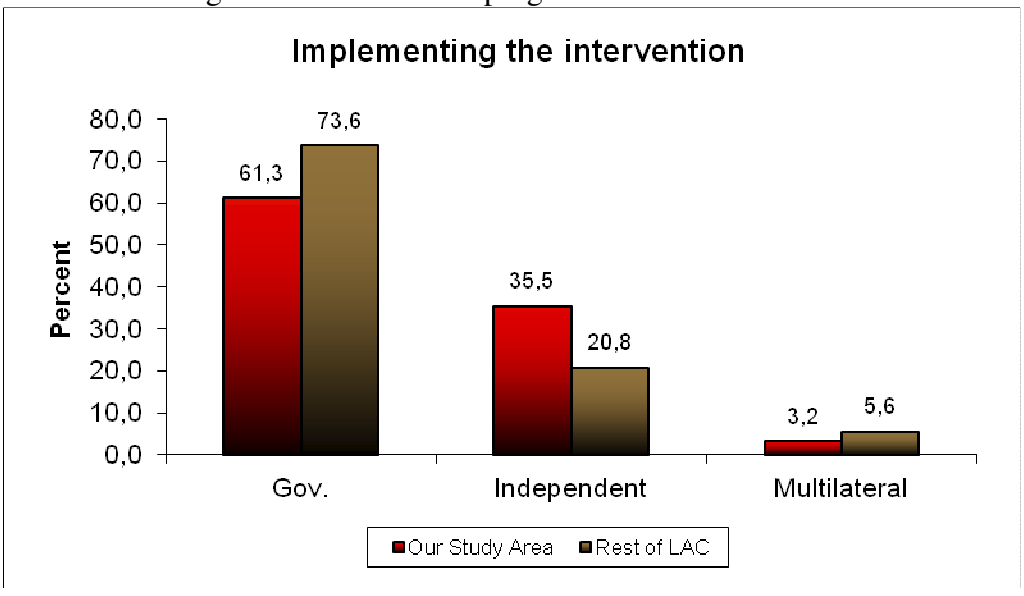
Only 18% of IE studies conducted by multilaterals succeed in involving local researchers for the analysis stage (Table 8). Independents do somewhat better and 69% of studies integrating local researchers at this stage are conducted by independent organizations. Yet, when looking at funding, we find that multilaterals fund 52% of studies with local research involvement. This two results put together are consistent with the fact that some multilaterals fund independent organizations to conduct IE studies, and these, in turn, employ local researchers for data collection and analysis.

Table 8: Local involvement in research depends on who is conducting/funding the research.

Conducting research	Gov.	Multilateral	Independent	Multil. & Indep.	Total
Data collection only	0.0	40.6	44.2	15.2	100.0
	0.0	82.4	52.6	70.0	63.3
Research and data col.	5.0	15.0	68.8	11.3	100.0
	100.0	17.7	47.4	30.0	36.7
Total	100.0	100.0	100.0	100.0	100.0
Funding research					
Data collection only	3.0	65.9	31.1	n.a.	100.0
	30.8	68.5	60.3	n.a.	63.5
Research and data col.	11.8	52.6	35.5	n.a.	100.0
	69.2	31.5	39.7	n.a.	36.5
Total	100.0	100.0	100.0	n.a.	100.0

Programs that are being evaluated are run by government, multilateral and independent bodies (Figure 9).

Figure 9: Who runs the programs under evaluation?

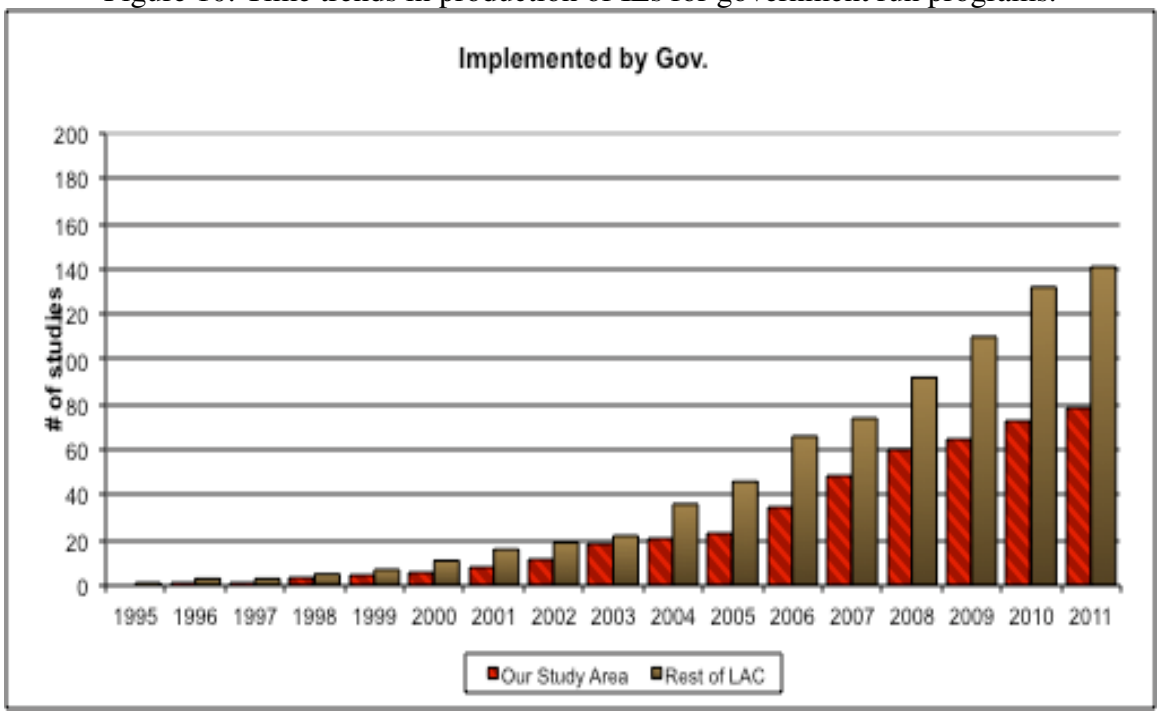


Most of them are mainly run by government agencies. They represent respectively 61% and 74% of all the programs that are evaluated in the study area and the rest of LAC. There are relatively more multilateral agency programs in the study area than in the rest of LAC. This finding is

consistent with the fact that there are relatively more IE studies funded by multilaterals in the study area than in the rest of LAC.

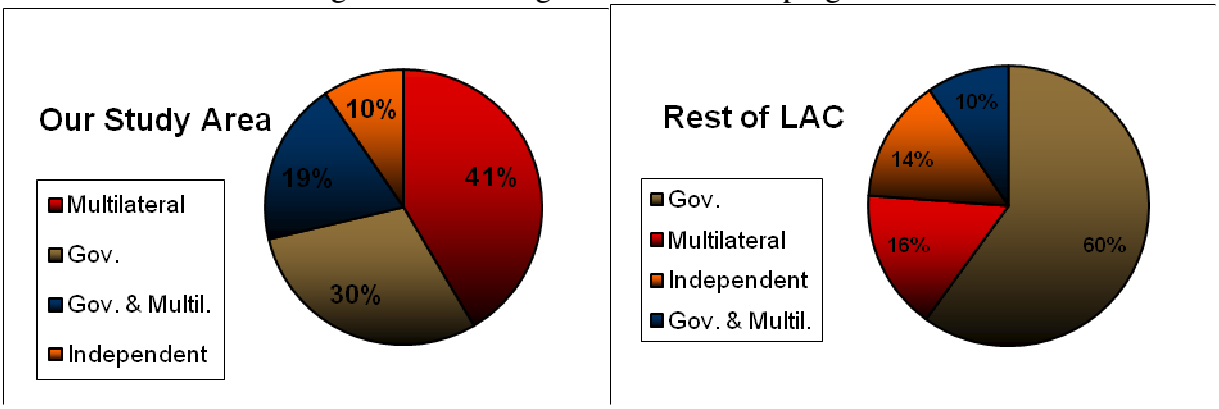
Looking at the evolution through time (Figure 10), we find that government-run interventions were always assessed, but the data suggests that the intensity at which these interventions are now evaluated is higher than in the past. Compared to Figure 1 above, we also find that the gap is much larger when we restrict the analysis to those interventions that include some participation of the public sector. That is, in our sample of countries, the role of NGOs might have been more relevant in explaining the global growth in the production of rigorous IEs.

Figure 10: Time trends in production of IEs for government run programs.



According to Figure 11, multilaterals fund 41% of the programs evaluated in the study. Governments come at a second place with 30%, followed by a combination of government and multilaterals (20%). Independent bodies fund only 10% of the evaluated programs.

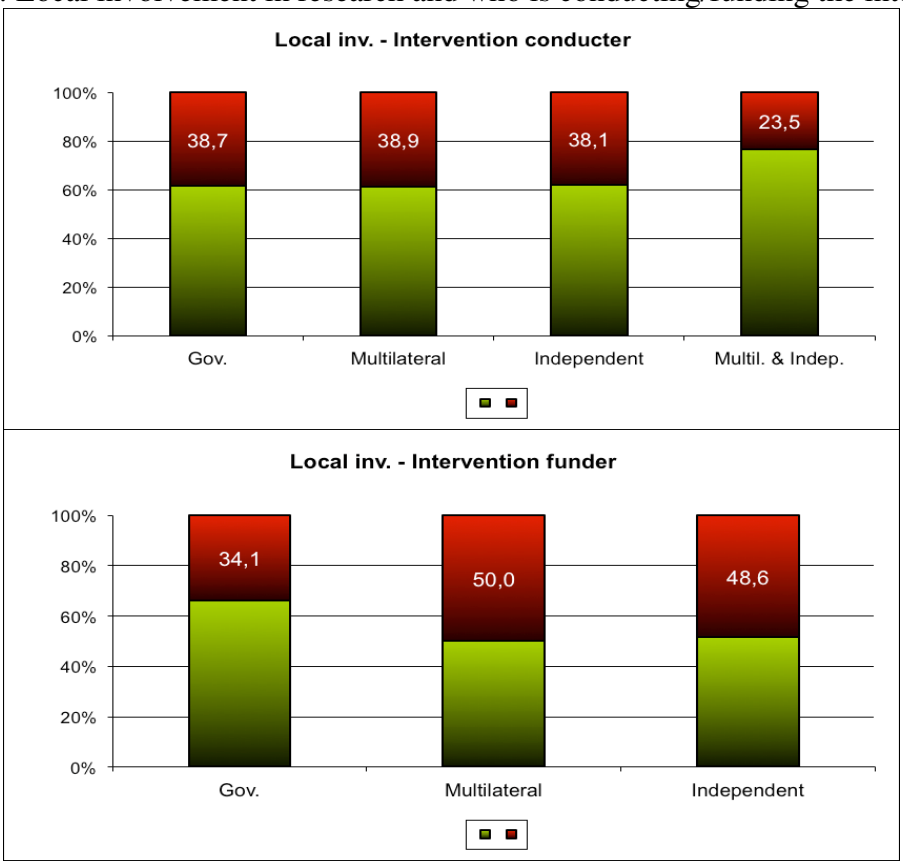
Figure 11: Funding for the evaluated programs.



In the rest of LAC, the pattern is quite different, with governments funding (60%) the majority of these programs, followed by multilaterals (16%) and independent bodies (14%), the remaining being funded by a combination of government and multilaterals.

Local researchers are about as likely to be involved in both research and data collection whether the intervention is conducted by an independent organization or a multilateral agency (Figure 12). Similarly, 62% of all completed studies assessing interventions funded by either government, multilateral or independent organizations, only involve local researchers to do the data collection effort. This is even higher when the intervention is jointly funded by government and multilateral organizations.

Figure 12: Local involvement in research and who is conducting/funding the intervention.



5. Lessons from Three Case Studies

The previous analysis provide important insights about the size and nature of the recent international wave of IE studies in our sample of countries, looking at the thematic focus, the type of methodologies used, the level of involvement of the government, the participation of local researchers, the sources of funding, etc. However, such quantitative analysis cannot tell us much about the extent to which the IE studies are being used for the design and implementation of policy making in the countries, and the factors facilitating and limiting such process. For that purpose, we now present findings from three case studies conducted in El Salvador, Peru and the Dominican Republic—three very different countries that have recently experienced some developments in their production of IE studies. As described in part 3 of section 2, we look at the list of IE studies identified in the country and try to identify key actors in the academic and policy areas, some of whom were later interviewed. In addition, we reviewed the local and international literature associated with the key processes analyzed, such as the Fomilenio in El Salvador, the Results Based Budgeting Initiative in Peru, and the Director of Juventud y Empleo in the Dominican Republic.

In this section, we present the conclusions we draw from the three case studies. A summary of the findings for each country can be found in Appendix C. We report here three features identified in the case studies that are relevant for understanding the space for policies regarding the

strengthening of the institutionalization of IEs for policy making[20]. The first one refers to the external shocks countries are facing in the production of rigorous IE studies, which is the result of increasing awareness among international donors and agencies about the appropriateness of using IEs for consolidating evidence based policy making. The second one discusses the characteristics of the embryonic local processes generated in each country in the light of the independence-relevance trade-off raised by Briceño and Gaarder (2010). The third one discusses the challenges to generate an institutionalization fundamental for effective policy design in the context of multisectoral interventions that are required for a specific social objective. We expand this discussion in the remaining of this section.

The three case studies show different examples of external shocks that have significantly altered the production of rigorous impact evaluations. In the Peruvian case, we have the Innovations for Poverty Action (IPA), which is a non-profit organization that uses and promotes the use of randomized control trials to identify what works best for helping the world's poor.[21] Fomilenio, a public office in charge of coordinating efforts against poverty under the MCC-GOES compact in El Salvador, has a clear mandate to help partner governments fund a well-defined poverty strategy, while establishing a learning system based on the most rigorous identification strategies, preferably RCTs, to estimate the impacts of the funded interventions.[22] As for the Dominican Republic, IE is restricted to a very limited set of programs which have substantial funding coming from the IADB and the WB. In this specific case, the original loan documents condition funds disbursement to produce a sound impact evaluation of the youth active labour market program "Juventud y Empleo".

These international forces play an important role in the three countries, but there are other international movements/organizations that may generate similar shocks in other poor countries in the region, such as the World Bank, IADB, J-PAL, 3IE, among others, which contribute to the international effort for promoting evidence-based policy-making around the world. The question is to evaluate to what extent such pushes may sustainably alter the production of rigorous IEs and promote the systematic use of IEs for policy making in our sample of countries. We start by first establish the quantitative importance of both shocks in the corresponding countries. As mentioned in appendix C, Peru has had a total of 31 programs/interventions with a relatively sound impact evaluation strategy over the past 15 years, which is the largest number within our sample of countries. Of those, we identified nine that were promoted and implemented by IPA. In the case of El Salvador, six of the 11 programs identified with sound IE strategies received funding through the MCC-GOES compact and are coordinated by Fomilenio. Although the participation of these institutions is sizable and similar in number in both countries, they are very different in nature. For the case of Dominican Republic, sound IE, which incorporated the evaluation components at the time of the program design, have been restricted to different rounds of the program "Juventud y Empleo".

---

[20] Following Briceño and Gaarder (2010), we understand *institutionalization* as a process of channeling program evaluation efforts through a formal system that accompanies program design and implementation, generate the IE studies, and define its use for policy making with appropriate benchmarks and analyzing trade-offs across interventions aiming towards some common goals.

[21] IPA has as affiliates experts in development economics from leading universities such as Harvard, Yale, MIT, LSE, among others (see http://www.poverty-action.org/). They have gradually opened offices in developing countries around the world to promote the use of RCTs and to facilitate fieldwork and monitoring.

[22] The Millennium Challenge Corporation (MCC) is a US foreign aid agency that aims to contribute to the reduction of poverty and the achievement of the MDGs.

A first aspect is the extent to which these shocks involve the governments themselves. Although IPA exclusively promotes the use of RCTs, their work in Peru has not yet been able to involve the government. Most of their work focuses on microfinance and is carried out in association with regulated and non-regulated microfinance institutions (MFIs). This is clearly not a negative point, as the microfinance sector is one of the friendliest in terms of their use of hard-evidence to guide innovations to improve financial services to the poor, at the international level and also in Peru. So, the IPA projects are likely to have a significant influence in the way Peruvian MFIs serve their referred population. However, its potential to influence the way public policy is created by the Peruvian public sector, is currently negligible. Obviously, this is not due to lack of effort on behalf of IPA, who is adamantly interested in participating in the impact evaluation of public programs. However, IPA's motivation is mainly academic and therefore mainly interested in conducting RCT. The government agencies are generally reluctant to engage in an RCT as they impose additional costs for the implementation stage for their programs.[23] Indeed, we have not found any public program in Peru with an experimental design to determine its impacts.

Another aspect that probably goes against a more meaningful contribution of IPA work on the institutionalization of the use of IEs in Peru, is the fact that they seldom involve a local researcher as an author. For instance, a researcher based in Peru will have more interest in, as well as more mechanisms to influence the way policies are designed and implemented in the country. One has to keep in mind though, that IPA makes significant efforts to promote the use of its results to guide the fight against poverty, but it is probably true that their main target audience is the international donors and policy spheres rather than the local ones.

The MCC-funded programs present a very different picture with respect to the participation of the local and national government in El Salvador. As mentioned above, and in appendix C, the MCC works through agreements with the GOES, so that the implementation is always conducted by local or national governmental agencies. Furthermore, a special agency is generated to coordinate the efforts against poverty under the agreement, called compact. In the case of El Salvador, this agency is called Fomilenio. This also means that the MCC and Fomilenio have to deal with limited human resources. Training offered for key policy makers played a decisive role. This training included components to help them design procurement processes and to increase awareness about the importance of evidence-based policy-making. As a result of the training, the government agreed to have two programs evaluated using an experimental design (out of six), despite initial reluctance (see Moreno et. al., 2010).

A key point to understand the possible contribution of the MCC agreement is that it deals with a subset of the Salvadoran public sector and has a finite duration of five years. Thus, the question is whether the strengthening of capacities achieved with the agreement can expand to the other sectors and can have sustainable effects on the way policy is designed and implemented in El Salvador. Fomilenio officials indicate that coordination meetings with implementing sectors (ministries) have gradually incorporated the participation of the Evaluation Office of the Technical Secretariat of the Presidency (STP). The secretary of the Presidency indeed presides the consulting committee of Fomilenio, and shows interest in expanding and sustaining the use of IEs for the design and implementation of social programs. In the next sub-section, we explore the strengths

[23] See Moreno, et. al., (2010) for a list of the usual justification for the public officers' reluctance to accept an experimental design.

and weaknesses of such a unit for becoming a champion of the institutionalization of IEs for policy making in El Salvador.

For the Dominican Republic case, there has not been any institutionalization or evaluations beyond the ones mentioned above. IEs appear so far as a by-product of international lending rather than a genuine demand from government actors.

We identified embryonic processes of institutionalization of the use of IEs for policy making in Peru and El Salvador, but not in the Dominican Republic. However, they differ in their origins and their level of development. In the case of Peru, institutionalization is initialized with the Results Based Budgeting approach and included in the Law of Public Budgeting in 2007. It is the General Direction of Public Budgeting (DGPP) of the Ministry of Economics and Finance (MEF) that are in charge of implementing the approach. In El Salvador, the driving process for institutionalization is led by Fomilenio, which should be conducted by the Technical Secretariat of the Presidency (STP) at the end of the MCC-GOES Compact.

A first key difference between the two identified processes is that the Peruvian one is mainly locally driven while the Salvadoran one was initiated via the external shock generated by the MCC-GOES compact. The Peruvian process starts with the growing awareness and increasing evidence on the way public resources are being wasted under some of the most important public programs for the poor. This evidence developed by local and international researchers was being accumulated over the years and became increasingly exposed in media during the first decade of the century, generating a space for initiatives that could bring some order and sense to the implementation of social programs.[24]

Both processes have been operating for several years now, and have led to some important achievements. They face a crucial juncture in which they need to consolidate their efforts towards the institutionalization of IEs for policy making in the corresponding countries. The Peruvian RBB process have generated several process evaluations that have helped reorganize some of the key programs through consensual agreements with the sectors involved, and budget reallocations have followed in favor of programs that have successfully adopted the recommendations (see appendix C). The Salvadoran Fomilenio, on the other hand, has led the decision to attach rigorous impact evaluations to each of the programs they fund. In two cases, Fomilenio was able to use an experimental design, despite the logistical adjustments they often require.[25] They have also trained policy makers on the theory and practice of impact evaluations, which have likely been instrumental in getting the support of the implementing sectors for the rigorous identification strategies.

Both processes are in crucial junctures to consolidate their efforts to institutionalize IEs for policy making in their corresponding countries. The Peruvian RBB team considers next step key to insert the use of rigorous impact evaluations in their process, so that budget reallocations can be guided not only by performance indicators associated to intermediate results, but by a causal link between performance and results linked to the programs' ultimate goals. As part of the DGPP, they are

[24] See Alcázar (2003) among others, as examples of variants of benefit incidence analysis (BIA) done by local and international researchers that showed severe leakages among several key social programs in Peru.
[25] See section 2.b for a discussion of the implications of implementing an experimental design for an impact evaluation.

indeed able to meaningfully affect the design and implementation of social programs, just like the Chilean DIPRES is (see Briceño and Gaarder, 2010). A new challenge ahead will be on the means to guarantee a level of independence of their work, not only from the implementing sectors but also from the government as a whole, to avoid improper influences in the generation and dissemination of assessments. The RBB process is still exclusively a unit within the DGPP (MEF), without any participation from an external body, such as CONEVAL in Mexico, or DIPRES in Chile. Briceño and Gaarder (2010) add that independence of an oversight body also depends on the funding rules, the reporting structure and dissemination laws. Furthermore, they argue that in the case of the Chilean Dipres, the transparency rules for the dissemination of results and the international advisory panel are key elements for the credibility of the organization, considering that organizationally is clearly dependent of the Ministry of Finance. These elements need to be considered for the consolidation of the process started by the Peruvian RBB team.

The issue of independence is also relevant for the embryonic Salvadoran process, since the unit that has become in charge of Fomilenio's achievements is the Secretary of the Presidency. However, their major challenge would be to sustain those achievements after the conclusion of the MCC-GOES compact in 2012.

6. The supply of training in IE methods in LAC

The surge in the production of IE studies has come together with important methodological innovations within experimental and non-experimental approaches, some of which have not been easily followed by researchers and policy makers in our sample of countries. Thus, many organizations interested in promoting the use of IE studies for policy making in these countries have been required to implement training programs for these key agents, not only to support the production of IE studies but also to spur demand for them and their use in designing new programs or adjusting policies. In this section, we present the results of a systematic search to identify who has been doing such efforts in the region.

We searched for information about training efforts, whose documentation can be found on the internet by usual subjects in the area and the region: The World Bank, the Inter-American Development Bank (IADB), the Abdul Latif Jameel Poverty Action Lab (J-PAL), Innovations for Poverty Action (IPA), the International Initiative on Impact Evaluation (3IE), the Impact Evaluation Network (IEN) of the Latin American and Caribbean Economic Association (LACEA), the Millennium Challenge Corporation (MCC), among others[26].

We were able to identify a total of 39 courses held in the region between 2001 and 2011[27]. By far, the most important institution has been the National Institute of Public Health (INSP), from Mexico, that has been running their workshop on Impact Evaluation of Population, Health and Nutrition Programs in Cuernavaca since the beginning of the last decade, with the support of USAID. That workshop has been directed towards scientist from all countries in the region and of

[26] An additional effort was made with institutions like IADB, the World Bank, IPA and J-PAL, for which some representatives assisted in completing or verifying information.

[27] You can find the full list in Appendix D, with additional information about locations, materials, type and number of participants, when available. Although likely not exhaustive, the time trends and actors identified provide useful information for any institution interested in fostering this process in our sample of countries.

different disciplines that work with such issues, including economists, although they were not from the prevalent discipline within the audience. Other important actors have been IPA (5), the World Bank (4), the IEN (4), J-PAL (3), among others. It is very important to note that most of these courses have taken place after 2006, which shows the increasing importance of these activities.

In addition, we were able to identify the number of participants for 23 of the courses listed in Appendix D. For that sub-sample, the average number of participants was 49 per course, with many of them including both, researchers and policy makers. This average, however, vary significantly across training institutions. The Mexican INSP courses had between 15 and 20 participants each year, with a duration of about three weeks. The World Bank courses, on the other hand, had between 100 and 200 participants each time, but lasts only 3-5 days.

In sum, it is clear that IE training is becoming increasingly common in the region. However, it is likely that more efforts are needed to expand outreach in countries with weaker research capacities, and to intensify the treatment to combine training with technical assistance, especially in the case of local researchers in countries with weaker capacities for this kind of research.

7. Conclusions

This study looks at the way impact evaluation studies are being produced and used for policy making in a sample of countries in Latin America and the Caribbean, a priori considered having less research capacities to absorb the current trend observed in other more developed countries in the region. The contribution of this study is threefold: (i) we performed a systematic search of the studies that evaluate the impacts of programs and policies with sound identification strategies, and analyzed time trends and key actors in the demand, production and funding of the studies (ii) we performed three case studies (Dominican Republic, El Salvador and Peru) to explore the institutional factors that work in favor and against the demand and use of rigorous impact evaluations for policy making, and (iii) we searched and identified training activities performed by main actors for the promotion of the production of IEs, and their use in the shaping of policies and programs.

Following Bouillon and Tejerina (2007), we conduct a systematic review of IE studies in a selected set of countries where the use of IE approaches for program evaluation is scarcer. We limited the systematic review to IE studies that offer a strong empirical strategy for the identification of the impact(s) of interventions, thus excluding studies based on beneficiary satisfaction and participation self-evaluation. The systematic review suggests that Latin American and Caribbean (LAC) countries have experienced a large increase in the number of IE studies conducted in the last decade, and the time trend in our sample of countries is similar to that one in the rest of LAC. Peru has been very productive and is clearly a leader in the first sample while Mexico leads the second group. In both areas, about 70% of the studies were produced after 2005. In terms of thematic focus, social protection programs make up for the largest share of the evaluated programs, 24% of the studies in our area of focus and 29% in the rest of LAC. This is partially due to the fact that most of the countries in the region have implemented a cash transfer program, but also to the example of the Mexican Progresa-Oportunidades program that benefitted from a rigorous impact evaluation strategy. In addition, accessibility to IE databases led to a

multiplicity of studies per program. In both our area of study and the rest of LAC, programs in the fields of education and entrepreneurship (including microfinance) were also assessed through rigorous IE studies. On the other hand, agricultural and rural development programs are more important in our area of focus, while urban development programs are more prevalent in the rest of LAC, which is likely a reflection of the difference in relative importance of rural and urban poverty in the two groups of countries.

Differences are also found in terms of the empirical strategy for identifying the impact(s), the source of funding, the involvement of the government or implementing agency and the involvement of local researchers. Randomized experiments are more common in our sample (43%) than in the rest of LAC (36%), although both groups of countries present a similar increasing trend in the use of RCTs. Matching methods is the most common method in the rest of LAC. Also, RCTs have been mostly used to assess CCT programs, and less so for the other types of programs. Job training and active labour market programs usually involve the use of matching and longitudinal empirical approaches.[28]

Multilaterals are more important in our sample of countries (77%) for the funding of the IE studies than in the rest of LAC (52%), and such funding seems to decrease the likelihood of participation of local researchers in authorship of studies. However, the participation of local researchers is generally low (in only 40% of the studies), although the proportion is increasing over time. Independent organizations (mostly NGOs) are more likely to be running the programs that are assessed in our study area (36%). In the rest of LAC, programs are mostly run by governments (74%).

To learn about how and if impact evaluations are used for policy making, we performed three case studies, one in El Salvador, one in the Dominican Republic and another one in Peru. We found an increasing trend in the production of rigorous IEs spurred by some external factors in the three countries, which are quite different in each case. In El Salvador, the external shock comes from the MCC-GOES compact that promoted the use of rigorous identification strategies to determine the impact of the interventions funded under the agreement. In the case of Peru, the shock came from the presence of Innovations for Poverty Action (IPA) in the country that promoted the use of randomized experiments, mainly for the evaluation of innovations in microfinance products. A key difference between these two shocks is that the Salvadoran one involved governmental agencies as implementing units, while IPA worked mostly with non-governmental microfinance institutions (MFIs). In the Dominican Republic, the external push for IE comes from IADB and WB. They earmarked some program funding for IE studies on specific programs.

We identify embryonic processes for the institutionalization of the use of IEs for policy making in El Salvador and Peru, but not in the Dominican Republic. Although the processes are very different in nature in the two countries, they both face crucial junctures at this moment, which will determine their expansion and sustainability. In the case of El Salvador the process is associated with the external shock, and the production and use of IEs through Fomilenio will be threatened by the end of the agreement with the Millennium Challenge Corporation (MCC) in 2012. Hence, it is crucial to strengthen the Secretary of the Presidency so that they can sustain the gains and expand them to the rest of the GOES. In the case of Peru, the Results Based Budgeting (RBB) initiative

[28] Although more recently there is evidence of such programs using RCTs.

was driven by internal forces. It has already accomplished significant achievements in the organization of social programs and the reallocation of budgetary resources based on performance. The next step is precisely the systematic use of IEs for budget allocations, and it will require major adjustments by the new administration that took office in July 2011. Furthermore, both processes are currently located within the executive branch. There is still a need to build a proper balance between independence and relevance (Briceño and Gaarder, 2010).

In sum, we see from the systematic review and the case studies that many of the countries in our sample are facing some external shocks in favor of the production of IE studies, but they differ in their intensity as well as in the likelihood to affect public policy making in the countries. One sound hypothesis is that knowledge of the most modern methods of impact evaluation, and of the way to use them for the design and adjustment of policies and programs, is a key determinant for the adoption of IE studies for policy making, and such resource is rather scarce in our sample of countries, mainly from Central America, the Caribbean and the Andes. Indeed, such hypothesis seems to be supported by primary international actors in the production and use of IEs. Many of them are taking action and offering training activities, not only aimed for local researchers but also local policy makers. However, it is likely that more such efforts are needed to expand outreach in countries with weaker research capacities, and to intensify the treatment to combine training with technical assistance, especially in the case of local researchers of countries with weaker research capacities of this kind.

The case studies also show that we cannot overlook the need to support the construction of institutional frameworks in favor of a systematic use of IE studies to increase accountability of public action against poverty. The political economy of such processes is very complex and there are often opposing forces that lose power with increased accountability. Such support, though, is not likely to be standardized and would require a clear diagnostic of the political economy behind the current institutional framework in each country.

References

Alcázar, Lorena; José López-Calix; Erik Wachtenheim (2003). "Las pérdidas en el camino: Fugas en el gasto público (transferencias municipals, vaso de leche y sector educación". Instituto Apoyo, 172 pp., Lima.

Bouillon, César P. and Luis R. Tejerina (2007). "Do We Know What Works? A Systematic Review of Impact Evaluations of Social Programs in Latin America and the Caribbean". Inter-American Development Bank, Washington DC.

Briceño, Bertha and Marie Gaarder (2010), "Institutionalization of government evaluation: balancing trade-offs", *Journal of Development Effectiveness* 2 (3): 289-309, September.

Crone, Amy (2008). "MCA Monitor Analysis: El Salvador Field Report". Center for Global Development Report, August. E-version available at the following link: http://www.cgdev.org/files/16595_file_El_Salvador_FINAL.pdf.

Heckman, James J. and Jeffrey A. Smith (2005). "Assessing the Case for Social Experiments". *Journal of Economic Perspectives* 9 (2) : 85-110, Spring.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith and Petra Todd (1998). "Characterizing Selection Bias Using Experimental Data," *Econometrica* 66(5): 1017-1098.

Heckman, James J. and Jeffrey A. Smith (1999). "The Pre-programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies," *Economic Journal* 109 (457): 313-48.

Heckman, James J., Robert Lalonde and Jeffrey A. Smith (1999). "The economics and econometrics of active labor market programs," Handbook of Labor Economics, in: O. Ashenfelter and D. Card (ed.), Handbook of Labor Economics, edition 1, volume 3, chapter 31, pages 1865-2097 Elsevier.

Lindquist, Evert (2001). "Discerning policy influence: Framework for a Strategic Evaluation of IDRC – Supported Research" School of Public Administration, University of Victoria, Victoria.

Mesa de Concertación para la Lucha contra la Pobreza (2010). "Reporte de Seguimiento Concertado del Programa Articulado Nutricional: Evaluación del Año 2009". E-version available at the following link: http://www.mesadeconcertacion.org.pe/documentos/documentos/doc_01460.pdf.

Millennium Challenge Corporation (2006). "Millennium Challenge Compact with El Salvador: Executive Summary". E-version available at the following link: http://www.mcc.gov/documents/agreements/compact-112906-elsalvador-executivesummary.pdf

Moreno, Lorenzo, Larissa Campuzano, Dan Levy and Randall Blair (2009). "Towards Closing the Evaluation Gap: Lessons from Three Recent Impact Evaluations of Social Programs in Latin America and the Caribbean". *Well-Being and Social Policy* 5 (2): 1-23, Second Semester.

Parodi, Sandro (2005). "Evaluando los efectos del Seguro Integral de Salud (SIS) sobre la equidad en la salud materna en el contexto de barreras no económicas al acceso a los servicios". GRADE manuscript (e-version available at http://cies.org.pe/files/active/0/parodi.pdf.

Perova, Elizabeth; Renos Vakis (2009). "Welfare impacts of the "Juntos" Program in Peru: Evidence from a non-experimental evaluation". World Bank manuscript (e-version available at http://www.mef.gob.pe/contenidos/pol_econ/documentos/Perova_Vakis_JuntosIE.pdf, March.

Rosas, David (2006). "Impact Evaluation of PROJoven Youth Labor Training Program in Peru". OVE/EPPER-04/06, July.

Skoufias, E. (2001). "PROGRESA and its Impacts on the Human Capital and Welfare of Households in Rural Mexico: A Synthesis of the Results of an Evaluation by IFPRI". International Food Policy Research Institute, Washington, D.C.

Valdivia, Martin (2010). "*Contracting the Road to Development: Early Impacts of a Rural Roads Program*". PEP-PMMA Working Paper 2010-18, October

Weyrauch, Vanesa and Gala Díaz Langou (2011). "Sound Expectations: from Impact Evaluations to Policy Change", International Initiative for Impact Evaluation Working Paper #12.