

From Research to Policy:

Using Evidence from Impact Evaluations to Inform Development Policy

Iqbal Dhaliwal and Caitlin Tulloch¹

J-PAL, Department of Economics, MIT

Abstract

Recent years have seen a tremendous growth in impact evaluations of development policies, and an increasing trend towards considering rigorous evidence while making policy decisions. But there is much greater scope for incorporating results from impact evaluations into policy decisions, even in the presence of political and administrative constraints. We discuss how evidence is currently incorporated into policymaking, the constraints on greater adoption of evidence-based policy, and suggest specific ways that stronger policy–research partnerships can help overcome these constraints. Recognizing that there are many different channels to influence policy, we provide examples from the Abdul Latif Jameel Poverty Action Lab’s (J-PAL’s) policy outreach work in the hope that this paper can help in greater sharing of best practices from the impact evaluation field on how best to use evidence to inform policy.

¹ The authors are grateful for the comments they received from Rachel Glennerster, Mary Ann Bates and from Howard White who organized the panels where many of J-PAL’s ideas in this paper were first presented. The authors gratefully acknowledge generous funding from the National Endowment for Science, Technology and the Arts (NESTA), for writing a previous version of this paper. All views expressed in this paper are the authors’ personal views and do not necessarily reflect J-PAL’s position.

1. Introduction

Billions of dollars are spent every year on development programs, but until recently there was relatively little rigorous evidence on the true impact these programs have on the lives of the poor. Different programs targeted at the same policy outcome can have very different results, but without clear evidence on their final impact there is little guidance for policymakers on which program to choose. In recent years, rigorous impact evaluations of social programs have emerged as a robust tool to guide social policy in developing countries. In particular, randomized impact evaluations that allow for precise measurement and attribution of impact can help policymakers identify programs that work and those that do not, so that effective programs can be promoted and ineffective ones can be discontinued. Recent years have also seen greater awareness of the need to use such evidence in policy decisions.

But we believe that there is still much greater scope to incorporate evidence from rigorous evaluations in policy decisions: while the use of rigorous evaluations in assessing development programs has increased exponentially in the past few years, they still only cover a small subset of development programs. Moreover, this scarcity of rigorous evidence is compounded by the technical language in which it is often presented (typically in academic journals) and there is need for greater effort to communicate the results to a policy audience. Finally, the growing body of evidence itself poses a dilemma—a policymaker may have difficulty comparing different studies, especially if there is not clear guidance on how to relate new evidence to the existing body of research. While some development organizations (broadly defined as governments, non-governmental organizations (NGOs), foundations and international development organizations) have the technical capacity and a mandate to incorporate evidence in their decisions, the above constraints can make evidence hard to access and use for many policymakers.

When deciding between alternative programs, policymakers have to take into account a number of factors including political constraints, administrative capacity, technical feasibility, time pressures and limited finances. We do not believe that these considerations can or should be abandoned in favor of purely evidence-based policy—these will always be salient for policy decisions, therefore addressing them is crucial for the successful implementation of any program. But we argue that there is much greater scope for incorporating evidence in decision-making, even in the presence of such constraints, and that closer partnerships between researchers and policymakers can foster more evidence-based policy. Researchers can work to make results more accessible and understandable so that policymakers have the information necessary to make choices informed by evidence. At the same time, policymakers can partner with researchers to provide local knowledge on which problems are most in need of evaluation and facilitate evaluations of their programs, thereby generating even more rigorous evidence from the field and creating a virtuous feedback loop.

In discussing potential ways to improve the scope of evidence-based policy, we draw on our experience at the Abdul Latif Jameel Poverty Action Lab (J-PAL). J-PAL was founded in 2003 as a center in the Economics Department of the Massachusetts Institute of Technology (MIT) with a mission ‘to ensure that policy is based on scientific evidence

and research is translated into action.’ With regional center on five continents, J-PAL’s worldwide network of 70 researchers have undertaken more than 325 randomized impact evaluations of development programs in 50 countries in partnership with development organizations. J-PAL has also trained hundreds of researchers and policymakers on how to conduct randomized evaluations (REs), and has been active in disseminating research results about ‘what works’ to policymakers around the world and in promoting the scale-up and replication of effective programs.

However, as we realize almost daily in our work, policymaking in general, and using evidence in policymaking in particular, are complex issues. We do not claim or believe that J-PAL’s strategy for policy outreach is either the only or the best way to inform policy – many organizations have had significant policy influence in their fields and may have followed a completely different path to success. This paper is also not a literature review of studies analyzing the role of scientific evidence in policymaking, nor have we done a rigorous empirical evaluation of alternative approaches to influencing policy. Rather, this paper is an attempt to share our current understanding, based on our experience, of the gaps that we have observed between research and policymaking, and to present what are, in our opinion, a few useful strategies for bridging this divide. Similarly, we do not claim or believe that REs are the only way to generate rigorous evidence that can help inform policy. There are multiple research methodologies that provide scientific evidence to inform development policy, including regression discontinuity designs, statistical matching and instrumental variable estimates, and these methodologies also produce reliable and useful evidence. The suggestions in this paper for strategies to strengthen the links between research and policy apply just as much to any of these methodologies as they do to randomized trials. Our hope in writing this paper has been to spur dialogue and identify successful policy outreach strategies among other organizations whose mission, like ours, is to promote evidence-based policy.

2. Methods

This paper is based on the collective experiences of a number of our colleagues at J-PAL, gathered through shared work as well as from past presentations, interviews and talks. We also spoke with staff at many of the organizations with which J-PAL partners, including individuals at governments, international development organizations, NGOs, policy and research think tanks and foundations. In these discussions we attempted to understand the priorities of their organizations in making programmatic decisions, and where evidence falls in the spectrum of priorities. Additionally, we asked about the channels through which evidence entered the organization, and what kind of evidence was most successful at informing them. Particular focus was given to examples where evidence had or had not been successful at influencing decisions, and the reasons for this. We have kept the identities of all interviewees anonymous per their request, and use the pronoun ‘her’ for all of them, irrespective of gender.

3. Background: potential and actual use of evidence in policymaking

This section highlights some of the key considerations that factor in policymaking, describes how policymakers can benefit from the results of rigorous impact evaluations and discusses some of the constraints that prevent evidence from being used in policymaking.

3.1. Policy decisions are made based on many considerations

While there are many well-designed and executed policies, and many others that need no further evidence, we believe that there is significant potential for much more of development policy to be driven by rigorous evidence on program effectiveness. For example, if a policymaker wants to find the best intervention to increase student attendance, which intervention should they implement: constructing new buildings, community involvement, treating children for intestinal worms, conditional cash transfers or something else? In choosing between such alternative programs, policymakers have to consider many factors including political constraints on policy choices, administrative capacity to implement a program, technical feasibility of a proposal, time pressures to deliver results quickly and limited finances that preclude ambitious programs. But often such decisions also get made based on a combination of ideology ('we should not give parents 'bribes' to do the right thing for their children'), ignorance ('treating children for intestinal worms is a health intervention and has nothing to do with school attendance'), inertia ('we have been running this teacher training program for decades') or instincts ('I have a feeling that cash transfers won't motivate parents to send their children to school'). This can lead to programs and policies that are sub-optimal in terms of their impacts, costs or both. The problem of 'bad' policy choices is further compounded by the fact that once launched, programs are very hard to roll back, even if they are later found to be failures because of the political and personal challenge for the program architects, sponsors and implementers in accepting failure, and because of lobbying by entrenched and powerful interest groups like contractors, service providers, bureaucrats and local politicians who benefit from continued program funding.

At the same time, we are encouraged by how many policymakers are willing and often keen to consider evidence in decision-making. This willingness has been accelerated by demands for greater accountability and larger impacts from development programs by civil society in developing countries, and increased scrutiny of development spending by funders and donors. But even when policymakers are willing and eager to incorporate evidence in decision-making, they face the challenge of often not knowing where to find such actionable evidence, since most such results are either presented at research conferences or published in academic journals. When they do find rigorous evidence, policymakers may have difficulty interpreting it because it is mostly written for academic audiences in technical language. They also face the challenge of synthesizing evidence and drawing lessons from different research studies that may have been conducted in different years and contexts, and sometimes show seemingly conflicting results.

As a result, there is a wide variation in the use of evidence in policymaking. At one end of the spectrum are some international development organizations and large foundations that have full-time technical staff that analyze various research studies to

inform their organization's decisions. In recent years, these organizations have shifted to using the results of rigorous and independent impact evaluations from the field in making funding decisions, rather than relying solely on qualitative surveys, anecdotes or less rigorous evidence (for example, a before–after analysis, or a simple comparison of final outcomes of beneficiaries to non-beneficiaries). At the other end of the spectrum are governments, NGOs and foundations, both in developed and developing countries that may not have access to or the technical resources to analyze evidence from research studies.

3.2. *Policymaking can benefit from using evidence from REs*

While rigorous impact evaluations using different methodologies can provide critical insights into policy, REs are particularly well suited for a number of reasons, discussed below:

3.2.1. *REs measure impact rigorously.* REs compare the outcome of interest (for example, test scores in schools) of beneficiaries who received a program (the treatment group) to another group (the comparison group) that is similar in all respects except that it did not receive the program (for example, free textbooks). Measuring outcomes in this comparison group is as close as one can get to measuring how program participants would have fared without the program (the counterfactual). Like clinical trials, REs are designed to measure the impact *caused* by a program, so policymakers can confidently use this evidence to make important decision like whether to scale up a program, make design changes or discontinue it.

3.2.2. *REs can provide key insight into why programs succeeded or failed.* Researchers can design evaluations in a way that the different treatment arms, administrative data collected and quantitative and qualitative surveys provide key information on the underlying mechanisms that contribute to the success or failure of a program. To ensure that the right data is collected, it is important to delineate a theory of change, specifying what mechanisms the program is thought to act through and what specific outputs are supposed to be affected.

For example, J-PAL affiliates Abhijit Banerjee, Esther Duflo and Rachel Glennerster evaluated Indian nurses' attendance at work by installing time clocks and providing incentives based on attendance (Banerjee *et al.* 2008). After 16 months, there was no discernible impact of the program on nurse's attendance, in contrast to previous studies that had shown that incentives for attendance could have a large and significant impact.² Throughout the evaluation, researchers collected administrative data on when exemptions were given to nurses who missed work. This, and qualitative interviews with stakeholders, showed that the program was extremely unpopular with nurses, who over time petitioned their supervisors for more exemptions. Because supervisors, who were the enforcers of the incentives, could be pressured into relaxing the conditions of the incentives, there was no improvement in behavior. These findings suggest that future programs offering incentives for service provider attendance should ensure that incentives are automatically applied without giving blanket discretion to supervisors to relax conditions.

REs often also include evaluations of multiple variations of programs, allowing

² For a summary of this evaluation, see <http://www.povertyactionlab.org/evaluation/incentives-nurses-public-health-care-system-udaipur-india>

policymakers to understand which components are driving effectiveness. For example, Nava Ashraf, Oriana Bandiera and Kelsey Jack evaluated an incentive program designed to motivate volunteers to distribute female condoms, thereby preventing HIV infections (Ashraf *et al.* 2012). They tested several variations of an incentive scheme, including small financial rewards for condom sales, large financial rewards for condom sales and a non-monetary reward where higher sales volume was recognized with a star placed on a public ‘sales thermometer’. Agents who received non-financial incentives were the most effective at generating female condom sales: volunteers in this treatment group sold twice as many packs of condom as agents in any other group.³ On the other hand, financial incentives were effective only for the poorest of volunteers, for whom the value of the payment was relatively large. Because they tested multiple variations of a program, the authors could show not only that non-financial incentives tap into people’s intrinsic motivation for the cause but also that leveraging intrinsic motivation was more effective than providing financial incentives to drive sales of preventive health products.

3.2.3. REs provide practical information to help facilitate and guide scale-ups. Because REs are performed in real-world situations, often with implementing partners who could themselves expand the program if it were proven effective, they can yield many valuable practical insights beyond simple estimates of program effectiveness. Since many evaluations also measure intermediate outcomes (for example, health staff attendance and not just final outcomes like infant mortality) and collect qualitative data, evaluations can provide real-time feedback to the implementer that allows them to make rapid adjustments to the way the pilot is being implemented to prevent the program from failing for avoidable reasons. Such learning from the pilot can then be incorporated into the scale-up of the program.

Field evaluations also provide critical proof-of-concept for an innovative new program, which can be crucial to get political, administrative and financial support for a larger scale-up. And because independent evaluations provide an objective and widely accepted ‘record’ of the true impact of a program, they can play an important role in institutionalizing change and ensuring continuity in organizations, such as governments, where civil servants are frequently transferred and new administrations try to reset the policies of their predecessors.

3.2.4. REs are easier for non-technical audiences to understand. The randomized methodology, although less common in the social sciences, is nonetheless much more familiar to policymakers than many other research techniques due to its use in medical trials. REs are also intuitively easier to explain than non-experimental or quasi-experimental methods that rely on statistical and econometric methods that often remain a ‘black box’ for many policymakers. And because the underlying program is actually implemented in the field, it can be relatively easier for the policymakers, implementers and program beneficiaries to see the program in action and understand the details of the program design, costs and impacts. This can be particularly useful in building support among various stakeholders for a wide scale-up of a program, and for sustaining the program when its initial champions or creators are no longer in power.

³ For a summary of this evaluation, see <http://www.povertyactionlab.org/evaluation/role-incentives-distribution-public-goods-zambia>

3.3. *If it is not disseminated well, even evidence from REs can be difficult for policymakers to interpret and use*

Even though evidence from REs is often easier to understand, it may not contribute to the policy debate and policy formulation unless it is disseminated well, for several reasons:

3.3.1. *Evidence is often presented for an academic audience.* Much of the rigorous evidence on program impact is presented in technical papers that are targeted for an academic, rather than a practitioner, audience. Because they are published in economics journals, many papers focus on the underlying economic intuition, econometric techniques used to arrive at unbiased estimates of program impact, how the observed impact validates or contradicts economic theories and numerous regression tables. Moreover, since these journals focus on issues of economic theory and analysis, many researchers and academic journals often do not draw strong policy conclusions based on evaluation results. Further, since the main focus of most research papers is on the design of the study and the results, many facts that most interest policymakers, such as context, implementation details and costs, are not covered in sufficient detail for policymakers to draw conclusions for their context.

3.3.2. *It is difficult to distinguish between evidence of different quality.* Research results should ideally be judged according to the scientific rigor of the methodology that produced them, so their policy conclusions can be given an appropriate amount of confidence. But in our experience it often requires significant technical expertise to discern the distinction between evidence of different quality. It also requires non-trivial time and effort, which many policymakers simply do not have, and as a result, many of them require assistance to understand which evidence is a reliable and rigorous guide for policy.

For example, a policymaker at a European government department dealing with development issues indicated that because many colleagues were unfamiliar with the different methodologies for generating evidence, they did not perceive any particular advantage to evidence from rigorous impact evaluations as compared to ‘homemade pre-post program evaluations’ based on administrative data or unscientific surveys.

Lack of understanding of what constitutes rigorous evidence can have an even more pernicious effect—it can deter the use of rigorous evidence in favor of less rigorous evidence selected to suit immediate needs. For instance, a development organization that competes with other agencies for funding from the US government performs rigorous evaluations of its programs to calculate its rate of return on program investments and to facilitate organizational learning. But this also means that they find that some of their programs are not actually having the intended impact. Because the government does not require rigorous impact evaluations of all of its grantees, most of the other competing agencies do not conduct such evaluations and end up presenting less rigorous evidence as proof that their programs are effective, making the more evidence-based organization’s work look less effective, and making it harder for them to get grants from the government.

3.3.3. *It can be difficult to compare evidence from different contexts.* It can be difficult to draw robust policy conclusions when comparing evidence from different countries. For example, when presented with evidence of the impact of deworming on reducing student attendance in Kenya, policymakers in India are likely to initially discount the policy lessons because of the differences between the two countries in terms of culture, institutions, infrastructure, geography, economy and so forth. In such a case it would be necessary for someone with an understanding of both the Kenyan study and the Indian context to explain to the Indian policymakers how, on the key drivers of impact, including the underlying economic or behavioral theory and cause of the problem, the two countries share similar features that make it likely that a similar impact would occur.

3.3.4. *It can be complicated to compare evidence from alternative programs aimed at the same policy outcome.* One of the biggest challenges in development policy is that policymakers are required to make choices between a number of *prima facie* ‘good’ options. For instance, it is not immediately clear what is the best way to reduce diarrheal diseases in rural settings where it is prohibitively expensive to provide piped water. Is it by distributing chlorine at schools, supplying it directly at home, or making it available at the source of water supply (a spring, well and so forth)? Should it be supplied at full cost, at a subsidy or totally free? All of these options come with robust theories of change, passionate advocates and anecdotal, qualitative and even rigorous experimental and non-experimental evidence that supports the common-sense logic of these approaches. In such situations, the challenge for a decision-maker is to choose the program that will have the biggest impact at the lowest cost—that is the most cost-effective program. But there are complicated assumptions and decisions that need to be made in creating a formal cost-effectiveness analysis, and this deters policymakers from conducting this type of analysis, even when rigorous evidence is available (Dhaliwal *et al.* 2011).

3.3.5. *Demand for evidence is time-sensitive.* Another factor complicating the use of evidence in policy decisions is the unpredictable time span in which policy decisions are commonly made (Lomas 2000). An evaluation and the associated academic paper can easily take two or more years to complete, and even with careful planning it is almost impossible to make this coincide with when policymakers really need evidence for decision-making. In France, for example, the Ministry of Education is currently evaluating the ‘of Excellence’ program, which gives boarding school spots to children from disadvantaged backgrounds.⁴ This evaluation was begun in 2009, and results are expected in early 2012, but the Ministry of Education has discussions with the Ministry of Budgeting every year for the annual allocations for the public boarding school system, and even a few weeks delay in the evaluation results being released could mean that the evidence would not be used until next year’s decision.

3.3.6. *Researchers are constrained in their ability to engage policymakers on a continuous basis.* Many of the previously described obstacles could be overcome if there was an active and continuous dialogue between researchers and policymakers (Lomas 2000), but this happens far less than it needs to for a number of reasons. First, there are relatively few

⁴ For a summary of this evaluation see <http://www.povertyactionlab.org/evaluation/boarding-schools-secondary-students-disadvantaged-backgrounds-france-internats-dexcellenc>

researchers doing rigorous evaluations of development programs compared to the very large number of policymakers in governments, NGOs, foundations and international development organizations around the world who could benefit from their findings. Given the specialization of researchers among sectors like education, health or governance, the effective pool of researchers in a particular sector is even smaller. This creates significant demands on the time of these researchers to discuss their findings with policymakers, and limits the number of policy outreach requests that they can respond to effectively.

Second, academic researchers who are often conducting some of the most innovative and ambitious research are not rewarded within their profession for presenting results to practitioners or explicitly trying to inform policy. In particular, those who have yet to get tenure face strong penalties if they take time off from producing papers to invest in policy outreach. Surprisingly, this sentiment can even extend to researchers at policy organizations. A full-time professional researcher at an international development organization who has done important work on poverty issues told us that they saw their role being confined to doing research and it was up to the policymakers in other development organizations to read, understand and interpret it whichever way they liked. This idea that research findings will gradually percolate into the policymaking dialogue, sometimes called the ‘enlightenment model,’ can create barriers to the use of evidence to inform decisions if researchers are not available to discuss the particulars of their research, or interpret their findings for interested practitioners (Weiss 1979).

Third, the comparative advantage, training and interest of many researchers, especially those in academia, is in doing research, writing papers for academic journals, and explaining their results to a technical audience. Informing policy requires a number of activities that researchers are not trained in and often do not have an inclination for, including identifying the best organizations to partner with, networking with key policymakers to build long-term relationships and trust, presenting results in a non-technical manner, and relating the policy implications of their research. This outreach therefore requires staff with strong technical skills in development economics and practical experience in policy formulation or implementation.

Fourth, informing and influencing policy is an effort-intensive, risky, and long-term commitment for researchers. It takes a long time to build relationships with policymakers and navigate all of the procedural and bureaucratic hurdles to get project approvals. There are numerous risks of such partnerships, including the difficulty in overcoming political hesitation to publish impact results (especially when they show the program in a poor light), and the often small and incremental nature of the policy changes that follow from such efforts. But perhaps one of the biggest risks that researchers face is that of frequent transfers of key officials in governments and the subsequent ‘orphaning’ or discontinuation of a project when its policy champions are transferred (or administrations are voted out). A recent study found that members of India’s top civil service – the Indian Administrative Service (IAS) – have an average tenure of just 16 months (Iyer and Mani 2012). These frequent transfers and the lack of institutionalized mechanisms for knowledge sharing with successors lead to a significant lack of continuity in decision-making and fulfillment of prior commitments. Researchers are often frustrated when, after months of building relationships and finally convincing a policymaker about the importance of certain research

findings, they have to again start from scratch with their successor who has not received any information about her organization's previous partnerships or may not share the enthusiasm for continuing that partnership.

These factors can hinder the development of strong linkages between researchers and policymakers, and sometimes even lead to a mutual feeling of mistrust between the two groups. Researchers may feel that policymakers are not responding sufficiently or quickly enough to what they believe is convincing evidence, while policymakers feel that researchers are too narrowly focused on the 'theoretical, perfect-world' situation with disregard for the practical 'real-world' constraints confronting policymakers.

4. Findings: promoting evidence-based policy

Having discussed the barriers that can prevent the utilization of evidence in policymaking, we now turn to some specific strategies for policymakers and researchers to partner to better incorporate evidence in decision-making. The tables 1 and 2, below, provide a list of some of these steps that researchers and evaluators can take themselves to promote effective partnerships with policymakers, and similarly some strategies for policymakers to strengthen partnerships with researchers. The section after that will go into strategies which overcome many of these obstacles in more detail.

4.1. Promote a culture of 'evidence-based decision-making' at policy organizations

In the past few years, some organizations have moved towards requiring a review of all relevant evidence as part of the formal process of proposing new programs to fund or implement. But this is not yet widespread and because of the range of evidence of varying quality that exists, it can be hard for decision-makers to be certain they are giving due weight to rigorous evidence, and easy for some to cherry-pick the evidence that supports their case. Thus merely requiring that evidence be considered is unlikely in itself to lead to a 'real' move towards evidence-based policy; it is important to also build the capacity of policymakers to become better users and producers of evidence:

4.1.1. Encourage policymakers to become better users of evidence. One key factor in promoting the use of rigorous evidence in policy is to build the capacity of practitioners to find, assess and incorporate rigorous evidence in their work (Lomas 2007). J-PAL conducts executive education courses every year during which affiliated professors and senior staff discuss the pros and cons of various research methodologies, use case studies to critique various research reports and encourage attendees to design their own evaluation with feedback from the trainers. To date, more than a thousand participants have been trained at J-PAL courses around the world, including staff of federal and state governments, international development organizations, NGOs and foundations. Many of the participants at these courses have gone on to design their own evaluations, while others are in key decision-making roles in development organizations where they are active users of research findings. J-PAL also works with organizations that have potential for significant policy influence (for example, senior civil servants in India or staff at very large NGOs like Pratham and BRAC) to design custom courses that help their key decision-makers become better users of evidence.

TABLE 1: OBSTACLES & STRATEGIES FOR FOSTERING RESEARCH-POLICY PARTNERSHIPS

RESEARCH/EVALUATION OBSTACLES	RESPONSES THAT ENCOURAGE RESEARCH-POLICY PARTNERSHIPS
1. Disconnect of evaluator from program design	Willingness to engage with policymaker in the concept and design of the program to be evaluated. E.g. Rajasthan Police
2. Unilateral decisions on design of evaluation	Work with policymakers to understand the underlying problem, and how the evaluation can be structured to answer key questions E.g. Immunization
3. Inflexibility in evaluation approach	Consider alternate evaluation design to accommodate political constraints and field realities without compromising rigor and objectivity. E.g. Minister's District
4. Only interested in "ideal", but long-term outcomes	Construct additional short term evaluation outcomes while continuing to design long-term measures. E.g. NRHM attendance vs. health outcomes
5. Evaluations that only measure impacts, not underlying causes	Qualitative data collection during impact evaluation to understand if program implemented per plan, and what worked or not and why. E.g. Nurses breaking machines
6. Begin policy engagement only at start of project	Actively participate in policy conferences, meet key policymakers, contribute to civil society debate via op-eds, books etc. in region of interest. E.g. Bihar conference
7. End relationship at completion of project	Willingness to stay engaged as a "technical" resources for policymaker even after publication, especially for scale-ups. E.g. Pratham (which in turn provided expertise for Ghana)
8. Not report negative results	Maintaining rigor and absence of bias in evaluation and reporting results despite above close relationship. E.g. Flour fortification with Seva Mandir
9. Shifting evaluation objective	Register hypothesis ahead of time to avoid allegations of data mining. E.g. J-PAL and partners for RCTs
10. Discuss only one narrow study (own research)	Link research to the entire body of evidence. E.g. DC Education Evidence Workshop, Cost-Effectiveness Analysis, "Poor Economics"
11. Focus only on new evaluation, not replications or scale-ups	Be willing to evaluate the replication of a program found to have succeeded in another context and well suited for this problem, not just "new and innovative" programs. E.g. Bihar deworming
12. Excessive use of technical jargon	Frame discussion in easy to understand language, communicate in a style policymakers are familiar with, and customize outreach to audience. E.g. Policy Bulletins, Briefcases
13. Funding for evaluation may be difficult for policymaker to raise	More organizations require impact evaluations in the programs they fund and dedicated funding groups makes this easier. E.g. DIME @ WB, 3ie, IGC, J-PAL Initiatives, IPA Funds

4.1.2. *Encourage policymakers to become better producers of evidence.* While examining evidence from completed evaluations is a valuable exercise, there will always be new interventions for which evidence does not exist. This may be because a program is entirely new, or adds an innovation to an existing idea. Funding and implementing organizations should encourage rigorous impact evaluations of such interventions wherever technically feasible and desirable, and active involvement of policymakers in the evaluation of programs is one of the best ways to get their buy-in in utilization of the results (Ward *et al.* 2009).

There are a number of ways to encourage such evaluations. Funding and implementing organizations can include funds for Monitoring and Evaluation (M&E) in the pilot program budget. Researchers can, through presentations at policy conferences, writings and meetings with policymakers, emphasize the need for evidence-driven policy. And organizations can lay down clear guidelines to encourage their staff to assess if an evaluation would be useful. For instance, the recent USAID Evaluation Policy outlines the conditions when an evaluation of its programs is required, stating that, ‘any activity within a project involving untested hypotheses or demonstrating new approaches that are anticipated to be expanded in scale or scope through US Government foreign assistance or other funding sources will, if feasible, undergo an impact evaluation’ (US Agency for International Development 2011). USAID has also created the Development Innovation Ventures (DIV), which provides funding for innovative development ideas and helps the Agency rigorously test these ideas to determine which are most effective at helping the poor. The initiative is also focused on taking proven ideas and scaling them up, in partnership with innovators and developing countries.

Finally, J-PAL also encourages organizations to become better producers of evidence by building their ‘in-house’ evaluation expertise. For example, in Haryana state in India, J-

FIGURE 2: OBSTACLES & STRATEGIES FOR FOSTERING RESEARCH-POLICY PARTNERSHIPS

POLICY OBSTACLES	RESPONSES THAT ENCOURAGE RESEARCH-POLICY PARTNERSHIPS
1. Political agenda trumps evidence	Target policymakers who are open to evidence, so they use it as an input along with other factors like political agenda, budget constraints and administrative capacity. e.g. TCAI in Ghana
2. Low policymaker capacity to consume, generate, or institutionalize evidence	Train staff at implementing organizations, establish M&E divisions, recruit competent people and motivate them via formal linkages with leading academics. e.g. Government of Haryana
3. Short-term horizon of policymakers	Combine short term outcome measures with long term outcomes and by encouraging phased rollouts to give opportunity to evaluate programs before large scale-up. e.g. NRHM
4. Risk-aversion and failure-avoidance by policymakers	Set up institutions that allow innovation and risk tolerance. e.g. Chile Compass Commission, French Evaluation Fund
5. Inability to build coalitions to support new programs	Work with other government agencies that are most receptive to evidence, even if they are not directly the social development departments. e.g. Finance, Governor’s Special Cell
6. Change in “rules of the game” after evaluation starts	Sign MoU (terms include phased roll out, control group, sample size, timeline, funding, data ownership, and scale-up). e.g. Government of Karnataka
7. Poor institutional continuity due to transfer of officials or new administrations	Everything listed in #6, plus build wider culture of evidence-based policy in the organization via capacity building and evidence workshops. e.g. Training of Indian Civil Servants (IAS)
8. Lack of pressure from civil society or legislature to conduct evaluations	Convince these institutions to demand evaluations via participation in civil society debate (op-eds, workshops, legislation). e.g. Mexican legislature created CONEVAL

PAL is helping to build the capacity of the state's Education Department to produce and use evidence for decision-making by helping set up an in-house M&E division. Similarly, J-PAL is partnering with the NGO Pratham in a collaborative evaluation of a literacy program that will build the capacity of their staff to perform rigorous evaluations of their own programs.

4.1.3. Use field evaluations as an opportunity to build strong long-term relationships between policymakers and researchers, while maintaining objectivity in reporting results. Policymakers can be better encouraged to use evidence in their decisions when they have closely partnered with researchers in all steps of the evaluation design and have benefited from feedback from the field to tackle unanticipated implementation roadblocks (Lomas 2000). Researchers and policymakers can jointly disseminate the lessons from the program and its evaluation to other policymakers so they can benefit from these dual perspectives. Such a collaborative process can encourage evidence-based decision-making at policy organizations.

For example, many evaluations by J-PAL affiliates often begin well before the program implementation starts with intense collaboration between researchers and implementers to discuss the underlying problem, perform a needs assessment for the various stakeholders, and consider various possible solutions along with the associated theory of change. Researchers share the results of previous evaluations in that sector, and work with implementers to help design promising interventions to test. Such partnerships can be with governments, international development organizations or NGOs as the implementing partners in the field and there are advantages and disadvantages in partnering with each, as described below:

Governments in developing countries are often the biggest funders and implementers of social programs, and working with them offers the chance to evaluate programs on a much larger scale than almost any other development organization. They also have the financial, technical and personnel capacity to widely scale up a program if it is proven effective in the pilot form. Work with governments can however involve long and cumbersome bureaucratic approval processes, a significant risk of projects being discontinued when the civil servants who championed the program are transferred, and wide variation in the skills and enthusiasm for change among civil servants. There is also the possibility of civil servants trying to influence researchers to effect changes in program design or in the publication of results to accommodate political pressure.

Multilateral or bilateral development agencies offer the chance to implement programs at a large scale, and can bring significant funding for the program. They also have some very skilled and experienced staff, both at the headquarters and regional offices, with experience working in numerous countries. However, their policy priorities can change significantly when their national governments change, and they often have a very specific geographic focus corresponding to their home country's strategic interests. For instance, in FY 2010, two of the largest recipients of funding from USAID were Pakistan and Afghanistan, whose total allocation dwarfs transfers to other countries. Similarly the UK's Department for International Development (DFID) has identified '27 focus countries on which it will now concentrate its bilateral funding. From DFID's

previous list of ‘priority’ countries, 16 countries have now been removed’ (DFID Research and Evidence Division April 2011).

Local NGOs in developing countries can be faster and more flexible in implementing new approaches and usually have very dedicated staff. Further, if NGOs do not have a guaranteed stream of funding for their activities in the absence of evidence on program effectiveness, they may face better incentives to rigorously evaluate their programs. But if there is already sufficient political and financial support for their activities without rigorous evidence, then rigorous evaluations could actually create risk in case the evaluation reveals that the program has little or no impact (Pritchett, 2002). It can also be difficult to identify a reputable, reliable and effective NGO from among a large and heterogeneous field in many developing countries and many NGOs also lack the resources or scale of operations to implement an evaluation with sufficient sample size, or to act on evidence via a large scale-up of effective programs. Moreover, programs tested with NGOs can sometimes face skepticism when presented to governments, because of the differences in scale and institutional design between government and NGO operations.

4.2. *Facilitate partnerships of researchers with policymakers*

As development organizations build a stronger culture of evidence-based policymaking, there is a need to facilitate ongoing partnerships between researchers and policymakers (Ward et al. 2009). As described earlier, there are many barriers to such partnerships, but it is important to facilitate them as not only do policymakers benefit from close interaction with researchers, but researchers also have much to gain from such partnerships. First, policymakers understand well the pressing issues facing their constituents, the local context and what the primary constraints on program options are, and they can therefore guide academics to the most relevant research questions, and also give them a sense of the difficulties that new programs may encounter. Second, researchers are reliant upon their implementing partners for the smooth implementation of any of the programs they are testing. If a program is incorrectly implemented, it becomes more difficult to draw strong conclusions about its results. A close feedback loop between researchers and policymakers ensures that such concerns are addressed quickly and effectively so that programs do not fail due to avoidable implementation problems.

Third, when policymakers see the researchers contributing positively by providing evidence from existing research and giving feedback on program design, they are more likely to be motivated to support considerations of evaluation design. This is particularly true in REs, where a clear separation of treatment and comparison groups is necessary and often requires the policy partners to be advocates to their colleagues, helping to explain why randomization will yield solid results and what can be gained from a rigorous evaluation. An example of this kind of advocacy is visible in an ongoing J-PAL project with the Government of Karnataka, where the researchers were actively involved in all stages of program and evaluation design and in response the government agreed to change many features of the program based on learning from previous J-PAL studies.⁵ They also changed

⁵ For a summary of this evaluation, see <http://www.povertyactionlab.org/evaluation/integrated-medical-information-and-disease-surveillance-primary-health-centers-india>

the rollout plan from 100 per cent coverage in two pilot districts to 50 per cent randomly chosen coverage in five districts, enabling a rigorous impact evaluation.

Fourth, if researchers stay engaged with policymakers, then the latter are more likely to approach the researchers for subsequent partnerships for evaluating other programs leading to a virtuous cycle of evidence-based policy. Fifth, while there is currently great interest in funding and promoting evaluations, if after a few years funding organizations feel that there is no linkage between the results of that research and policy, it is conceivable that the relevance of evaluations and its funding could diminish. On the other hand, if the results from rigorous evaluations are seen to feed into policy, it can attract other researchers to the field. For instance, recent innovative evaluations in governance (for example, issues around measuring corruption, using information to strengthen voter control over politicians and the role of community monitoring in strengthening public services) have helped establish the important role that evaluations can play in governance research.

We now turn to a discussion of some of the institutional strategies that J-PAL employs to facilitate new partnerships between researchers and policymakers that can complement those undertaken by researchers and policymakers as listed above.

4.2.1. Respond to requests from policymakers for evidence and partnerships. J-PAL seeks to facilitate the transfer of learning from impact evaluations to governments, foundations, NGOs and international development organizations. Consequently, J-PAL is frequently asked by policymakers to discuss evidence on a pressing policy challenge or to help identify researchers to help design and evaluate innovative programs. J-PAL has dedicated staff who respond to such ‘opportunistic’ requests by sending relevant information, making presentations, matchmaking researchers to policymakers and providing support to researchers to facilitate their interaction with policymakers.

4.2.2. Targeted outreach conferences for policymakers to disseminate evidence and match-make partnerships. J-PAL’s dedicated Policy Group complements opportunistic outreach with targeted outreach to organizations that work on a particular development issue, or in a region where the learning from J-PAL’s research could be particularly salient. When there is sufficient evidence about the development challenges of a particular region, and signals of ‘responsiveness to evidence’ among local policymakers, J-PAL identifies organizations and individuals that have the most potential to impact policy and focuses dissemination efforts on this group. Often, this involves organizing an event like a conference or workshop that features presentations by researchers and their policy partners to discuss the details of an evaluated program. Such events also provide valuable opportunities to make new contacts within organizations and to get insight into the most pressing policy or research questions that policymakers want answered.

For instance, in 2010, J-PAL identified Bihar (a state in India that has some of the lowest indicators in health and education) as a place where evidence from J-PAL studies was particularly relevant, and where the new political leadership had demonstrated a strong commitment to improving development outcomes. J-PAL reached out to the Government of Bihar and organized a joint conference that brought together researchers from the J-PAL network, their field partners from relevant projects, top politicians and civil servants from Bihar and development organizations that work in that state to discuss

pressing development issues. The conference addressed problems as diverse as how deworming children can tackle endemic health and education challenges, how double-fortified salt may hold promise in the fight against iron-deficiency anemia and how politician report cards can strengthen democracy.⁶ Based on discussions with the Government that immediately followed this event, the state government agreed to conduct a massive school-based deworming campaign in partnership with Deworm the World, that reached 17 million school-children aged 6–14 years from February to April 2011.⁷

In addition to such regional conferences, J-PAL also organizes thematic conferences that are primarily aimed at developed country foundations and international development organizations that are responsible for allocating funds to large development programs.

In 2011, J-PAL organized a conference on agricultural technology adoption with the Center of Evaluation for Global Action (CEGA) and USAID to highlight the role of technology in agricultural production, identify technologies that are appropriate for Africa and make matches between policymakers and researchers.⁸ For such matchmaking conferences, J-PAL staff identify organizations working in that sector from around the world, and talk to them about their research priorities and the programs that they need evaluated. Staff screen out organizations that do not have specific research questions or where their program may not be able to support the research design, and then invite the remaining groups to a conference. Simultaneously, researchers whose work focuses on related questions are invited to attend and give presentations on their research interests and past partnerships so that researchers and policymakers are able to find common questions from which new evaluations and partnerships can evolve.

4.2.3. Fund special ‘Initiatives’ to identify pressing areas for research and encourage a coherent research agenda. J-PAL has recently started creating new funds, termed ‘Initiatives,’ to promote original research, policy outreach and capacity building in areas of pressing policy need. The Initiatives begin with a comprehensive literature review to identify what we do not know about achieving the outcome of interest (for example, technology adoption in agriculture, or strengthening governance in developing countries). The resulting ‘review paper’ identifies the main areas for future research. Feedback is widely solicited from policymakers who specialize in the particular area, so the review paper acts a medium for dialogue and ideas exchange between a number of researchers and policymakers to arrive at a research agenda that is widely accepted by both researchers and policymakers. This paper then serves as the basis for a Request for Proposals (RFP), in which researchers and their field partners are invited to submit proposals to fund evaluations of innovative programs that address the research agenda. J-PAL also organizes conferences and outreach activities to disseminate the policy lessons from the research funded by these Initiatives.

4.2.4. Organize policy–research collaborations with large development organizations, including national governments, to tackle key policy challenges. J-PAL also responds to requests from large development organizations to set up policy and research collaborations aimed at

⁶ Details at <http://www.povertyactionlab.org/south-asia/bihar-conference>

⁷ Details at <http://www.povertyactionlab.org/scale-ups/school-based-deworming>

⁸ Details at <http://atai-research.org/about-atai/news/atai-matchmaking-conference>

finding answers to particularly challenging policy questions. For example, the federal Government of Chile requested J-PAL to convene a commission to identify the most pressing social problems facing the country, brainstorm innovative programs to tackle these problems, and help evaluate them.⁹ The ‘Compass Commission’ consisted of leading academics and policymakers from Chile and around the world who met several times and submitted their report to the Chilean President in summer of 2011.

4.2.5. *Create policy tools that allow researchers to effectively communicate their findings.*

J-PAL’s dedicated Policy Group works to strengthen the linkages between research and policy by creating policy publications highlighting the results of field evaluations and disseminating knowledge about ‘what works’. Some of these publications include:

1. *Policy Summaries*: For each evaluation conducted by a J-PAL affiliate, the Policy Group creates an ‘evaluation summary’: a two-page synopsis of the relevant policy questions, context of the study, details of the program being evaluated and the results of the evaluation.¹⁰ These summaries are targeted at a non-academic audience, and are therefore written in a non-technical style and are available in a searchable database on the J-PAL website.
2. *Policy Briefcases*: For evaluations that address a particularly relevant question for development practitioners, J-PAL creates expanded summaries called ‘briefcases.’ Around four pages in length, these print publications provide a more comprehensive summary of the project and allow outreach to a larger audience.¹¹
3. *Cost-Effectiveness Analyses*: One way to analyze results from multiple evaluations of programs addressing the same policy goal is to combine them in a cost-effectiveness analysis which calculates the ratio of the amount of ‘impact’ each program achieves to the cost incurred to achieve that impact. Cost-effectiveness analyses (CEA) can be extremely useful as it both summarizes a complex program in terms of an illustrative ratio of impact to costs that is easy to understand, and also allows the reader to use this common measure to compare multiple programs conducted in different contexts and different years.

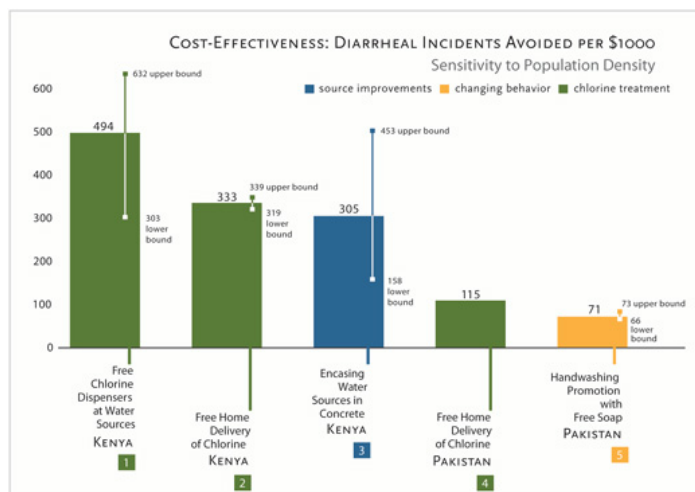
But relatively few impact evaluation studies include cost data in the published paper, and what data is available is typically presented in a wide variety of formats that make cost-effectiveness analysis time-consuming and complex. J-PAL’s Policy Group collects impact and cost data from programs that aim to achieve the same policy goal (for example, reducing diarrheal disease in children), and calculates the cost-effectiveness of each program, expressed as the cost of achieving one unit of impact (for instance, the cost to avert one incident of diarrhea). Graphs of the comparative cost-effectiveness of these programs and the underlying calculations provide policymakers the opportunity to examine why a particular program looked more or less cost-effective (see Figure 1).

⁹ Details at <http://www.povertyactionlab.org/LAC/compass-commission>

¹⁰ The database is available at <http://www.povertyactionlab.org/evaluations>

¹¹ J-PAL’s print publications can be accessed at <http://www.povertyactionlab.org/policy-lessons/publications>

Figure 1. Cost-Effectiveness: Diarrheal Incidents Avoided per \$1000 Spent

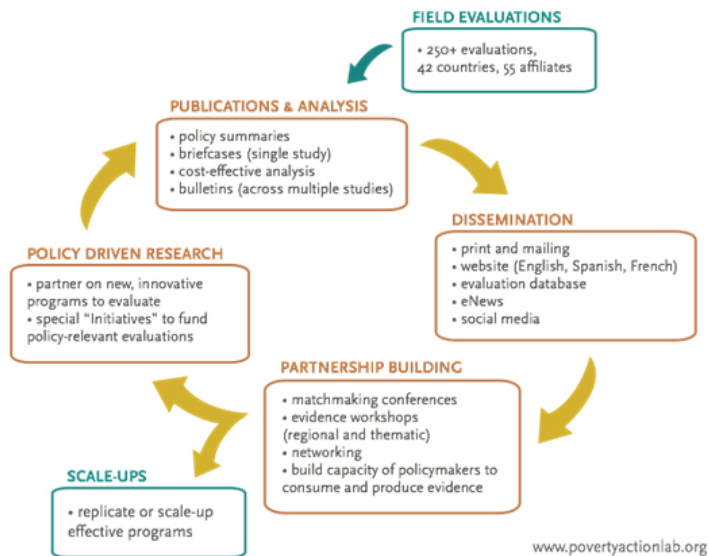


For a complete discussion of cost-effectiveness methodology in a policy context, including the large number of assumptions that need to be made to allow comparison and some of the reasons why researchers do not conduct enough CEA see Dhaliwal *et al.* (2011).

4. *Policy Bulletins:* The results of these CEA, as well as syntheses of evaluations which cannot be quantified in CEA, are presented in longer print publications called ‘bulletins.’ These documents provide a unified message that has been developed by the researchers from all the included evaluations about the overall learning from the included impact evaluations on particular development issues. J-PAL has seen great demand for these bulletins because they address policymakers’ difficulty in drawing coherent lessons from multiple evaluations, especially when sometimes these evaluations can appear to present conflicting results. Bulletins also provide more information on the sensitivity of policy conclusions to different contexts and assumptions, such as local institutions or the cost of program inputs.

J-PAL uses all these tools to facilitate policy outreach and build long-term relationships with policymakers. Sometimes these relationships begin when policymakers approach J-PAL or its affiliates for help identifying evidence regarding a particular policy challenge they are facing. In other cases, policy–research relationships begin after J-PAL identifies policymakers who are in positions where evidence-based policymaking could have significant benefits. Sometimes such relationship building begins even before a policymaker is in a position to directly use rigorous evidence, but has the potential to be in a position to encourage evidence-based policy in the future—for example, during executive training courses for senior civil servants. Figure 2 summarizes the key activities of the Policy Group:

FIGURE 2: J-PAL'S POLICY GROUP



5. Discussion: bringing effective programs to scale

Effective dissemination of research results to policymakers and strong policy–research partnerships can result in an interest from policymakers in scaling up effective programs. In such cases, J-PAL works in close partnership with the interested policymakers to compare the key details of the original program and the context-specific factors like the severity of the problem, the capacity of local implementing agencies to undertake the project, and the service delivery models. If this process reveals that the program has the potential to be successfully replicated in the new context, J-PAL works with the implementing organization to design a ‘policy pilot’ that tests the program in a small area. It is our experience that in most cases it is preferable to implement a replication or a ‘policy pilot’ before scaling up the program for a number of reasons.

First, because of the smaller size and therefore lower cost of implementing a pilot, policymaker and funders (internal or external) are more likely to approve such pilots to minimize the risk of allocating resources to a new program, especially if there is some level of uncertainty or debate about the optimum program design or expected outcomes.

Second, such policy pilots often incorporate important tweaks to the design or implementation process based on the learning from the original evaluation or in response to the differences in context between the two locations. As such, these pilots help the implementer understand how well the original impacts and costs generalize to this new context and provide additional data on proper implementation for the larger scale-up. For instance, J-PAL is working with its sister organization, Innovations for Poverty Action (IPA), as it partners with the Government of Ghana to pilot a program offering remedial

education through contract teachers in primary schools. This program is based on the NGO Pratham's remedial education program that was originally evaluated by J-PAL researchers in India, and if it is found to be effective, it could be scaled up to the rest of the country. This replication will provide useful information to help the Ghanaian policymakers learn what challenges and issues they may need to consider for a successful scale-up.

Third, when policymakers are actively engaged in the evaluation or pilot process, they acquire a much better appreciation of the importance of specific features of the successful program design and are less likely to oppose or modify it in a way that weakens the program. An example is the evaluation by J-PAL affiliates Abhijit Banerjee and others on increasing full immunization rates, which found that a predictable and regular supply of vaccines coupled with effective demand-side incentives for parents can lead to a large increase in full immunization rates for children in rural India (Banerjee *et al.* 2010). Many policymakers who have read the results of this study and want to use a similar program to tackle the problem of low immunization rates in their area tend to overestimate the impact of supply-side improvements and underestimate the impact of demand-side incentives, despite evidence from this study that the demand-side incentives caused the larger increase in full immunization rates. Some base this belief on the grounds that 'our context is different', while others are ideologically opposed to incentives on the grounds that they are tantamount to 'bribing' parents to do what is right for their children. Because of this overemphasis on the supply-side, many policymakers may be inclined to design a scale-up that includes only the supply-side intervention. In such situations, it is very useful to have a low-cost pilot study or replication evaluation where different variations of the program are tested so policymakers can see the relative impact of demand and supply side interventions in their context.

While not enough programs evaluated by J-PAL affiliates have been scaled up for us to draw general lessons, we provide some examples below of how different types of evidence in different situations led to scale-ups or widespread policy changes (for a comprehensive exposition on the components of the scale-up process, see Cooley and Kohl 2006).

5.1. A policy organization may have the capacity to massively scale up its own program after it is evaluated and found to be effective

Sometimes an implementing organization may have the capacity to use evidence from an evaluation of its own pilot to massively scale up that program. A well-known example of this is Mexico's PROGRESA program, which gave cash transfers to poor families conditional upon children's attendance at school and at doctor's appointments. Evaluations found significant effects on school enrolment as well as health outcomes. Shortly after the program was piloted and scaled up, a new party took power in Mexico. The strength of the evidence from these evaluations, as well as the immense popular support the program enjoyed, likely contributed to make it politically infeasible to discontinue the program, and it continued under the new name of Oportunidades (Gertler 2000, Schultz 2004).¹²

¹² For the history and political economy of PROGRESA, see http://www.brookings.edu/global/progress/pap_total.pdf, or

5.2. *A policy organization may leverage evidence to raise money to scale up its operations*

In 2001, J-PAL affiliates Abhijit Banerjee, Shawn Cole, Esther Duflo and Leigh Linden evaluated a pilot program of the education NGO Pratham's Balsakhi program which provided remedial education to school children in public primary schools in Vadodara and Mumbai cities in India. The evaluation found that the program had significant impact in raising the learning outcomes of participating students. Pratham used these results to demonstrate the effectiveness of its program and raise funds from Hewlett Foundation and Gates Foundation to massively expand this program to cover more than 300 of the 600 districts of India.¹³

5.3. *A program found to be effective in an evaluation is so highly cost-effective and simple to implement that other organizations readily replicate or scale it up*

J-PAL affiliates Michael Kremer and Edward Miguel, in their 1997–2001 study in Kenya, showed that school-based deworming was highly cost-effective: at an estimated cost of less than 50 cents per child per year, school-based deworming reduced the incidence of infection by 25 percentage points and reduced school absenteeism by 25 per cent. The relationship between taking a deworming pill and reduction in morbidity associated with worm infections is generalizable to areas that have similar worm loads. The program is also inexpensive to implement, as it piggybacks on the existing school infrastructure, requires minimal training of the teacher and is easy to administer. Following the evaluation, Michael Kremer, Esther Duflo and others from the Young Global Leaders forum assisted in the founding of Deworm the World, an NGO dedicated to promoting deworming policies in developing countries worldwide. Using the evidence from the original evaluation in Kenya, as well as similar evaluations in other countries, Deworm the World has helped make deworming a policy priority for both education and health organizations worldwide.¹⁴

5.4. *An evaluation provides critical and timely evidence on a very salient policy debate*

Sometimes timely and highly relevant evidence can spur organizations to make changes to policies or program designs. For instance, a 2006 study by J-PAL affiliate Pascaline Dupas and Jessica Cohen found that charging even small positive prices considerably decreased demand for insecticide treated bednets, and women who paid positive prices were no more likely to use the bednets than those who received them for free (Cohen and Dupas 2010). The study, and others, provided critical evidence in the midst of a raging debate on whether such preventive health products should be given for free or at a positive price. Soon many organizations changed their policy from charging a positive price for preventive health goods to distributing them for free.¹⁵

5.5. *An organization with a mission to promote evidence-based policy assists in a strategic scale-up*

In 2010, Innovations for Poverty Action (IPA) and J-PAL identified Ghana as a country that

<http://www.ifpri.org/sites/default/files/pubs/divs/fcnd/dp/papers/fcndp118.pdf>, or <http://www.bwpi.manchester.ac.uk/resources/Working-Papers/bwpi-wp-14211.pdf>

¹³ Details at <http://www.povertyactionlab.org/scale-ups/remedial-education>

¹⁴ Details at <http://www.povertyactionlab.org/scale-ups/school-based-deworming> and <http://www.dewormtheworld.org/>

¹⁵ Details at <http://www.povertyactionlab.org/scale-ups/free-insecticidal-bednets>

could benefit from innovative programs to improve the learning outcomes in primary schools. Staff discussed the NGO Pratham's remedial education model with Ghanaian education ministry, facilitated visits by them to schools in India where the program was being implemented, prepared prospective cost-effectiveness estimates for donors and provided significant technical expertise to the Ghanaian authorities to facilitate a large scale replication cum evaluation of a similar remedial education program that is currently underway.¹⁶

Having an organization that provides support to policymakers in the scale-up process can be a critical factor for success. While J-PAL does not itself implement the scale-up of programs found to be effective, it provides technical guidance to policymakers including sharing key insights from the original study, helping the implementing organization plan the logistics for the scale-up process, and providing assistance in monitoring and evaluating during the scale-up process. Such support can be especially critical for governments, as many of the policymakers there are so busy in implementing the large number of existing programs and reacting to crises that there is a very high value added of assisting them in program design, implementation planning, and providing continuous objective feedback from the field. As the Secretary of a key social department that implements numerous poverty alleviation programs in one of India's largest states once told us, *'Don't just tell me what the best strategy is, come and help me implement it.'*

6. Conclusion

Throughout this paper, we have deliberately used the term 'influencing' or 'informing' policy, rather than 'impacting' it, because it is not possible to apply the same rigorous criteria for measuring impact on the policy process. There is no counterfactual for what would have happened in the absence of a particular piece of evidence or a particular outreach strategy, and so we cannot claim to know exactly what the causal impact of policy outreach is. However, it is still instructive to examine the ways in which the research is disseminated to the policy audience, and look at examples where particular pieces of evidence have moved the debate or contributed to the adoption of a proven effective program. Over time, as J-PAL increases its policy outreach activities, there will be more learning to share. For now, the purpose of this paper has been to discuss why we believe partnerships between researchers and policymakers are critical to improving outcomes of development policy, the factors that help or hinder the use of evidence in policymaking, and what researchers (evaluators), policymakers and organizations like J-PAL with a mission to promote evidence-based policy can do to encourage such partnerships.

Even when policy decisions are based on evidence, they must take into account other factors like administrative capacity, political constraints, technical limitations, time pressures and budget limits. J-PAL's policy outreach works to make it easier for policymakers to include evidence as one more critical input in their decision-making process. REs are particularly well suited for providing evidence to the policy process because they measure impact rigorously, can provide key insight into why programs

¹⁶ Details at <http://www.poverty-action.org/remedialeducation/scalingup>

succeeded or failed, provide practical information to help facilitate and guide scale-ups and are easier for non-technical audiences to understand.

J-PAL tries to promote evidence-based policy by encouraging policymakers to become better users and producers of evidence and by using field evaluations as an opportunity to build strong relationships between policymakers and researchers while maintaining objectivity in reporting results. Similarly, J-PAL encourages researchers to work with policymakers by responding to requests from policymakers for evidence and partnerships, and performing targeted outreach to policymakers to disseminate evidence and match-make partnerships. In recent years, J-PAL has begun creating special funding Initiatives that help overcome some of the obstacles described above, and entering into policy-research collaborations with large development organizations, including national governments, to tackle key policy challenges. To support all of these activities, J-PAL's dedicated Policy Group creates policy tools that allow researchers to effectively communicate their findings. Success in the above efforts can facilitate the scale up of effective programs, often following a policy pilot in the proposed context. Organizations are also more likely to replicate a program when the underlying intervention is very simple to implement, highly cost-effective and generalizable to other contexts, and when an evaluation provides strong and actionable results in the middle of a particularly vigorous policy debate.

While there is a lot more to learn about what makes particular evidence more likely to be incorporated in policy, our experience has been that policymakers are more likely to use evidence in decision-making if that evidence is: *Unbiased* (independent evaluation not driven by any particular agenda); *rigorous* (used best methodology available and applied it correctly); *substantive* (provides insights that are novel or evidence on issues where there is a robust debate); *relevant* (fits the policymakers' context, needs and problems); *timely* (available when policymaker needs it to make decisions); *actionable* (comes with a clear policy recommendation); *easy to understand* (links theory of change to empirical evidence and presents results in a non-technical fashion); *cumulative* (draws lessons from not just one program or an evaluation but also the larger body of evidence in that area) and *easy to explain to constituents* (it helps greatly if researchers regularly inform policymakers and other stakeholders through conferences, policy papers, opinion pieces and so forth).

Throughout this paper we have given many examples of how J-PAL tries to foster strong partnerships between researchers and policymakers, while recognizing that there are many different ways to inform policy, and the approaches outlined in this paper may not be the best or most appropriate in all circumstances. Our hope is that this paper can provide some useful information to evaluators or organizations that are looking to do more policy outreach as to what may work or not work in that process, and also encourage other organizations whose mission, like J-PAL's, is to ensure that 'policy is driven by evidence and research is translated into action' to document the results of their efforts so that the broader evaluation community can benefit from these shared experiences.

References

- Ashraf, N., Bandiera, O., and Jack, K., 2012. *No margin, no mission? A field experiment on incentives for pro-social tasks*, CEPR Discussion Papers 8834. London: CEPR.
- Banerjee, A.V. and Duflo, E., 2011. *Poor Economics*. New York: Public Affairs.
- Banerjee, A.V., *et al.*, 2010. Improving immunization coverage in rural India: clustered randomized controlled evaluation of immunization campaigns with and without incentives. *British medical journal*, 340, c2220.
- Banerjee, A.V., Glennerster, R., and Duflo, E., 2008. Putting a band-aid on a corpse: incentives for nurses in the Indian Public Health Care System. *Journal of the European economic association*, 6 (2–3), 487–500.
- Cohen, J. and Dupas, P., 2010. Free distribution or cost sharing? Evidence from a randomized malaria prevention experiment. *Quarterly journal of economics*, 125 (1), 1–45.
- Cooley, L. and Kohl, R., 2006. *Scaling up- from vision to large-scale change*. Washington, DC: Management Systems International.
- DFID Research and Evidence Division, April 2011. *Operational plan 2011–2015*. London: Department for International Development.
- Dhaliwal, I., *et al.*, 2011. *Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education*. Cambridge, MA: MIT Press.
- Gertler, P.J., 2000. *Final report: The impact of PROGRESA on health*. Washington, DC: International Food Policy Research Institute.
- Iyer, L. and Mani, A., 2012. Traveling agents: political change and bureaucratic turnover in India. *The review of economics and statistics*, 94 (3), 723–739.
- Lavis, J., *et al.*, 2003. How can research organizations more effectively transfer research knowledge to decision makers? *The milbank quarterly*, 81 (2), 221–248.
- Lomas, J., 2000. Connecting research and policy. *Isma Canadian journal of policy research*, 1 (1), 140–144.
- Lomas, J., 2007. The in-between world of knowledge brokering. *British medical journal*, 334, 129–132.
- Office of Evaluation and Oversight, 2006a. *Assessment of the 2004 Project Completion Reports (PCRs) produced under the bank's new PCR guidelines*. Washington, DC: OVE.
- Office of Evaluation and Oversight, 2006b. *Evaluation of the IDB's studies*. Washington, DC: OVE.
- Pritchett, L., 2002. It pays to be ignorant: a simple political economy of rigorous program evaluation. *Journal of policy reform*, 5 (4), 251–269.
- Schultz, P.T., 2004. School subsidies for the poor: evaluating the Mexican PROGRESA poverty program. *Journal of development economics*, 74 (1), 199–250.
- US Agency for International Development, 2011. *USAID evaluation policy: learning from experience*. Washington, DC: USAID.
- Ward, V., House, A., and Hamer, S., 2009. Knowledge brokering: the missing link in the evidence to action chain? *Evidence & policy*, 5 (3), 267–279.
- Weiss, C.H., 1979. The many meanings of research utilization. *Public administration review*, 39, 426–431.