FINAL REPORT

Strengthening Institutions to Improve Public Expenditure Accountability Project

JANUARY 6, 2014

PRESENTED TO:
Ms. Ramona Angelescu Naqvi and
Ms. Savi Mull
Global Development Network
Delhi, India

PRESENTED BY:
NORC at the
University of Chicago
Mawadda Damon, Samuel Haddaway,
Ritu Nayyar-Stone, Aparna
Ramakrishnan
4350 East-West Highway
Bethesda, MD 20814
Tel. +1-301-634-9300

NORC at the UNIVERSITY of CHICAGO

This page left intentionally blank

Program Identification Details

Table 1.1: Programme Identification Details

GTF Number	CN-164
Short Title of Programme	Strengthening Institutions to Improve Public Expenditure Accountability
Name of Lead Institution	Global Development Network
Start Date:	16.10.2008
End Date:	15.10.2013[1]
Amount of DFID Funding:	GBP 4,993,829
List all countries where activities have taken or will take place	Table 1.4
List all implementing partners in each country	Table 1.4
Target groups- wider beneficiaries Society; the poor	Immediate beneficiaries are the selected research institutions whose research and communication capacity will be built for providing rigorous evidence based policy options for Public Expenditure Management (PEM).
Lead Author	Samuel Haddaway, Mawadda Damon, Ritu Nayyar-Stone, Aparna Ramakrishnan NORC at the University of Chicago 4350 East-West Highway Bethesda MD 20814, USA 1-202-286-3194 damon-mawadda@norc.org
Other people contacted to prepare this report	Table 2.4

[1] The closing date of the project is January 15, 2014 given that all activities and reports are to be completed within three months of the contract end date of October 15, 2013.

This page left intentionally blank

Table of Contents

Annexes – bound separately

Table of Exhibits

List of Acronyms

AST	Advanced Social Technologies, Armenia
BIA	Benefit Incidence Analysis
CBA	Cost Benefit Analysis
CBPS	Centre for Budget and Policy Studies, India
CEA	Cost Effectiveness Analysis
CEDS	Center for Economics and Development Studies, Faculty of Economics, Padjadjaran University, Indonesia
CFP	Call for Proposals
CIPPEC	Center for the Implementation of Public Policies Promoting Equity and Growth, Argentina
CIUP	Research Center of the University of the Pacific (El Centro de Investigación de la Universidad del Pacífico), Peru
CRC	Center for Research and Communication, University of Asia and the Pacific, The Philippines
CSEA	Center for the Study of the Economies of Africa, Nigeria
DFID	Department for International Development, UK
EGAP	Graduate School of Public Administration and Public Policy, Tecnológico de Monterrey, Mexico
EPRC	Economic Policy Research Centre, Uganda
ESRF	Economic and Social Research Foundation, Tanzania
FUNDESA	Fundación para el Desarrollo de Guatemala, Guatemala
FY	Financial Year
GBP	Great British Pound
GDN	Global Development Network
GDP	Gross Domestic Product
GTF	Governance and Transparency Fund
IBP	International Budget Project
IEA	Institute of Economic Affairs, Kenya
ISODEC	Integrated Social Development Centre, Ghana
M&E	Monitoring & Evaluation
MTEF	Medium Term Expenditure Framework
MTR	Mid-Term Review

NGOs	Non-Governmental Organizations
NORC	NORC at the University of Chicago
PBA	Program Budgeting Analysis
PCS	Policy Community Survey
PEM	Public Expenditure Management
PO	Participating Organization
PRAD	Policy Research and Development Nepal, Nepal
PRO	Policy Research Organization
PS	Policy Simulation
R4D	Results for Development Institute
RNPs	Regional Network Partners
RTI	Right to Information
SC	Steering Committee
TI	Transparency International
US	Unnayan Shamannay, Bangladesh

Executive Summary

The “Strengthening Institutions to Improve Public Expenditure Accountability” project aimed to strengthen the capacity of the 15 participating policy research organizations over a five-year period to monitor and analyze public expenditure choices, processes, and impacts and to engage constructively with policy officials to recommend improvements. The project’s ultimate goal was more capable, accountable, and responsive governments in the countries where the project operated. Populations in the countries where partner organizations are located were anticipated to benefit tangibly in the mid-term from improved government performance. The project management team consists of representatives from the Global Development Network (GDN), the lead organization, and the Results for Development Institute (R4D), the technical partner.

Participating organizations (POs) were to, over the five-year project life, perform four distinct but related budget analysis activities in the health, education, and water services sectors—program budgeting analysis; cost effectiveness analysis; benefit incidence analysis; and, development of evidence-based policy options for improving the results of public expenditures. The project implementation team provided capacity building in the corresponding analytic methods and communications/dissemination through workshops, mentoring through technical advisors, peer-engagement and learning, and the development of a resource-rich, easy to access website that is open to POs and others. This was a “learning by doing” approach wherein POs were able to implement the analysis and communications techniques and approaches learned in the series of analyses and reports required by the project while receiving technical support from advisors and the project management team.

It was expected that POs would master new skills and acquire valuable human capital that would be of service to their organizations as well as equip them for future work in the area of public expenditure monitoring. The project also aimed to deliver timely analyses effectively into the policy process with the explicit goal of improving public services and improving peoples’ lives.

The project included a robust monitoring and evaluation program to assess progress made in reaching the project's outcomes. This report builds off of five years of data collection and reporting for the impact evaluation and mid-term and final implementation evaluations.

Relevance

The project is relevant to increasing voice, accountability, and responsiveness. The countries targeted are ones in which voice and accountability were limited yet there was enough space for civil society to operate and constructively engage with the government to increase voice, accountability, and responsiveness.

The project was also aligned with national or local government priorities - all POs were to work in the health, education, and water sectors of their countries – key sectors of any country. They were to utilize research and analysis tools taught to them by the project to complete a program budgeting analysis (PBA) and benefit incidence analysis (BIA) at the general sector level before choosing a specific topic within each of the sectors for which to conduct a cost-effectiveness analysis (CEA) and then a "policy simulation" (PS) where-in they presented policy recommendations on the topic. All POs, except for two, considered government priorities in the selection of specific sector topics.

Impact

It is important to note that we cannot attribute direct causality to the project for the changes observed in the outcomes calculated using the survey data. Instead, we use anecdotal evidence and the opinions of respondents regarding the contribution of the project to observed changes in outcomes of interest over time.

Overall Impact

Quality of research

POs demonstrated an improvement in their report writing capabilities. Overall report scores increased by about seven percentage points over the observation period. Notably, the first reports produced by the project were on average of lower quality than non-project reports but by project end the project-produced reports were on average of higher quality than non-project reports, suggesting that POs at first struggled with the analysis and writing of project-proscribed reports but by project end had greatly improved.

Communications effectiveness

There was an increase in the policy community's contact with and awareness of POs. In particular, they were seeing greater mention of POs and their work in the media (newspapers, magazines, TV, radio). They had somewhat more frequent contact with POs through direct mail or email and had somewhat increased their use of POs' websites. This increase in contact and awareness was especially true for the second half of the project. GDN had responded to midterm evaluation findings that had showed a continued weakness in communications capabilities among POs and had hired a consultant to focus on improving communications training; this seems to have had an impact.

The policy community's perception of PO publications improved over the course of the project, particularly regarding how interesting and informative they were and their timeliness - areas where there were large statistically significant increases.

For those that attended the POs' events, the content and organization of events were perceived to improve greatly over the course of the project by the majority of respondents. Similarly, for those that use the POs websites, they were perceived as easier to use and containing more relevant information.

Despite these positive outcomes, there is room for improvement. A few POs mentioned limits in communication and dissemination skills and constructive engagement with the policy community that affected their ability to fully achieve their objectives regarding changing policy. The majority of policy community respondents met with POs only one or two times a year or less to discuss policy-related issues in 2013. Suggestions for enhancing policy impact from the policy community respondents included making research products more accessible to the public by utilizing media and in-country policy networks, producing deliverables that focused on specific policy issues, developing strategic partnerships with other organizations engaged in similar work for wider dissemination, and conducting consultations with concerned government officials and community organizations to better tailor their work to the specific contexts.

Perceived effectiveness in the policy arena

Our data show mixed results regarding POs' effectiveness in the policy arena over the project period. We find that POs are viewed as generally being more effective, with a large statistically significant increase in the perception that POs focus on issues that are of high priority and smaller statistically significant increases in the perception that they are valuable sources of research and provide helpful policy recommendations.

When asked specifically about whether the POs have an impact on public policy or administration, public expenditure quality, and the budget making process, there is less of a clear picture of positive change and no change in regards to the budget making process. Over half of respondents felt there was an improvement in whether the POs positively impacted public policy or administration but we found no statistically significant increase in score for this question; less than half of the respondents thought the POs had an impact in holding the government accountable for public expenditure quality but we found a modest statistically significant increase in the score for this question. There is clearly no change in the POs' influence on the budget making process as there were only about a third who thought there was any change and we found no statistically significant change in score for this question.

Individual Outcomes

Use of research produced by the project

Members of the policy community found the project-produced research to be of interest and relevant. Of those from the policy community who interacted with the project-produced research, 78 percent used the information for policy or administration-related purposes. Of these 78 percent, a further 80 percent shared the information with others. Policy community respondents interacted more with health and education research than water research most probably due to the challenges faced by the water sector. They also interacted more with program budget analyses and benefit incidence analyses, produced during the first half of the project, than with the cost-effectiveness analyses and policy simulations that were only just completed at the end of the project.

Policy impact outcomes of individual POs

We found that the policy community that had interacted with the project-produced research generally believed that it had contributed to policy changes, although to a somewhat limited degree. On average across all the countries, the majority of respondents felt the project-produced research was somewhat responsible for a change in policy (34 percent) or that it made a small contribution (28 percent). The extent of this contribution was limited since only four countries had any respondents who felt the project-produced research was mostly responsible for a change in policy – notably in India this was about one third of respondents.

Almost all POs spoke of having some sort of influence on policy; only two felt they were unsure of whether any of their project-produced research had a policy impact. Research produced in the health sector was the most frequently used with six POs mentioning their contribution to policy in this sector;

this is followed by the education sector in which four POs mentioned their contribution to policy. The water sector had only one PO that mentioned that their research had any contribution to policy.

Effect of the project on direct beneficiaries

Overall, the project seems to have contributed to improving competence levels in communications and performing program budget analyses, benefit incidence analyses, cost effectiveness analyses, and policy simulations in POs. By project end the majority of POs felt they had reached a "normal" competence level and had made changes to their communications practices due to what they learned from the project.

Economy

An analysis of the allocation of funds across major areas of spending and unit costs found that expenditures were appropriate to the context. About half of all spending was on grants to participating organizations, about 11 percent went to indirect costs, and monitoring and evaluation costs were within the standard of three to ten percent of the total budget.

Efficiency

Management and Coordination

Overall, the project was well-managed and did not serve to hinder the attainment of objectives. Rather, POs spoke of GDN's flexibility in adapting to their particular circumstances, their responsiveness to feedback obtained from POs on project implementation, citing some changes the project made based on their recommendations. However, there were some management and communication challenges, particularly in negotiating responsibilities between GDN and R4D at the beginning of the project but these were handled by the project team and by project end all viewed the project management quite favorably.

Dissemination of skills and knowledge among PO staff

The project was focused on improving the capability of the PO institution, rather than individual researchers, which required dissemination of the knowledge gained from those involved with the project to the rest of the institution. Substantial sharing of knowledge gained on research techniques is evident at nearly all POs. Materials were shared with others and informal conversations were had in over 85 percent of the POs throughout the project duration.

Attrition of project staff at POs did create significant challenges to attaining project efficiency. New staff had not attended the previous project trainings, the working relationships developed with project partners were lost, and it took extra to bring new staff up to speed.

Were the risks properly identified and well managed

GDN has paid close attention to possible risks that could be faced by the project; defining risks annually and taking risk mitigation measures. Starting with the project concept note that identified 4 potential risks, each annual report revisited possible risks and rated the likelihood and impact of the risk and provided a mitigation plan. One of the main risks was staff attrition at the POs; GDN managed this by increasing the number of researchers they funded to attend each of the training workshops from one to three. To mitigate the risk of delays in deliverables, GDN added a full time project associate dedicated to frequent communication with POs and keeping track of deliverables.

Effectiveness

Achievement of objectives

The project has showed promising evidence of attainment of some results and limited success in attaining others. There was significant progress towards building the capacity of POs - the quality of report writing increased and communications practices have been somewhat improved. Most of those who had interacted with the project-produced research had used the research but there was limited impact of this research on policy. It is important to note that the DFID requirement for the timing of the final evaluation report did not allow enough time after the end of project activities to measure the full effect of the cost-effectiveness analyses and policy simulations. The development of benchmarks for the quality of public spending was only partially attained and activities are not yet complete for the development of a Knowledge Portal containing comparable data from the participating countries. The project was also not as successful in creating a strong network of institutions; communication and sharing of experiences and analysis results among POs outside of the global conferences could have been much more extensive.

Program Approach

The program approach was generally effective and appropriate to obtain the desired objectives. The majority of POs felt the project was mostly suited to their needs and the training workshops, mentoring program, and peer-to-peer reviews were good mechanisms for building capacity although the execution of the mentoring program faces some mixed reviews.

Equity

While the direct beneficiaries of the project were the participating policy research organizations, one of the objectives of their research was to conduct a benefit incidence analysis and determine whether vulnerable groups were benefitting equally from government programs. In addition, some of the specific topics chosen by POs for their cost-effectiveness analyses and policy simulations target improving services for vulnerable populations. While it is difficult, not to mention too early, to conclude whether these analyses had a direct impact on disadvantaged groups, some initial impact on policies were reported.

Value for Money

The project represents value for money, especially with respect to economy and efficiency, and to limited extent effectiveness. In the evaluator's opinion, the benefits have outweighed the costs. With respect to effectiveness, shortcomings of the project were a lack of focus and support to POs on communication and dissemination of their results from the beginning of the project and sustained engagement with the policy community from project inception to ensure the uptake and direct use of research results in policy.

Sustainability

Based on our analysis of the capabilities of POs in the project impact and effectiveness sections of the report, there is evidence that POs have acquired sufficient skill levels in public expenditure analysis and dissemination of results to continue this type of research without the project technical support .

Retention of the staff with these skills by POs is a key issue for sustainability. Attrition during the project was high; only four organizations retained their original project teams, yet POs have managed to improve their capabilities based on our analysis. A promising sign is that all current project staff, with the exception of one PO, intend to remain at the POs beyond the end of the project.

All PO organizations indicated that they would like to continue research on transparency and accountability issues although four Asian POs and two Latin American POs mentioned that funding was a limiting factor for this continued research.

Although we determined that the creation of a network of POs involved in the project was only partially successful in our analysis in Section 3.5, all POs indicated they thought they would contact other POs in the project after project end, signally that the network would not disappear.

Replicability

In the evaluator's opinion, the program can and should be replicated if all the innovative aspects of the project are maintained (as outlined in the innovation section), continuous support and feedback is provided to participants for the analytic work, the five-year project duration is maintained since it takes significant time to master the analysis techniques and disseminate results.

Innovation

What is innovative about the project is that it employs the following five elements together effectively: a solid conceptual framework supported by rigorous empirical evidence, recruitment of organizations with a clear interest in the approach, a highly structured "learning by doing" approach wherein the techniques learned are immediately applied, using a "constructive engagement" approach to effectively communicate results to policymakers, and a peer review and learning element.

Acknowledgements

Many people have generously contributed to establishing a set of indicators that can be employed to determine if the capacity building paradigm being employed by this project is effective. First and foremost are the leaders and staff at the policy research organizations participating in the project. Unquestionably, the data requirements for this evaluation are significantly greater than those for most similar projects. Our partners have been wonderful overall in responding to the evaluation data needs. We sincerely appreciate their hard work.

We are also indebted to our colleagues at the Global Development Network and the Results for Development Institute for their guidance and support. In particular, we thank Savi Mull, Ramona Angelescu Naqvi, Pooja Sarin, Kaushik Ganguly, and Vasundhra Thakur from the Global Development Network and Courtney Tolmie and Courtney Heck from the Results for Development Institute for all they have contributed.

We would like to thank and acknowledge our colleague Ray Struyk who is fully responsible for the design of this evaluation and all reports through the mid-term evaluation. Dr. Struyk retired from NORC prior to completion of this evaluation.

1. Short Introduction to the Program

The Strengthening Institutions to Improve Public Expenditure Accountability project (hereafter the "project") aimed to improve development outcomes by increasing the effectiveness with which governments allocate and use their resources. It was to strengthen analytical underpinnings of the policy debates around public expenditure priorities and their impact, thus improving the governance of public service delivery in 15 countries (Table 1.4). Through this approach the project intended to achieve four key outcomes:

- Expanded institutional and individual capacity for public expenditure monitoring and analysis, development of policy alternatives, and constructive engagement in a peer learning environment
- Increased use of evidence-based policy reforms in social services and infrastructure
- Internationally comparable information on public expenditures, incidence (who benefits), effectiveness, and policy alternatives that will begin to build benchmarks for the quality of public spending
- Creation of a strong network of institutions to share training materials, templates for analysis and communication.

On the analytic side, participating organizations (POs) were to, over the five-year project life, perform four distinct but related budget analysis activities in the health, education, and water services sectors—program budgeting analysis; cost effectiveness analysis; benefit incidence analysis; and, development of evidence-based policy options for improving the results of public expenditures. The project implementation team provided capacity building in the corresponding analytic methods and communications/dissemination through workshops, mentoring through technical advisors, peer-engagement and learning, and the development of a resource-rich, easy to access website that is open to POs and others. This was a "learning by doing" approach wherein POs were able to implement the analysis and communications techniques and approaches learned in the series of analyses and reports required by the project while receiving technical support from advisors and the project management team.

Over the five years of the project, numerous capacity building workshops were held and dissemination events were organized. These are listed in Tables 1.2 and 1.3.

Table 1.2: Capacity Building Workshops

Type of activity	No. of events	Date(s)	Location(s)	Partner Attendees	Partners who did not attend
PEM Launch Meeting	1	Dec-08	New Delhi, India	Ghana, India, Kenya and Peru	Additional partners were not selected by then
Global Workshops	5	May 2009, January 2010, February-March 2011, April 2012, April 2013	Washington DC, Prague, Bangkok, Istanbul, Jakarta	Global	Bangladesh due to visa issues (May 2009), and Ghana due to visa issues (Feb-Mar 2011)
Regional Workshops	11	July 2009, April 2010, June 2010, June 2010, July 2010, July 2010, July 2011, Oct 2011, July - August 2012, September 2012, December 2012	New Delhi, Dhaka, New Delhi, Mombasa, Buenos Aires, Buenos Aires, Nairobi, New Delhi, New Delhi, Buenos Aires, Nairobi	Regional	Regional workshop-only partners from the region invited; Ghana due to elections (December 2012)

Table 1.3: Dissemination Events

Title of Event	Date	Location	Partner Attendees
Presentations at the African Economic Research Consortium (AERC) Bi-Annual Workshop	4 June, 2010	Mombasa, Kenya	Ghana, Kenya, Nigeria, and Uganda
Presentation at the Economics Education and Research Consortium (EERC) Workshop	26-28 June, 2010	L'viv, Ukraine	Armenia
Presentations at the East Asian Development Network (EADN) Annual Forum	3-4 August, 2010	Bangkok, Thailand	Indonesia and Philippines
Presentation at the International Society for Equity and Health (ISEqH) Conference	26 - 28 September, 2011	Cartagena, Colombia	Guatemala, Peru and Philippines
Policy Dialogue: Regional Policy Seminar on Effective Public Service Delivery in Health and Education	12 October 2011	New Delhi, India	Armenia, Bangladesh, India, Indonesia, Nepal and Philippines

Title of Event	Date	Location	Partner Attendees
Effective Public Service Delivery in Education, Health and Water: Evidence from GDN's Study in India, Nepal and Bangladesh	21-22 December 2011	Dhaka, Bangladesh	Bangladesh, India and Nepal
Governance and Public Service Delivery in Education, Health and Water	08-12 July, 2012	Madrid, Spain	Guatemala
Policy Dialogue: Promoting Effective Public Expenditure and Service Delivery in Health, Education and Water	9 April, 2013	Jakarta, Indonesia	Armenia, Argentina, Bangladesh, Ghana, Guatemala, India, Indonesia, Kenya, Mexico, Nepal, Nigeria, Philippines, Peru and Uganda
Policy Dialogue: Viabilidad de las Finanzas Públicas en América Latina	25 September, 2013	Mexico City, Mexico	Argentina, Guatemala, Mexico, and Peru
Policy Dialogue: Making Smarter Policies - Improving Health and Education Outcomes in Africa	4 November, 2013	Abuja, Nigeria	Ghana, Nigeria, Uganda and Kenya

The project was focused at the institution level and it was expected that POs would master new skills and acquire valuable human capital that would be of service to their organizations as well as equip them for future work in the area of public expenditure monitoring. The project also aimed to deliver timely analyses effectively into the policy process with the explicit goal of improving public services and improving peoples' lives.

Table 1.4: Participating Organizations

Organization Name	Abbreviation	Country
Advanced Social Technologies	AST	Armenia
Centre for Budget and Policy Studies	CBPS	India
Center for Economics and Development Studies, Faculty of Economics, Padjadjaran University	CEDS	Indonesia
Research Center of the University of the Pacific	CIUP	Peru
Center for the Implementation of Public Policies Promoting Equity and Growth	CIPPEC	Argentina
Center for Research and Communication	CRC	Philippines

Organization Name	Abbreviation	Country
Center for the Study of the Economies of Africa	CSEA	Nigeria
Graduate School of Public Administration and Public Policy, Tecnológico de Monterrey	EGAP	Mexico
Economic Policy Research Centre	EPRC	Uganda
Economic and Social Research Foundation	ESRF	Tanzania
Fundación para el Desarrollo de Guatemala	FUNDESA	Guatemala
Institute of Economic Affairs	IEA	Kenya
Integrated Social Development Centre	ISODEC	Ghana
Policy Research and Development	PRAD	Nepal
Unnayan Shamannay	US	Bangladesh

Using the knowledge gained from the project, GDN is expecting to produce four knowledge products: a course module on public expenditure analysis, a white paper on cost-effectiveness analysis versus cost-benefit analysis, a data portal of public expenditure data with an interactive sharing program from the project participating countries, and a set of best practices and lessons learned for creating and disseminating research results in order to impact policy.

The project included a robust monitoring and evaluation (M&E) program to assess progress made in reaching the project's four projected outcomes listed above.

Eleven of the 15[2] policy research organizations from as many countries were selected through a competitive process for project participation. Over 200 organizations applied. The remaining four were pre-selected at an earlier date and were included in the project proposal. Over the course of the project period over 100 staff at the participating organizations were involved in the project.

The project was led by the Global Development Network (GDN). GDN engaged the Results for Development Institute (R4D) to coordinate the development and delivery of the majority of the technical content on public expenditure analysis to the participant organizations.

[2] One of the 15 organizations selected, ESRF in Tanzania, was dropped from the program in 2012 due to declining quality in deliverables, failure to sufficiently address feedback provided on final reports, and significant delays in submission. The rationale is further discussed in the Efficiency section of this report.

2. Evaluation Methodology

The final evaluation was organized in effect as two evaluations: an impact evaluation which is part of the Strengthening Institutions to Improve Public Expenditure Accountability Project's M&E program and an implementation evaluation that was organized specifically at only two points in time - once for the mid-term review and again at the end of the project. The final evaluation covers the implementation period of the project from the first training workshop in mid-2009 until September 2013 whereas all project activities were completed by January 15, 2014[3]. It is important to note that this evaluation does not cover the full impact of the project because the DFID deadlines required a complete final evaluation at project end. This required data collection prior to the end of activities in order to allow enough time for analysis and report writing. Moreover, some of the higher level impacts of the dissemination of the last reports produced by the project are expected to occur beyond the formal end of project activities. The description below first outlines the impact evaluation and then turns to the implementation evaluation. The impact evaluation provides hard data on the project's outcomes while the implementation evaluation addresses the efficacy and efficiency of its operations.

Impact Evaluation

The fundamental objective is to determine if the capacity building model employed in this project is successful. To make this determination requires that certain participating organizations' activities and practices be monitored over time and that outcome indicators be defined and measured for at least two points in time. Over the five-year project period, baseline data was collected in mid-2009, a mid-term evaluation conducted in 2010-2011, and a final evaluation has been prepared in 2013 on the project's success in achieving its goals. In addition, activity monitoring was conducted in 2010 and 2012 through surveys of participating organizations.

In light of the paucity of rigorous evaluations of programs designed to strengthen the capacity of policy research and advocacy organizations, GDN sees this evaluation as providing high payoff for the donor community. As the importance of strong governance becomes better documented and the role of policy research organizations in helping governments improve their performance is more appreciated, more and

[3] The data for the final evaluation was collected between September and November 2013. This is prior to completion of all the project activities and was necessary in order to complete the final evaluation report by the deadline specified by GDN.

larger capacity building programs are being launched. Yet there are few evaluation results from earlier efforts to inform these new undertakings.[4]

One can anticipate the project having five types of impact on organizations under the first and second project outcomes listed at the beginning of the Short Introduction to the Program section:

1. Improved skills and analytic rigor in conducting public expenditure monitoring projects,[5]
2. Some transference of these skills to an organization's other projects,
3. Increased incidence of shaping policy debates and evidence-based policy making due to improved communications and dissemination practices,
4. Increased effectiveness in shaping policy debates in the organization's focus areas, and
5. Increased grantee interest in public expenditure accountability projects, i.e. organizations find work in this area rewarding and devote more resources to it.

Because items 1 and 2 are so closely related and because some participants may not have previously carried out accountability projects, we examine evidence of an overall quality increase in analytic work. All five impacts are being measured by the evaluation.

Structure. The evaluation employs a before-and-after (reflexive) design, with comprehensive data obtained by the 2009 baseline survey and at two future points (2010 and 2013) over the span of the five-year project. The first follow-up round occurred about mid-way through the project in late fall-early winter 2010 and the second between September and November 2013.

Indicators and data collection. Below are statements explaining the basic measurement strategy for the objectives of improving analytic skills, increasing the incidence of evidence-based policy making, and organizations' perceived effectiveness in shaping policy debates.

---

[4] The evaluation structure is informed by a recent review of past evaluations of projects aimed at strengthening the capacity of think tanks , which also laid out a conceptual framework for technically stronger future evaluations. R. Struyk, *Options for Supporting Independent Monitoring Organizations* (Washington, DC: NORC, paper prepared for the Brookings Institution Transparency and Accountability Project, 2008, processed).

[5]Accountability in this context is the acknowledgement and assumption of responsibility for actions, products, decisions and policies, including administration, governance and implementation within the scope or role of a particular position, be it civil servants, politicians or members of parliament. Hence, the project is supporting adjustments to policies and procedures that make it easier to identify what actions have been taken and who is responsible for them.

Transparency can be understood as the availability and accessibility of relevant information about the functioning of government policy making, procedures, and program implementation. Transparency as a concept covers event transparency (open information about inputs, outputs, and outcomes), process transparency (open information about transformations that take place between inputs, outputs and outcomes), real-time transparency (information released immediately) and retrospective transparency (information available only after time passes). C. Hood, "What Happens when Transparency Meets Blame-avoidance?" *Public Management Review*, vol.9, no.2, 2007, pp.191-210.

- *Improved analytic rigor—review of organizations' analytic reports.* A critical review of a sample of analytic policy research reports was conducted at baseline to ascertain the degree to which they meet usual standards for such work. Participating organizations were asked to provide for review six reports produced before the project's initiation. Six more reports produced from the beginning of the grant period to the time of the second evaluation wave were reviewed for the mid-term evaluation. Reviewers were experienced, policy-oriented social scientists.

At baseline a total 83 policy research reports were assessed by 14 external / independent reviewers. Nineteen of the reports were in Spanish and the rest in English. At mid-term an additional 80 reports were reviewed which included 26 produced by POs for this project. At endline 77 reports were reviewed with included 41 produced by POs for this project. Both project and non-project reports were reviewed for every organization.

- *Usefulness of organization-provided information on accountability and other issues* to national and local decision makers and practitioners and contribution to evidence-based policy making.[6] The policy community survey (PCS), carried out to gather the necessary information, had two objectives: to measure (a) the perceived extent of evidence-based policy making and the role of organization- and other NGO-provided information in this process; and, (b) the extent to which a grantee organization's work, publications, events, and policy recommendations are viewed as valuable sources of information.

Respondents were identified by POs and comprised of (a) persons who occupy official positions at the national and local level with official responsibilities in the sector(s) where the organization is active (e.g., education and health); (b) leaders of NGOs in the topic areas on which they have been conducting accountability activities; and (c) other relevant professionals, such as journalists, educators, business leaders, and local members of the international community actively engaged in policy issues.

In July 2009 NORC sent a questionnaire for completion, either online through an eSurvey, via email or by the PO delivering a hard-copy questionnaire to targeted respondents. A similar mailing was made in November 2010 and September 2013. The method employed depended on the advice of the organization for each respondent on the approach more likely to result in questionnaire completion. NORC did initial follow-up with respondents via email but asked the PO to telephone and email as well if necessary. When response rates to the eSurvey in particular turned out to be lower than expected, the project offered those

[6] The survey broadly follows the procedures used by John Hird in his study of the effectiveness U.S. non partisan policy research centers sponsored by state legislatures in which he surveyed state legislators by mail about effectiveness. J. Hird, "Policy Analysis for What? The Effectiveness of Nonpartisan Policy Research Organizations," *The Policy Studies Journal*, vol. 33, no.1, 2005, pp. 83-105.

POs a small incentive fee to additional assistance in contacting respondents and encouraging them to complete the survey. This increased the number of responses, although most organizations said they were not interested in the fee.

Altogether, 170 responses were received at baseline and of these 169 were usable in the analysis. In the second round 218 usable responses were received. At the endline, 234 responses were received, of which 231 were usable.

Table 2.1: PCS Survey Response Rates

Country	Surveys Sent (a)	Surveys Completed	Percent Completed	Panel Respondents				
				2013 only	2009, 2011, and 2013	2011 and 2013	2009 and 2013	Unknown
Argentina	45	15	33%	47%	7%	7%	0%	40%
Armenia	20	14	70%	43%	7%	14%	0%	36%
Bangladesh	43	15	35%	7%	7%	7%	0%	80%
Ghana	34	12	35%	33%	17%	17%	0%	33%
Guatemala	137	61	45%	61%	3%	3%	2%	31%
India	40	24	60%	33%	4%	4%	4%	54%
Indonesia	35	16	46%	25%	6%	13%	6%	50%
Kenya	38	11	29%	9%	27%	9%	0%	55%
Mexico	30	7	23%	57%	0%	0%	0%	43%
Nepal	22	9	41%	22%	11%	11%	0%	56%
Nigeria	40	8	20%	25%	25%	0%	0%	50%
Peru	63	19	30%	37%	0%	11%	5%	47%
Philippines	33	5	15%	20%	20%	0%	0%	60%
Uganda	35	18	51%	44%	6%	22%	0%	28%
Total	**615**	**234**	**38%**	**39%**	**7%**	**8%**	**2%**	**44%**

(a) This is the total number of surveys sent by mail and email, minus emails that bounced back due to bad addresses.

To understand whether the PCS respondents reflect the views of those targeted by the project, in Tables 2.2 and 2.3 we compare the target audience of the project and the occupations of the PCS respondents, respectively. From these two tables, one can see that generally the respondents in the PCS are the targets of the project produced reports--government officials, staff at NGOs, and academics. Staff at think tanks, members of parliament, and members of the media, notably, are targets of project products but are underrepresented in the pool of PCS respondents.

Table 2.2: Audiences targeted in dissemination of project products

	Percent of POs	Argentina	Armenia	Bangladesh	Ghana	Guatemala	India	Indonesia	Kenya	Mexico	Nepal	Nigeria	Peru	Philippines	Uganda
National senior government officials (ministers, deputy ministers, or equivalent)	93	■	■	■	■	■		■	■	■	■	■	■	■	■
Think tanks	86	■	■	■	■	■	■	■	■		■	■	■		■
Advisories to senior officials	79	■	■		■	■	■	■		■	■		■	■	■
National level non-elected bureaucrats (civil servant, technical staff)	79	■		■		■	■	■		■	■	■	■	■	■
Advocacy NGOs	79		■	■	■	■	■	■	■	■	■			■	■
Academics	79	■		■	■	■	■	■		■	■	■		■	■
Members of parliament	64	■	■	■	■	■				■	■			■	■
Reporters/editors	57	■		■		■			■	■	■		■		■
Sub-national senior government officials (minister, deputy minister, or equivalent to minister or deputy minister in regional or local government)	50			■	■			■				■	■	■	■
Sub-national level non-elected bureaucrats (civil servant, technical staff)	50	■					■	■				■	■	■	■
International aid organizations	43		■			■		■	■		■	■			
Other types of NGOs	36			■	■			■		■	■				
Private consultants	29					■	■				■		■		
Other positions in the media	21									■	■	■			
Other government officials	14						■							■	
Other	0														
Number of Target Audiences Selected	**8.6***	**8**	**6**	**9**	**8**	**10**	**8**	**10**	**5**	**9**	**12**	**8**	**8**	**9**	**10**

**Mean of all countries.*

Table 2.3: Occupations of PCS Respondents

	2009	2010-11	2013
Which area and position best characterize your current professional life?			
Government official	23%	31%	27%
Member of parliament	3%	6%	7%
Leader or staff at an advocacy NGO	14%	13%	10%
Leader or staff at another type of NGO	18%	7%	15%
Leader or staff at a think tank	8%	5%	4%
Leader or staff at an international aid organization	0%	11%	7%
Private consultant	0%	6%	7%
Reporter/editor	2%	2%	4%
Other position in the media	1%	1%	0%
Academic	10%	11%	8%
Other	21%	8%	10%
If government official: Can you further define your position relative to the government?			
National senior government official (minister, deputy minister, or equivalent in regional or local government)	18%	27%	17%
Sub-national senior government official (minister, deputy minister, or equivalent to minister or deputy minister in regional or local government)	1%	9%	12%
Advisory to senior official	24%	36%	10%
National level non-elected bureaucrat (civil servant, technical staff)	--	--	39%
Sub-national level non-elected bureaucrat (civil servant, technical staff)	--	--	10%
Other	57%	28%	12%

- *Tracking organizations' volume of accountability analyses.* Information on the share of projects and resources devoted to such projects was gathered in the baseline survey and in the 2010 and 2013 monitoring surveys via a PO survey. Questions in this area were changed between the baseline and monitoring survey to improve response accuracy. The 2010 responses now serve as the project's baseline. The PO survey also examined the effect of the project on POs capacity as self-reported by the POs.

Lastly, note that NORC also gathered information for each country on the relevant trends in governance and other areas. The list of initial information sources considered includes indices such as Freedom House's Freedom of the Press report, the World Bank's Worldwide Governance Indicators, the Open Budget Index, and the World Bank's Doing Business Report among others. These country-level developments could be very important for interpreting the changes in the impact measures just discussed.

Causality

It is important to note that we cannot attribute direct causality to the project for the changes observed in the outcomes calculated using the survey data. Instead, we use anecdotal evidence and the opinions of respondents regarding the impact of the project on perceived policy changes and research and communications capabilities of respondents.

It is also essential to remember that POs may have also received assistance from other organizations or programs over the monitoring period. We asked POs about this in the 2013 survey, and it was indeed the case. Half of the POs received technical assistance related to research or advocacy from sources other than GDN. In several cases the assistance covered similar topics as GDN (for example, PETS, budget tracking, public expenditure analysis). Furthermore, 62 percent of the POs received funding for research or advocacy work in transparency and accountability. Donor organizations included the International Budget Partnership (through the Inter-American Development Bank), United States Agency for International Development (USAID) and the World Bank. These additional forms of assistance should be considered when interpreting the results of this evaluation.

Implementation Evaluation

Implementation evaluations focus on how well a program operates: Is it efficiently managed? Are its activities well-structured and focused? Does it avoid wasteful expenditures? Are sufficient but not excessive resources provided to participants? This evaluation was organized to address the points set forth in GDN's TOR to the Evaluator, which substantially mirrors DFID's guidance for the final evaluation.[7]

The method followed was straightforward. Relevant project documents and reports were read and analyzed against the points raised in the guidance. These documents, i.e., individual documents or classes of documents, are listed in Table 2.4. Beyond those listed in the table, 163 policy research reports produced by the participating organizations have been critically reviewed for the impact evaluation, as noted above.

Project implementers were interviewed on project development and implementation and the leaders and key staff of all fourteen participant organizations were interviewed on the "services" they received, policy impacts they anticipated from their project-related work, and the quality of project administration from

[7] The Consultant's TOR is in Annex 2.

their perspective. A separate interview guide was developed for each interviewee group. Those interviewed are listed in Table 2.5.

Table 2.4: Documents Reviewed

Source	Document or Document Type[a]
DFID	Guidance on Commissioning a Final Review and Final Evaluation Final Evaluation Guidelines Additional Guidance for FERs and PCRs Making Governance Work for the Poor Measuring Change and Results in Voice and Accountability Work GTF Learning and Impact Strategy
GDN and R4D	Budget documents: proposal budget, cost-to-date distributions prepared at evaluator's request Agendas for Project Workshops 2013 Annual Report Annex A2 – Approved Program Logframe Annex G1 – Activity to Outcome Annex G2 – Policy Outcomes and Funding Opportunities Stage 1 – Most Significant Results List

a. These are "short titles" to convey the substance of the report. Full titles are available if needed.

Table 2.5: Persons Interviewed

Category	Organization
Project Management	R4D GDN
Participating Organizations	AST Armenia CBPS India CEDS Indonesia CIPPEC Argentina CIUP Peru CRC Philippines CSEA Nigeria EGAP Mexico EPRC Uganda FUNDESA Guatemala IEA Kenya ISODEC Ghana PRAD Nepal US Bangladesh

Key management and research staff from the organizations listed above were interviewed for the final evaluation. The information gathered from these various sources was analyzed to address the various points in the TOR given to the consultant by GDN. The names of persons interviewed are not provided to maintain confidentiality.

3. Findings in Relation to Standard Review Criteria

3.1 Relevance

Good governance requires not only top-down ‘supply side’ improvements such as policy and institutional reforms by the government, but also sustained domestic demand for good governance from civil society. This project is designed to help realize that principle by strengthening the demand side - the mechanisms by which the public can hold governments accountable in developing countries. It does this by focusing on developing the capabilities of policy research organizations (PROs) to conduct research in tracking budget expenditures, measuring the quality of service delivery, and proposing policy or administrative changes.

In this section we demonstrate that the project is relevant to increasing voice, accountability, and responsiveness. The countries chosen are ones in which voice and accountability were limited, as evidenced in Table 3.1.1 where none of the countries are above the 60th percentile and most are below the 50th percentile on the Voice and Accountability indicator of the Worldwide Governance Indicators from 2009, the year the project started. Yet , for this type of project to succeed, there must be a minimum level of freedom for civil society to operate without suppression and there must be access to budget information for an organization to analyze. We start by examining the civil society and policy environments in the selected countries and then look at the specific selection criteria for participating organizations within these countries to determine whether the choice of country and PRO were relevant to the objectives of the project.

Table 3.1.1: Worldwide Governance Indicators: Voice and Accountability 2009[8]

	Percentile rank	Score (-2.5 to 2.5)
Argentina	55.9	0.25
Armenia	25.6	-0.82
Bangladesh	35.1	-0.37
Ghana	60.7	0.50
Guatemala	37.0	-0.33
India	60.2	0.47
Indonesia	48.3	-0.05
Kenya	37.4	-0.32
Mexico	53.6	0.13
Nepal	30.8	-0.58
Nigeria	24.2	-0.85
Peru	50.2	0.04
Philippines	45.5	-0.12
Tanzania	43.6	-0.14
Uganda	33.2	-0.49

Operating environment

The operating environment in the project countries allowed enough space for PROs to operate while also presenting an opportunity to increase the use of evidence (data, research, and analysis) in policy making. At the start of the project, there was moderate use of evidence from research in policymaking, as shown in the data from the baseline survey in 2009 in Table 3.1.2. Looking at frequency of use of data, research, or analysis, it is of note that no country scores a three "this is standard procedure now," most government ministries are closer to two "this happens sometimes but is rather exceptional," and most parliaments are below two. There was little significant change in these ratings over the period of the project with the exception of India that saw an improvement in 2003 and Peru where the government rating improved significantly the last year but the parliament fell to a very low rating of almost no use at all.

Similarly, there was a moderate level of government and parliament willingness to accept input from thinks tanks or NGOs at the start of the project. Most scores are just about in the middle, a rating of around 2.5; of note is that African countries scored higher than other regions in this respect (they all score

[8] The Worldwide Governance Indicators consist of an aggregation of data sources drawn from a diverse variety of survey institutes, think tanks, non-governmental organizations, and international organizations. These sources are aggregated to provide a rating of -2.5 to 2.5 and countries' percentile ranking is provided for easy comparison. The Voice and Accountability indicator measures the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.

above 2.5 with the exception of the Ugandan parliament). Over time about half the countries signaled more openness to receiving input, showing some level of statistically significant positive change. The exception here is Argentina where government officials seem to have reverted back to baseline levels of openness after a relatively large increase at the midterm.

There also appears to be adequate access to budget information at the start of the project, with the exception of Nigeria. We assess the level of access to budget information by using the Open Budget Index[9]. Table 3.1.3 presents the scores for data collected from June to September 2009, at the start of the project, and in August to December 2011, the most recent data available. All countries show a score of 45 out of 100 or above. Four countries show increases of over five points from 2009 to 2011: Bangladesh, Indonesia, Mexico, and Uganda. However, three countries show a decline of over five points: Argentina, Peru, and the Philippines.

[9] The Open Budget Index measures public access to budget information. Civil society researchers rate countries on 92 questions; 58 of these questions focus on the public availability and comprehensiveness of the Executive's Budget Proposal. Each of the answers to the questions is given a score from 0-100 and all the answers are averaged to produce the final rating. The index was first published in 2006, and again in 2008, 2010, and 2012. More information can be found at http://internationalbudget.org/what-we-do/open-budget-survey/

Table 3.1.2: Policy Community's View on Government Agencies' and MPs' Use of Data and Analysis and the Willingness to Accept Input from Think Tanks by Country

	Frequency of Use of Data/Research/Analysis (a)						Willingness to Accept Input from Think Tanks and NGOs (b)					
	Govt Ministries			Parliament			Govt Officials			Parliament Members		
	2009	2010-11 (c)	2013 (c)	2009	2010-11 (c)	2013 (c)	2009	2010-11 (c)	2013 (c)	2009	2010-11 (c)	2013 (c)
Argentina	1.60	2.33	1.58	1.75	1.60	1.77	1.60	3.00***	1.64***	3.00	3.14	2.64
Armenia	2.00	2.00	1.92	1.50	1.73	1.73	2.25	3.44***	2.92	1.92	2.50	2.36
Bangladesh	1.73	2.13	1.85	1.10	1.79*	1.77	2.36	2.13	2.75**	2.00	2.45	2.58
Ghana	2.29	2.00	2.09	1.64	2.10	2.00	2.73	2.63	3.00	3.13	2.91	2.90
Guatemala	1.91	1.82	1.85	0.91	1.22	1.45	2.46	2.35	2.34	2.31	2.11	2.40*
India	1.50	1.63	2.24*	1.75	1.44	2.28**	2.45	2.38	2.95**	2.00	2.07	2.53
Indonesia	2.00	2.25	2.53	1.36	1.53	1.77	3.11	3.00	2.86	2.33	2.64	2.69
Kenya	2.33	2.00	2.11	1.89	2.00	2.11	2.83	3.29	3.30	2.89	3.50*	3.20
Mexico	2.20	2.14	1.00	1.40	1.67	1.67	2.20	2.73*	2.75	2.40	2.67	2.67
Nepal	2.00	2.20	2.33	1.67	1.20	1.86	2.33	3.00	2.71	2.44	2.60	2.29
Nigeria	1.92	2.00	2.50	1.78	1.86	1.83	2.91	2.57	3.17	3.00	2.71	3.17
Peru	1.71	1.78	2.27**	1.37	1.35	0.50***	2.55	2.47	2.60	2.24	2.13	2.00
Philippines	2.63	2.67	2.75	2.17	2.67	2.00	2.71	3.33	3.25	2.43	3.33	2.75
Uganda	2.75	2.35	2.40	2.00	1.93	2.25	3.25	2.71	2.82	2.33	2.81	3.29*

(a) Four point scale from 0 to 3, where 0= Essentially never; 1 = This practice is still quite rare with us; 2 = This happens sometimes but is rather exceptional; and 3 = This is standard procedure now.

(b) Four point scale from 1 to 4, where 1 = not willing at all and 4 = very willing.

(c) T-tests have been performed to test whether the average score was significantly different from the average of the previous survey round. * = P<0.10, ** = P<0.05, and *** = P<0.01

Table 3.1.3: Open Budget Index 2009 and 2011

	Score (0 -100; 100 the highest)	
	2009	**2011**
Argentina	56	50
Armenia	Not available	
Bangladesh	48	58
Ghana	54	50
Guatemala	50	51
India	67	68
Indonesia	51	62
Kenya	49	49
Mexico	52	61
Nepal	45	44
Nigeria	18	16
Peru	65	57
Philippines	55	48
Tanzania	45	47
Uganda	55	65

Selection Criteria for Participating Organizations. The selection criteria also helped ensure that the environment in which POs work is at least somewhat conducive to project success. To be eligible, applicants had to be from countries ranked as politically free or partially free by the Freedom House index. Selection criteria, shown in Table 3.1.4, included several points that also spoke to the transparency-accountability environment: whether the applicant had the likely ability to access necessary and quality data for the study proposed; whether the applicant had a plan to overcome country governance/transparency obstacles; and the value of previous public expenditure management studies in the applicant's country and the applicant's apparent knowledge and understanding of the studies. Finally, selection was dependent upon the feedback received on shortlisted applicants from both GDN's Regional Network Partners and DFID local country offices or British High Commissions.

Overall, the selection criteria and the way in which applicant organizations were rated were thorough and thoughtfully designed to select POs with a strong likelihood of contributing to and benefitting from the project.

Table 3.1.4: Criteria for Selecting Partner Organizations

No.	Criteria	Indicator
1.	Quality of Application	Quality of proposed methodology Quality of proposed initial work plan Well-defined project budget in line with project scope How well project fits into organization's institutional strategy
2.	Capacity of Applying Organization	Level (quantity and quality) of past experience in conducting quantitative policy analysis Level (quantity and quality) of past experience working with policymakers Presence and expertise of analytical staff (strength) Perceived ability to access necessary and quality data to complete study Organization's potential for expansion in the research area
3.	Feasibility and Value of Proposed Study	Quality of plan to overcome country governance/transparency obstacles for country's below a certain score Demonstrated understanding of the national budget processes Value of previous studies on PEM in the applicant-specific country
4.	Proposed Dissemination Activities	Quality and likely impact of communication and dissemination plan Responses of policymakers to applicant's past research/dissemination Presence of communications expertise on staff

While it seems that countries and POs were appropriately selected, there were several risks in initiating a project with a PEM focus, given the uncertain context in some of the partner countries. Attempts were made by the project implementation and management team to deal with these issues; but risks external to the project could only be mitigated to a limited extent. These issues included:

- Location of a few partners in politically unstable countries like Nepal and Bangladesh which sometimes restricted the extent of constructive engagement with the policy community to improve accountability.
- Poor governance arrangements, resistance to change, and general difficultly in assigning private property rights to water services burdened the management of water systems, and made policy reform in this sector controversial and challenging in some countries.
- In some countries, lack of political commitment limited some public officials' willingness to interact with or listen to recommendations by partners supported by the program.
- Since the project had a five-year duration, election cycles, turnover in political office and new public officials posed problems of continuity of contact with government departments for policy traction.

- In low and middle income countries, good high quality information and accurate data on public expenditures, usage and benefits was difficult to access and published with considerable lags. At times data was available in formats not amenable to policy or academic research on social issues.

Despite most of the above issues being external to the project and beyond the control of management and implementation two issues were resolved to some extent. To overcome data problems, partners used either secondary data or initiated their own independent surveys to obtain relevant information. To evoke interest from the policy community, creative communication strategies, citizens' forums, research competitions, and social media were used by the partners.

Governance priorities

The project was closely aligned with national or local governance priorities as is demonstrated by the project approach and the approach of the POs in selecting their research topics. The project approach was that all POs were to work in the health, education, and water sectors of their countries – key sectors of any country. They were to utilize research and analysis tools taught to them by the project to complete a program budgeting analysis (PBA) and benefit incidence analysis (BIA) at the general sector level before choosing a specific topic within each of the sectors for which to conduct a cost-effectiveness analysis (CEA) and then a "policy simulation" (PS) where-in they presented policy recommendations on the topic.

Half of the POs chose their sector topics primarily based on government priorities, the other POs considered other factors first such as data availability or researcher interest. Only two POs did not consider government priorities at all in the selection of some of their sector topics. An African PO met with policymakers to determine which topics were of the greatest policy significance while a Latin American PO was already operating according to government priorities, which determine their institutional objectives. Another Latin American PO chose a topic because it was widely debated in public discourse and was related to a segment of the population that had been neglected by policymakers. An Asian PO reviewed recent legislation to determine which topics were relevant to policy discourse.

- *"One of the requirements of GDN [is] to be able to make a difference in terms of politics, so I wanted to consider questions in terms of research but also that has impact on policy." (Latin America)*
- *"The other issue is we did consult within the policy community of what were the ongoing activities. This was mainly through the Department of Planning and Education and Health, that's where we mainly made these consultations. And these helped us to tell us these are the three or four things that are important now. Then we say let's choose this. Is the data there? No. Then let's look at the second option." (Africa)*

- *"Our country stresses education, health and water provision for dealing with poverty; to that extent, the requirements of the project, particularly the sectors, are consistent with the priorities of the national plan… The choice of the topic was based on the challenges that we face in the country – it's an issue that is being addressed at the national level. There's also the issue of interest of the researcher, but the main one was the country priorities." (Africa)*
- *"In the education sector we went to the website of the Department of Education and went to the website of Congress; they have a list of laws on education. We look at what are the pressing issues tackled by Department of Education and Congress; we try to bring down to three to five. We then met with the team to decide which are the more pressing issues." (Asia)*

Two other POs chose their topic based on the results of the PBA and BIA analysis that revealed an important issue.

- *"For the policy simulation for water sector initially we were thinking about difference in access to water. But when we saw that benefit incidence analysis after the first year proved that access to water is more or less equal, almost ideal physical access, but when we look at affordability we had to change our whole perspective on research. We understood that we must not do research on access but on affordability. The same idea came up after the very first meeting with the Water Committee. They said they are not that interested in our research on access, because we see more issue in tariff formulation. It was a two-stage approach." (Asia)*

The project is also aligned with DFID's Governance and Transparency Fund whose Theory of Change is modelled on the Capability, Accountability and Responsiveness (CAR) framework. The project fits most directly into the accountability element of the framework, which is defined as "the ability of citizens to hold leaders, governments and public organizations to account"[10]. The DFID white paper on Measuring Change and Results in Voice and Accountability highlights that interventions that strengthen civil society should consider the importance of a transition from voice to accountability, i.e. effectively engaging with the government to secure rights or transform relationships.[11] The project focuses on doing just that by not only strengthening the capabilities of POs to have a strong evidence-backed voice on key policy issues but by also helping them constructively engage with policymakers to use the evidence they generate to improve policies and secure the rights of citizens. The project also addresses the capability and responsiveness elements through its program logic that project financial and technical support to POs to develop strong and relevant research in the health, education, and water sectors will lead to an increased

[10] DFID, 2009. *Measuring Change and Results in Voice and Accountability Work*, London, DFID, p4.
[11] Ibid.

incidence of evidence-based policy (improving the capabilities of the government in making sound policies) that result in a government that is more responsive to the needs of its citizens.

3.2 Impact

This section measures the impact of the project by first giving an overview of overall project impact on the PO institutions capabilities and effectiveness in the policy community and then focuses on the impact of the research produced specifically for the project, achievement of individual PO objectives, and the effect of the project direct beneficiaries. We do this by drawing on the results of the external review of research reports, the PC survey, PO survey and in-depth interviews with the POs.

The project intended to achieve four key outcomes: expanded capacity for PEM; increased use of evidence-based policy reforms; internationally comparable information or benchmarks on public expenditures; and, creation of a strong network of institutions. Impacts discussed here relate to the first two outcomes, which were the focus of the impact evaluation; benchmarking and networking are presented in Section 3.5 on effectiveness. The indicators examined in this section include:

Overall Project Impact

- Quality of policy research,
- Communication effectiveness by POs,
- Policy communities' use of evidence to inform policymaking, and
- Perceived effectiveness of POs in the policy arena.

Individual Outcomes

- Use of research produced by the project,
- Policy impact outcomes of individual POs, and
- Effect of the project on direct beneficiaries

Overall Project Impact

It was expected that project support would increase institutional capacity of POs for public expenditure monitoring (PEM) and analysis, development of policy alternatives, and constructive engagement with the policy community in an effort to increase the incidence of evidence-based policy making, thus improving the governance of public service delivery. Overall, the project improved the research capacity, quality of research reports and dissemination of results by the POs. As of now it has had limited results on impacting policy and increasing the interest and use of research by the policy community. It is important to note that POs have only just completed the policy simulation reports and therefore constructive engagement with policymakers using these reports is currently under way at the time of this report writing. Therefore, this evaluation does not capture the full impact of all communication efforts.

Quality of research. We found that on average the quality of reports produced by POs increased by seven percentage points over the project period. Notably, the first reports produced by the project were on average of lower quality than non-project reports but by project end the project-produced reports were on average of higher quality than non-project reports, suggesting that POs at first struggled with the analysis and writing of project-proscribed reports but by project end had greatly improved.

At baseline (2009), midline (2011), and endline (2013), each PO was asked to provide six policy research reports for assessment by outside reviewers. At midline and endline, we allowed for a subset of reports to be those produced as part of the project.

The evaluation team, based on prior experience in working with a range of think tanks, believed that six reports selected by the organization as demonstrating its analytic and policy capabilities constituted a sufficient sample to give a valid assessment of an organization's written work. An examination of the score variance for individual POs at baseline shows the variance is generally small, increasing our confidence that we obtained a reasonable mean score from six reports.[12] Worth noting is that expanding the number of reports reviewed would be difficult because some of the more advocacy-oriented POs do not produce many policy research reports. For example, three POs were unable to supply six reports for review at baseline; one was only able to submit two.

The reviewers used a common scoring form that rated eight areas for problem definition and analysis performed and five areas for the conclusions drawn and policy recommendations made. These rating criteria are listed in Table 3.2.1 below. The same scoring form was used in all three rounds.

At baseline 77 reports were reviewed; for midline, 74; and for endline, 77.[13] Of the 74 reports reviewed at midline, 24 were prepared as part of the POs' work on the project (specifically, the reports were on program budgeting and benefit incidence analysis); of the 77 reports reviewed at endline, 41 were

[12] A standard variance measure is the coefficient of variation (COV), i.e., the standard deviation of a set of numbers divided by the mean value of the numbers, with the product multiplied by 100. Values of 50 and below are generally considered to indicate limited variance. Among the 14 organizations, at baseline only two had COVs greater than 50; 3 were under 20 and 5 in the 21-40 range.

[13] Note that these numbers differ from the numbers in the Mid-Term Review because here we dropped the observations from Tanzania.

prepared as part of the project.[14] An initially-formulated hypothesis was that these reports are stronger than others because of the project participants' close interaction with highly qualified project staff in preparing them. However, some POs did not have access to strong data sets for such analyses and all were constrained to using standard table templates and report outlines—factors which could weaken report quality.[15] Thus, we have no firm expectation as to how these reports compare to others under review.

Before turning to the findings, it is wise to recall that local analysis and report standards vary from country to country, and not all are consistent with "international standards." Therefore, it is possible for POs to produce reports that are accepted as providing a strong policy development basis while they are at the same time not consistent with policy research standards of multilateral donors and peer-reviewed journals. Reviewers were instructed to score the reports against international standards rather than local ones. It is important to note that the reports were not edited or revised for publishing therefore, we focus on changes in scores over the course of the project period rather than the average scores themselves.

We also note that while reviewers are published authors in peer-reviewed journals and they were given extensive guidance on scoring the reports, it is possible that some difference in the scores depends on who did the rating. Unfortunately, we did not have the resources for more than one reviewer to rate each report. To minimize the variance in report scores associated with reviewer-specific idiosyncrasies to the maximum extent possible we engaged the same reviewers to review reports from the same organizations in each round of reviews. However, two of the 14 baseline reviewers were not available at midline, and other baseline reviewers took their place. At endline, only 8 of the original 14 baseline reviewers were available, and we employed four new reviewers. Finally, organizations differ in their missions and in the type of work they do. Hence, a significant range of scores is expected. The descriptive results of the report reviews are in Table 3.2.1 below. Each of the 13 areas is rated from 1 to 10, with 10 indicating the highest score. Note that a subset of the 13 areas may not have been applicable to each report, so the N values vary.

Generally, the mean scores at all three rounds are within a range of 0.5 points. A3, A5, B1, and B2 are the only exceptions to this rule; A5 and B2 are the only questions with a range greater than one. Additionally, the trend in scores is not consistent – some areas increase from one round to the next, others decrease, and

[14] During the observation period POs were to prepare reports on program budgeting and benefit incidence analysis for the three sectors that are the program's focus. In principle, 90 reports could have been available for review, i.e., 3 sectors x 15 POs x 3 types of analysis. However, most POs consolidated reports in various ways. The 26 reports were essentially selected at random. Of those reviewed only one was for a single sector and one type of analysis. Seven POs combined analysis for all three sectors in a single report for program budget analysis and another consolidated report for benefit incidence analysis. POs from Argentina, Mexico and Nepal submitted a single comprehensive report for all analyses and all sectors.

[15] Moreover, GDN informed the NORC team organizing the reviews after the reviews were completed that these reports were not envisioned to be polished reports. The NORC team had mistakenly assumed the reports would be more finished

still others waffle back and forth with peaks and valleys at the midline. These results are overall fairly inconclusive, so we turned to more rigorous methods of analysis.

Table 3.2.1: Mean Scores of Report Reviews

		2009		2011		2013	
		Mean Score	N	Mean Score	N	Mean Score	N
A	**Organization and analysis**						
A.1	Is the issue well-defined and the case for its policy importance effectively made?	7.2	77	7.0	74	7.0	76
A.2	Is the issue defined or structured in such a way that a clear hypothesis or researchable question is stated?	6.4	77	6.3	73	6.7	76
A.3	Are all the relevant aspects of the issue included for analysis?	6.6	76	6.0	73	6.1	76
A.4	Are relevant previous studies on the issue in the country cited and built on?	5.8	75	5.2	74	5.4	76
A.5	Do the authors show knowledge of the relevant international studies on this topic?	5.4	77	4.3	74	5.7	76
A.6	Has the right type of information and data been assembled to address the issue? If not, what was omitted that should have been included? Where sample data are employed, is the sample correctly drawn to be representative? Is it sufficiently large for the necessary tests?	6.3	76	6.1	74	5.8	76
A.7	Are the methods employed appropriate? Are statistical tests used where needed?	6.2	76	5.7	73	6.1	76
A.8	Is the report well-organized and clearly and succinctly written?	6.5	77	6.3	74	6.2	76
B	**Conclusions**						
B.1	Are the conclusions based squarely on the paper's findings? (or do the authors go beyond the findings in effect expressing personal views or political opinions?)	6.6	77	7.2	74	6.2	76
B.2	If the conclusions call for action through government programs, is the cost realistically estimated? Is the administrative feasibility and complexity of the program considered?	4.5	64	3.8	56	5.2	64
B.3	Do the authors consider various options for addressing the issue and the merits of each, or focus exclusively on a single approach?	5.2	70	4.8	72	5.2	75
B.4	In general, do the authors draw out the full policy implications of the findings and make realistic suggestions for their use in changing current policies?	5.3	70	4.8	71	5.0	76
B.5	Where appropriate, do the authors suggest what	4.5	59	4.3	62	4.5	76

additional data could be collected and/or analysis undertaken to better answer the question posed or to answer additional questions the study raised?

We had to address two issues that prevented us from using a simple t-test to establish whether there is a statistically significant difference between the baseline and mid-term report scores. One is that the number of reports provided by some POs changed which can affect the difference between the mean score values of the three rounds. To see this consider that at baseline three missing reports were from a comparatively weak organization and had low scores. Now consider that at midline the missing reports were from an organization that had very high scores at baseline, with the weak organization submitting all six. The difference in mean scores would be artificially low because of this substitution.

The second issue arises from individual reviewers scoring on different scales despite the guidance given, e.g., different reviewers give higher or lower scores to the same reports. Some may score systematically higher or lower than the baseline reviewers they replaced. A related issue is the change of some reviewers between the rounds of reviews.

We addressed these issues by testing for a significant difference in scores between rounds by estimating a regression model in which the dependent variable is the report score (scaled to a percentage of possible points, so scores range from 0-100). The independent variables include dummy variables for the report round, dummy variables for each of the report reviewers,[16] dummy variables for each of the POs, two dummy variables indicating whether the report was produced as part of the project during the second and third round of report reviews, and a dummy variable indicating that a report submitted at baseline was completed earlier than the three year period specified.[17,18]

The results of greatest interest are the coefficients for the dummy variables indicating round and whether the report was part of the project (Table 3.2.2). “Round 2” indicates that a report was submitted at midline, “Round 3” indicates that it was submitted at endline, “GDN Round 2” indicates that it was a

[16] The coefficients of these variables show the mean difference in the scores awarded by the reviewer indicated by the dummy variable coefficient compared with those of the omitted reviewer. i.e., the PO for which no dummy is included in the model.

[17] POs were asked to submit reports from within a three year period prior to the project (2007-2009). It is possible that earlier reports were “cherry picked,” i.e., stronger but older reports were submitted. On the other hand, POs submitting older reports may have done so just to reach the six report “quota.”

[18] If the coefficient for a dummy variable indicating the round of reviews is statistically significant, then there has been a significant change in the scores from baseline to that round. Similarly, if the dummy variable indicating the report was produced as part of the project is statistically significant, then the project can be said to be raising the quality of analysis done for it compared to others.

project report submitted at midline, and "GDN Round 3" indicates that it was a project report submitted at endline. Note that our "base" is the Round 1 report scores and an individual report can fall in both the generic round variable and the GDN round variable.

After controlling for other factors, including the difference between project reports and non-GDN reports, Round 2 is positive and statistically significant in all three models, and Round 3 is positive and statistically significant in the models for Part A and the Overall Score. Overall report scores increased by about 7 percentage points over the observation period; this is a substantial accomplishment. Interestingly, the reports produced as part of the project at midline received significantly lower scores, on average, than other reports. This negative effect disappears in the GDN reports produced at endline for the overall score and Part B, suggesting that there was a steep learning curve for POs in producing GDN-quality reports which they collectively overcame by the end of the program, most notably in the conclusions and recommendations of the reports. By project end, on average a project-produced report's overall score was more than 7 percentage points higher than a baseline report, representing a very large improvement by POs in the analysis and writing of project reports.

Table 3.2.2: Key Coefficients of the Regression Analysis of Change in Report Scores

Model	Round 2	Round 3	GDN Round 2	GDN Round 3
Overall Score	6.91**	7.00*	-14.87***	1.11
Part A: Problem Definition and Analysis	5.39*	8.24**	-16.37***	-2.90
Part B: Conclusions and Recommendations	8.68**	4.19	-10.59**	7.91
* = P<0.10, ** = P<0.05, and *** = P<0.01				

Communications effectiveness. There has been a statistically significant somewhat large increase in the level of contact and awareness of POs among the policy community. The PO's communications with the policy community, including both events and publications is shown in Table 3.2.3. The top panel of the table, "Contact and Awareness", displays information on the frequency with which policy community members had interacted with information from the PO in various forms, ranging from face-to-face meetings to visiting its website. The mode was either "every couple of months" or "1-2 times a year" for all contact types. While self-reported increases over time in the frequency of interactions were never above 48 percent of respondents, the average score for all but one question had significant increases in ratings from baseline to endline. Moreover, for almost all questions, there was a large jump from midline to endline in the percentage of respondents answering "once a month or more often." This is perhaps due to the fact that NORC's midterm review highlighted a weakness in communication and resulted in the project engaging a consultant to focus on increased communication training. Between midline and endline In addition, the POs were likely near the peak of their advocacy activities for the project when the endline survey was fielded.

Somewhat surprising is the high frequency with which policy community members reported meeting with PO staff—16 percent said monthly and another 30 percent said every couple of months at midline, and for endline the parallel figures are 24 percent and 28 percent. We do not know the share of these meetings that were initiated by PO staff.

Whereas website usage and attendance at events were the least popular means of contact with POs, given that by project end about 20 percent of respondents had never visited the PO website or attended a PO event, those that did use the website or attended an event felt they had very significantly improved over the course of the project. Those using PO websites gave them high marks for being easy to use (in the panel labeled "Website") and perceived substantial improvement – the ratings increased significantly at midline and endline, and all improvements were large in magnitude.

Policy community ratings of POs' events were strong. In the endline survey, 49 percent of respondents or more gave the highest ratings for the POs' ability to attract the "right people" to events, organize events, and provide high quality content, improving from 44 percent in the baseline. In some countries it is difficult to get responsible government officials or MPs to participate in roundtables and similar events which undermines the events' usefulness. The problem does not appear to be widespread. Substantial improvements are seen in the questions: "Were the events well organized in terms of logistics?" which increased from 36 to 57 percent in its rating of "very well done;" while the question "How do you rate the contents of these events?" was given a rating of "very well done" by 35 percent in 2009 and improved to 51 percent in 2013.

We also note that the percent of respondents noting improvements increased significantly for all aspects of events both from baseline to midline and from midline to endline, and the magnitudes of these effects are very large.

The policy community's perception of PO publications is quite positive. In the first three topics under "Publications" more than 90 percent rated POs in the top two categories, while in the fourth about 75 percent did so. For each topic, about half of respondents expressed that publications had improved, and the mean ratings on each improved significantly from either baseline to midline, baseline to endline, or both.

Two observations are in order in assessing the information just presented on changes in research quality and communications. First, the detailed questions on events, publications, and website use were only asked to those who attended events or visited the websites, and therefore some of the sample sizes are small. Second, we cannot attribute direct causality to the GDN project for the changes observed since there is no comparison group to indicate or disprove if the improvement is just a natural progression in absence of the program.

Table 3.2.3: Communications Ratings

Contact and Awareness	Year	N	% of Respondents Answering: Once a month or more often	Every couple of months	1-2 times a year	Never	% of Respondents Rating as an Increase (b)	Significance (c)
How often do you receive letters, memos, reports or publications or correspondence from the organization?	2009	139	19	32	32	17	--	--
	2010-11	189	14	28	37	21	27	0.103
	2013	173	29	32	25	14	48	0.658***
In general, how often do you see the organization or its work mentioned in newspapers or magazines?	2009	129	19	45	30	6	--	--
	2010-11	172	18	40	37	5	36	0.802***
	2013	164	30	37	27	5	48	0.971***
How often do you have face-to-face meetings with staff from the organization?	2009	142	14	27	47	11	--	--
	2010-11	193	16	30	40	15	31	0.252
	2013	179	24	28	31	17	45	0.333
How often do you use the organization's website?	2009	138	12	36	32	30	--	--
	2010-11	189	10	28	31	32	25	0.401**
	2013	174	20	31	26	23	36	0.774***
How often do you see or hear the organization or its work mentioned on television or radio?	2009	118	11	38	36	15	--	--
	2010-11	164	10	28	41	21	33	0.921***
	2013	151	20	32	32	16	40	1.244***
In the past year how often have you participated in events that the organization arranged?	2009	142	6	20	46	27	--	--
	2010-11	188	11	15	45	28	31	0.361*
	2013	173	11	21	49	19	33	0.359*
Is the website easy to use?	2009	102	33	53	12	2	--	--
	2010-11	134	48	44	6	2	42	2.203***
	2013	137	55	36	4	4	56	1.782***
Does its website provide the information you need?	2009	107	11	67	18	4	--	--
	2010-11	142	31	46	17	6	42	1.971***
	2013	140	37	47	10	6	51	1.694***

See notes at the end of the table.

Table 3.2.3: Communications Ratings (continued)

Events			Very well done	Above the usual level	Typical	Below the average	% of Respondents Rating as an Improvement (b)	Significance (c)
Did the events include the right people, i.e., those most concerned and knowledgeable on the issue?	2009	106	44	43	11	1	--	--
	2010-11	138	49	30	20	1	50	1.795***
	2013	143	49	32	17	2	55	1.114***
Were the events well organized in terms of logistics?	2009	103	36	44	20	0	--	--
	2010-11	136	46	31	19	4	47	1.927***
	2013	144	57	22	20	1	61	1.347***
How do you rate the content of these events?	2009	105	35	46	19	0	--	--
	2010-11	138	38	42	19	1	51	1.859***
	2013	144	51	28	19	1	62	1.340***
Publications			**Very well done**	**Above the usual level**	**Typical**	**Below the average**	**% of Respondents Rating as an Improvement (b)**	**Significance (c)**
Is the information contained in them reliable?	2009	117	66	32	3	0	--	--
	2010-11	140	75	24	1	0	52	0.583**
	2013	143	71	29	0	0	63	0.141
Are they interesting and informative to read?	2009	117	60	35	5	0	--	--
	2010-11	141	66	33	1	0	52	0.348
	2013	142	78	20	1	0	71	0.720**
Are they timely?	2009	112	29	63	8	0	--	--
	2010-11	134	47	45	8	0	45	0.935***
	2013	137	48	47	4	1	55	0.921***
Have they changed the way you think about certain economic, policy or social issues?	2009	114	19	58	18	4	--	--
	2010-11	133	27	50	14	9	53	0.529**
	2013	140	26	51	17	6	55	0.383

(a) All questions used a four-level response pattern. The specific wording varied among questions. Net of respondents who said they did not know or did not want to answer.

(b) Percent of respondents who said in the follow-up survey there had been an increase or an improvement in this area over the past 18 months.

(c) Displays results of a statistical test of significance of the change between responses conducted ordered logit regression analysis with the rating as the dependent variable and independent dummy variables for each PO and survey round. The entry is the coefficient in the model for the rating being from the 2010-11 survey or 2013 survey. It shows the mean points of improvement or decline in the mean score relative to baseline, e.g., a coefficient of 0.3 for 2013 indicates that the average 2013 survey score is 0.3 higher than the average baseline survey score. * = P<0.10, ** = P<0.05, and *** = P<0.01

Despite the findings that there have been improvements in communications effectiveness, there is continued room for advancement. Interviews with POs highlighted that for a few, limits in communication and dissemination skills and constructive engagement with the policy community affected POs' ability to fully achieve their objectives regarding changing policy. An Asian PO remarked that lack of experience liaising with policymakers was a limiting factor, but included mention of ties with a few concerned government officials. Although commenting that its advocacy goals had not been achieved, another Asian PO mentioned that it had been progressively building a network of policy stakeholders, some of whom had solicited the PO for technical input.

- *"[Our goal is]To increase capacity of research output, and contribute to social change through policy improvement in development sectors. We have achieved to a large extent the first one, but only to a limited extent the second one. We did have improvement in capacity of research. This was a good opportunity for us to learn from global network. We don't have the experience of communicating with policy makers; that's why the first one is achieved but not the second one. (Asia)*
- *"… we realized there is a very big gap in putting together water sector advocacy…When we started it was a difficult climb. But as we move along we built up expertise because of our networking with people who work in the sector." (Asia)*

The results of our PCS survey suggest that PO outreach to the policy community could be improved. Over half the policy community respondents either have not met with POs or only sought out POs one to two times a year to obtain information for policy purposes over the last year although results vary widely by country. In the endline survey, we added a question asking PCS respondents whether they had met or called a PO within the last year to obtain information for policy purposes or to learn about the results of the organization's work. Overall, as shown in Table 3.2.4, just over half of the PCS respondents had met with the PO two times or fewer in 2013[19], while the remainder met with them at least every couple of months. The results for the individual organizations are extremely varied – for two POs, all policy community members had met with them in the last year, while for others, over half had not done so.[20]

[19] The survey question asks about frequency of contact "in the past year". The survey was fielded from Sept.-Nov. 2013 the time frame roughly corresponds to 2013.

Table 3.2.4: PCS Respondents' Meetings with POs

In the past year, how often have you met in person or talked by phone with [PO name] to get information for a general policy discussion or to learn of the results of the organization's work?					
	Once a month or more often (%)	Every couple of months (%)	1-2 times a year (%)	Never (%)	N
All countries	**17**	**26**	**35**	**22**	**174**
Argentina	15	31	23	31	13
Armenia	0	18	73	9	11
Bangladesh	0	40	60	0	10
Ghana	20	40	30	10	10
Guatemala	25	23	36	16	44
India	14	19	52	14	21
Indonesia	45	0	18	36	11
Kenya	38	63	0	0	8
Mexico	0	0	25	75	4
Nepal	0	20	40	40	5
Nigeria	0	67	0	33	6
Peru	7	14	29	50	14
Philippines	33	0	0	67	3
Uganda	14	36	36	14	14

NOTE: This question was only asked in the 2013 survey.

Lastly, members of the policy community were allowed the opportunity to provide additional feedback on the POs as part of the PCS survey. Respondents' suggestions on enhancing policy impact included making research products more accessible to the public by utilizing media and in-country policy networks. POs were encouraged to produce deliverables that focused on specific policy issues, and develop strategic partnerships with other organizations engaged in similar work for wider dissemination. Additionally, POs were asked to conduct consultations with concerned government officials and community organizations to better contextualize their work.

Policy communities' use of evidence to inform policymaking. As demonstrated in Table 3.1.2 in Section 3.1, the results of the PCS do not show much significant change in the use of research, data, and analysis to inform policy making. We also do not find much change in sources of information consulted to inform policy discussions, presented in Table 3.2.5. In the baseline, midline and endline survey the policy community was asked the average number of times they consulted with various information sources in the year prior to the survey. The overall breakdown between the different types of information sources remains the same over time. Searches on the internet are by far the highest with average consultations at 45.6 times on average over all three years combined. Consultation with local think tanks or policy research NGOs does show a slight improvement over the three years; starting at 6.1 on average in 2009, and increasing to 7.8 and 7.7 in 2010-11 and 2013 respectively although none of these were statistically significant changes.

Table 3.2.5: Consultation by Policy Community of Various Information Sources

Information Source	Average Number of Times Consulted in Year Prior to Survey			
	2009	2010-11 (a)	2013 (a)	All Years Combined
Relevant government ministry or agency	12.4	12.5	14.6	13.2
Government research institutes	4.6	6.2	3.5	4.8
International agencies such as World Bank, UNDP, etc.	7.4	6.4	8.3	7.3
Local think tanks or policy research NGOs	6.1	7.8	7.7	7.3
Local advocacy NGOs	4.4	4.2	5.4	4.7
Professors and university institutes	5.6	7.7	5.4	6.3
Search the internet [a]	41.6	50.9	42.8	45.6
Others	31.9	9.1	13.6	13.8

[a] All responses greater than 500 were recoded to 500.

(a) T-tests were performed to test whether the average number of times consulted was significantly different from the average of the previous survey round, but no statistically significant differences were found.

Perceived effectiveness in the policy arena. Our data show mixed results regarding POs' increased effectiveness in the policy arena over the project period. We find that POs are viewed as generally being more effective but when asked specifically about the POs having an impact on public policy or administration, public expenditure quality, and the budget-making process, we see less of a change over time, particularly on the budget-making process. Based on data from the PCS, Table 3.2.6 displays the distribution of answers across the four response categories for questions in two areas—general effectiveness and effectiveness on transparency and accountability issues. There are sharp distinctions in the ratings and in their progress over time. POs were seen as focusing on issues of high priority, being valuable sources of information and providing helpful policy recommendations; within the general effectiveness group, the ratings on the first, second, and fourth questions were indeed strong, with 50 percent or more of respondents giving the highest score and nearly no one giving the two lower ratings. The fourth question also shows improvement over time, with the percent of respondents answering "very much" increasing from 51 percent to 60 percent to 71 percent. These three questions are on the very heart of policy research organizations' work and speak well for the POs.

In contrast, the ratings for the third question in this group - on whether POs actually have a positive impact on public policy or administration - were lower and they did not change much over time. Only about 30 percent of respondents gave ratings in the highest category in each year, a distinctly lower share than for the other three questions. Between 14 percent and 19 percent also rated POs in the lowest two categories, which is a higher share than for the other three questions.

Did the policy communities' perceptions of the POs change over time? We employed two approaches to measure this. The questionnaire asked respondents both (1) to provide a rating on the 4-point scale for each performance area, e.g., "Does the organization focus on high priority issues?" and (2) to score whether the organization's performance in an area had improved, stayed the same, or declined over the period. The second question was asked at midline and endline because we were concerned that respondents included in multiple survey rounds could be inconsistent in their ratings. For example, while implicitly meaning to indicate an improvement a respondent could actually record a lower value for an area in the midline survey than he did in the baseline survey, i.e., a decline. Additionally, the second and third rounds of the survey included a number of persons not surveyed at baseline. For these reasons one might prefer this measure over the comparison of baseline and post project survey ratings.[21]

We performed Chi-square tests to check whether the proportions of respondents answering "increase," "decrease," or "no change" were equal. In all cases, the null hypothesis of equality was rejected. Since all results are the same, indicators of significance for this test are not included in the tables.

As shown in Table 3.2.6, the proportion answering "improvement" was generally a majority. Note that in reading the tables if the percentage of responses is under 50 percent, a majority believed that there was no improvement, i.e., performance declined or stayed the same.

To determine whether a score changed over time, i.e., comparing the 4-point scale ratings at baseline and follow-up, is statistically significant it is not appropriate to employ a simple t-test using individual response data. The reason is that the number of respondents from each of the sample countries can vary between the survey rounds. If the response numbers are higher from countries with lower scores on average, the overall mean would be lower even though the mean for most countries could have increased. Using country means is problematic for two reasons: (a) the reduction in the number of observations from several hundred to 42 (14 POs for 3 rounds) makes it more difficult to detect significant differences; and, (b) the small number of observations used in computing the mean values for some POs makes the values less accurate.

[21] Because of the way confidentiality controls were structured for the surveys, we are not able to identify all those who participated in both survey rounds.

This issue was addressed by estimating a regression model in which the dependent variable is the rating given for the question. Independent variables include a dummy variable for each PO (except one that must be omitted to permit model estimation) and a dummy variable for each of the follow-up survey rounds. The dummies for organizations control for variations in the number of responses from each organization in the three survey rounds.

Returning to Table 3.2.6, one sees in the final two columns the conundrum of change ratings pervasive in the results. On the one hand, respondents indicated that POs' general effectiveness improved since the previous survey round (next to last column). Around 50-60 percent indicated an improvement in this area for the indicators in the first group. Smaller percentages of respondents, only around 35 percent, said POs performed better in having an impact on transparency and accountability issues.

On the other hand, we did find statistically significant and positive differences between baseline and midline ratings (one question) and baseline and endline ratings (four questions) on the four point scale (last column). The area with the most drastic increase is the perception that POs are focusing on high priority areas – this rose by an average of 0.931 points from baseline to endline. The other increase larger than 0.5 is in the rating of holding governments accountable for public expenditure quality from baseline to endline, which rose by 0.566 points. Other statistically significant increases – on the helpfulness of policy recommendations from baseline to endline, the rating of the PO as a valuable source of research and statistics from baseline to endline, and the focus on high priority issues from baseline to midline – are all between 0.4 and 0.5.

The increase in the ratings of the PO agendas appears to be a major improvement. The remaining results generally indicate a positive (but more modest) change in policy communities' perceptions of POs' performance. Notably, the ratings of the POs' influence on the budget making process remained quite poor throughout the program.

Table 3.2.6: Rating for Policy Process Roles

	Year	% of Respondents Answering: (a) Very much	Some-what	Not much	Not at all	% of Respondents Rating as Improvement (b)	Significance (c)
General Effectiveness							
Are the organization's policy recommendations helpful?	2009	53	43	5	0	--	--
	2010-11	48	45	7	1	63	-0.104
	2013	59	36	5	0	68	0.436*
Is the organization a valuable source of research, including data and statistics?	2009	56	39	5	0	--	--
	2010-11	54	36	9	1	62	0.050
	2013	61	31	9	0	57	0.410*
Does the organization's work positively impact public policy or administration?	2009	31	54	12	2	--	--
	2010-11	28	52	17	2	50	-0.044
	2013	32	53	14	1	60	0.344
Agenda: Does the organization focus on issues that are of high priority?	2009	51	49	1	0	--	--
	2010-11	60	36	3	1	51	0.454*
	2013	71	27	2	0	61	0.931***
On Transparency and Accountability Issues							
Does the organization have an impact in holding the government accountable for public expenditure quality, i.e., efficient and honest use of public resources?	2009	13	50	29	8	--	--
	2010-11	19	39	33	9	35	0.353
	2013	19	48	24	9	46	0.566**
Does the organization have an influence on the budget making process in terms of openness, quality, or equity of budget choices?	2009	15	50	24	11	--	--
	2010-11	18	42	26	13	36	0.331
	2013	21	35	26	17	4	0.336

(a) All questions used a four-level response pattern. The specific wording varied among questions. Net of respondents who said they did not know or did not want to answer.

(b) Percent of respondents who said in the follow-up survey there had been an improvement in this area over the past 18 months.

(c) Displays results of a statistical test of significance of the change between responses conducted ordered logit regression analysis with the rating as the dependent variable and independent dummy variables for each PO and survey round. The entry is the coefficient in the model for the rating being from the 2010-11 survey or 2013 survey. It shows the mean points of improvement or decline in the mean score relative to baseline, e.g., a coefficient of 0.3 for 2013 indicates that the average 2013 survey score is 0.3 higher than the average baseline survey score. * = P<0.10, ** = P<0.05, and *** = P<0.01

Individual Outcomes

In this section we focus in on the use of the particular research reports produced by the POs for the project and the reported uptake and use of the report findings by the policy community.

Use of research produced by the project. We found that members of the policy community found the project-produced research to be of interest and relevant. Of those from the policy community who interacted with the project-produced research, 78 percent used the information for policy or administration-related purposes. Of these 78 percent, a further 80 percent shared the information with others.

The use of research produced by the project first requires it to be publicized and available to the policy community via event and reports/policy briefs/memos. In the endline survey we asked PCS respondents if they had attended events or read publications resulting from the PO's participation in the project. Table 3.2.8 shows that more respondents had interacted with the health- and education-related research than in the water-related research, and that more respondents were involved in program budget analysis and benefit incidence analysis than in cost effectiveness analysis and policy simulations/recommendations. This can be explained by the fact that (i) there were many issues with implementation of water sector research due to problems with obtaining data (this is discussed in further detail in Section 3.4); (ii) that PBA and BIA analyses were mostly completed by the third year of the project and had greater opportunity for dissemination, and (iii) the CEA were only just completed and many PS reports had not yet been submitted at the time of data collection for this report.

Table 3.2.7: Use of GDN PEM Research

Have you attended any events or read any publications (reports, policy briefs, or memos) that presented the results of any of the following research and analysis produced by [PO Name] between 2010 and the present? Percent answering "Yes"

	Program Budget Analysis			Benefit Incidence Analysis			Cost Effectiveness Analysis			Policy Simulations/ Recommendations				
	Health	Edu-cation	Water	Health	Edu-cation	Water	Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3	None	N
All Countries	**48**	**48**	**26**	**42**	**44**	**20**	**31**	**22**	**15**	**30**	**25**	**12**	**23**	**172**
Argentina	38	46	8	46	38	8	15	23	8	31	23	8	23	13
Armenia	64	45	36	45	36	27	55	36	27	36	18	9	9	11
Bangladesh	55	64	45	45	55	18	9	0	36	18	0	0	9	11
Ghana	50	50	40	50	50	50	40	20	20	30	20	20	30	10
Guatemala	40	51	11	38	56	9	40	18	0	42	38	7	22	45
India	74	74	58	63	47	37	32	47	42	26	42	11	11	19
Indonesia	55	36	27	55	36	36	27	0	9	36	9	36	27	11
Kenya	75	63	63	50	63	38	25	25	--	--	13	--	13	8
Mexico	0	0	0	0	25	0	0	0	0	0	0	--	75	4
Nepal	0	25	0	0	25	0	25	0	25	50	25	25	50	4
Nigeria	33	33	0	50	50	0	50	17	0	33	0	0	50	6
Peru	50	25	8	42	25	8	17	33	17	8	42	0	25	12
Philippines	0	0	0	0	0	25	0	0	0	0	0	0	75	4
Uganda	50	50	36	29	36	21	36	36	29	36	21	50	7	14

NOTE: This question was only asked in the 2013 survey. "Topic 1", "Topic 2", and "Topic 3" were all personalized for each organization and are within the health, education, and water sectors, respectively. Two organizations did not have relevant topics to ask about for all analysis types.

Table 3.2.9 presents further details about the use of the project research and analysis used by the policy community. Of those who attended events or read publications listed in Table 3.2.8, a respectable 78 percent used the information from the reports in a wide variety of ways. Of those who did not find it useful, fully half said the reason was that the topic was not of personal interest, and few cited reasons related to the quality of research. This speaks highly to the quality of work done within the project. In Table 3.2.10, we explore the sharing of the GDN PEM research results by the policy community. Over 80 percent of respondents who used the research results also shared it with others. Most often, they shared with people inside their own organization, but also frequently with friends (48 percent), academics (40 percent), and government departments/ministries (37 percent). This, too, suggests that the policy community found the research results very useful and well-executed. Most respondents – 63 percent - said that they were very likely to use the information within the next year, and none said they were not at all likely to do so.

Table 3.2.8: Reasons for Using POs' Research and Analysis

Did you use any of the information contained in these research and analysis reports?			Why did you not use this information? Lacking in data	Weak analysis	Topics were not national priorities	Topics not of personal interest	Not timely
No	22	→	9	22	22	52	17
Yes	78	↓					

How did you use this information?	% "Yes"		To what degree did you use the research and analysis information in the activity or to make the change? Exclusively	Mostly	Somewhat	Very little
As evidence in policy debates	58	→	5	54	38	3
As evidence in advocacy efforts	42	→	5	55	36	5
To raise public awareness of an issue	60	→	8	56	33	3
To make changes in policy	32	→	3	45	42	9
To make changes in legislation	19	→	0	50	30	20
To make changes in budget allocation	31	→	9	44	38	9
To make changes in implementing current policies	30	→	6	42	42	10
To provide access to information	55	→	8	42	42	7
Other	12	→	8	25	58	8

NOTE: This question was only asked in the 2013 survey.

Table 3.2.9: Use and Sharing of POs' Research and Analysis

Did you share any of the information contained in these research and analysis reports with anyone else?	
Yes	81
No	19
With whom did you share this information?	
Inside my organization:	
Within my department	77
With another department	46
Outside my organization:	
A government department/ministry	37
Parliament	18
Research organization	30
Advocacy organization	26
Other NGO	32
Multilateral/Bilateral organization	23
Academics	40
Media	26
Private sector	28
Personal friends	48
Do you anticipate using this information within the next year?	
Very likely	63
Somewhat likely	28
Marginally likely	9
Not at all likely	0

NOTE: This series of questions was only asked in the 2013 survey.

Policy impact outcomes of individual POs. In this section we attempt to understand whether the research produced by the project had an impact on policy. We found that the policy community that had interacted with the project-produced research generally believed that it had a contributed to policy changes, although to a somewhat limited degree . Almost all POs spoke of having some sort of influence on policy; only two felt they were unsure of their policy impact.

The PCS survey asked respondents who had interacted with the project-produced research whether they thought it had contributed to any policy changes in the health, education, or water sectors. As we see in Table 3.2.11, on average across all the countries, the majority of respondents felt the project-produced research was somewhat responsible for a change in policy (34 percent) or that it made a small contribution (28 percent). Looking at each country individually (ignoring those with less than five responses and those that said they didn't know) we see in all cases that the majority felt there was at least some sort of contribution of the project-produced research to policy change. The extent of this contribution was limited since only four countries had any respondents who felt the project-produced research was mostly responsible for a change in policy – notably in India this was about one third of respondents. A fair share of respondents, including 4 of 7 in Bangladesh and more than 20 percent in 7 other countries, reported that they did not know whether the PO's project-produced research had contributed to a policy change.

Table 3.2.10: Perception of POs' Direct Impact on Policy

Has the research and analysis in these reports contributed to any policy changes in the health, education, or water sectors in your country?						
	Mostly responsible for a change in policy (%)	**Somewhat responsible for a change in policy (%)**	**Made a small contribution to a change in policy (%)**	**Had no effect on policy (%)**	**Don't know (%)**	**N**
All countries	**8**	**34**	**28**	**7**	**24**	**137**
Argentina	0	10	40	20	30	10
Armenia	0	30	50	20	0	10
Bangladesh	0	14	29	0	57	7
Ghana	0	43	14	14	29	7
Guatemala	8	24	49	5	14	37
India	32	26	11	0	32	19
Indonesia	11	33	22	0	33	9
Kenya	0	86	0	0	14	7
Mexico	0	0	0	100	0	1
Nepal	0	33	33	0	33	3
Nigeria	0	67	0	0	33	3
Peru	0	40	30	10	20	10
Philippines	0	0	0	0	100	1
Uganda	8	62	0	0	31	13

NOTE: This question was only asked in the 2013 survey. 93 of 237 respondents skipped this question due to survey skip patterns.

The POs were also interviewed on whether PEM-produced research had been used by policy community stakeholders, and if so, by whom. Table 3.2.1X summarizes these comments by region and sector.

Table 3.2.11: Number of POs that mentioned specific contribution to policy

	Sector			
	Health	**Education**	**Water**	**Total**
Latin America	2	2	0	4
Asia	2	1	1	4
Africa	2	1	0	3
Total	**6**	**4**	**1**	**11**

Among the three sectors, research produced in the health sector was most frequently utilized by policy stakeholders. Six POs mentioned specific uses of their work. An Asian PO's BIA helped expand the Ministry of Health's understanding of health service access, though the PO could not definitely state that its recommendations had also been utilized. Another Asian PO's recommendations improved state-level program planning; this PO attributed its success in influencing policy to a strong partnership with government counterparts. A Latin American PO achieved similar success, as its research was utilized by multiple government entities to enhance conditional cash transfer programs. Another Latin American PO's work in catastrophic diseases helped inform policy discussions among government officials, media and academics. Some of the key findings of an African PO's CEA recommending use of long-lasting insecticidal mosquito nets were utilized by government stakeholders. Another African PO engaged in a similar CEA study was able to use the results to convince government officials that the public needed to be better informed about this measure in order to allay fears about safety. In contrast to these six POs, an additional PO indicated that while the findings of their health sector analysis had been presented to senior officials in the Ministry of Health, they were unsure of policy impact.

- *"Some of the results of benefit incidence analysis for health sector were even more detailed and more informative than Demographic and Health Survey (DHS), which is done every fifth year in the country and one of the major reference points for the Health Ministry. We could provide additional data for them, in some of the aspects, because there was more deep analysis to them. We found confirmation of the recent BIA data on health services utilized in rural versus urban areas, which was a total surprise to the Ministry of Health. These types of deep analysis are being utilized. Unfortunately, I cannot say that for the entire paper, especially the recommendations."* (Asia)
- *"For the health insurance, since you're in continuous touch with them for data and support, for next year planning phase they have asked us to come and present some of the findings based on that they tried to redesign their planning. We have recommended something to [state] government; few recommendations they had adopted ...."* (Asia)
- *"Yes, in terms of the health sector…this is where we have got most of the feedback. It is being used. In one of our results we showed that using of indoor residual spraying was the most cost effective as compared to net, but the government had not taken up the initiative to allay people's fears on what chemicals are being used in sprays…We did the research and tried to find out why it was not being used. And what we got from those who had failed to take it on was 'We don't know what's contained in that chemical.' So we went back to them and told [them] that people don't feel it is safe…So that ability to be more forthright with this information even before it's asked of you we felt was not being done. So this is what we told them. We have a good feeling that they're going to do this."* (Africa)

Four POs highlighted specific uses of their education research products. An Asian PO's BIA was utilized by the Ministry of Education to develop higher education policy, and segments of the final paper were included in a publication released by government officials. As captured by the media, one Latin American PO's research helped initiate a dialogue with top officials in the Ministry of Education around issues concerning teachers' union representatives. Another Latin American PO's research was utilized by multiple government entities to inform improvements in conditional cash transfer programs. An African PO incorporated their project work into another project that monitored teacher absenteeism in primary schools, which was expanded by the government into a national-level study that the PO was asked to evaluate. In addition to these four POs, an Asian PO mentioned that while their work had not been directly utilized, its importance was recognized by a senior state government official. An African PO had engaged government officials in discussions regarding their work, but was uncertain of policy impact.

- *"We have had the greatest success in [the] Education sector. In the end of the second year of the project, the Minister of Education picked up our findings, specifically the benefit incidence analysis, and used it to formulate selling points for government to use new strategy for higher education. This later became a parliament approved strategy. The graphs that we did are in a parliament-approved document." (Asia)*

- *For the [health and education] analysis, it was used by [government officials involved in these programs]. They were very interested about how to improve the program, because during the last government there were a lot of difficulties and controversies with population about implementation of the program. The new government elected in 2011 decided to take into account some of our recommendations, in order to improve the implementation of the program. Now they are working with some insights that we had presented to the Minister." (Latin America)*

> "The education research documents have been most extensively used; the Minister of Education and Deputy Minister have recommended and given feedback on those documents. They even appear in one picture in the media with the documents in their hands. It has affected the understanding of the central problems in the education sector on behalf of teachers unions' representatives, which is something new for [us] to have influence in this particular area. Of course it has been used by media and academics; we have no problem on that side." (Latin America)

- *"Yes, as part of this research, which we combined with some other research, we looked at some of the challenges in delivering education services in primary education. One of the key findings was that there is a lot of absenteeism by teachers. So using another project, we tried to examine whether it's possible to more regularly monitor these teachers and then change their behavior. And that we did it by piloting the use of just a simple mobile phone that either the Head Teacher punches in teacher* *attendance. Six months after doing this, the government came up with even a larger project, larger than ours; they were not just only targeting teachers, they were even looking at pupils. We're doing a weekly basis; they're doing it on a daily basis…They've come back to us and said 'We may need you to evaluate this program. Since you've started it, and you know how it works.'" (Africa)*

Research conducted in the water sector was least utilized by policy stakeholders. Only one PO indicated that their findings are helping to inform policy decisions. This PO, based in Asia, was able to encourage government officials to invest in improving existing infrastructure. Another Asian PO commented that they were more highly regarded for their work in the health and education sectors, and were unaware as to whether their work in the water sector had affected policy change. For two Latin American POs, their water sector findings had not yet been disseminated at the time of interview since the final reports were under development.

- *"Government used to dedicate few resources to rehabilitation and maintenance of existing water system. They are always tempted to finance new projects and abandon the old ones. We showed that how with less money you can enhance the coverage of drinking water." (Asia)*
- *"The water research has not been used because it is not ready. After we finish that [policy simulation], we had to think on what is useful and who we're going to interact with to share the results of that research. The discussion on water is not high profile here compared to other sectors, so we have to think carefully." (Latin America)*

Many POs were able to speak more to the use of the PBA, BIA, CEA and PS analysis than to successes achieved within particular sectors. An Asian PO observed that both PBA and BIA were utilized by policymakers, but that analysis from the latter tools has not been as influential since it was only recently disseminated. PBA findings disseminated by a Latin American PO were also utilized by policymakers as well as PS to some extent, though BIA and CEA were not similarly taken up. Another Latin American PO achieved success in sharing key findings from use of all four tools, engaging media, academia and policymakers at the national level. For a third Latin American PO, the results of their PBA and BIA were utilized by think tanks, media and government stakeholders, though the CEA was not as valuable since the tool is less commonly used in the country. This PO was also able to disseminate a standard methodology developed for PBA and BIA to members of the policy community for wider use. An African PO was successfully able to disseminate PBA and BIA findings to media and policymakers, but faced challenges in sharing results from the PS due to lack of necessary data in the education and water sectors.

Another African PO reported that their BIA will be utilized by academia, and relevant findings have been discussed during a recent meeting with Parliament on the national budget.

- *"I am using CEA more because it is very easy to teach adults. We haven't used that much the results of the policy simulations. We are in the process of cleaning the research, beginning to do more communications results – like booklets, books on CEA…We are going to present those results to many of the project managers of the programs, and also in the media it has appeared significantly. One of things that happens a lot here in discussion of the budget is the percentage of the budget that has been executed by the sector or project. We have trying to get out of that question; it gives us zero information – you can spend 10 percent doing nothing, or 0 percent doing a lot. We are telling instead these are the things you need to be concerned on how much goes to this program, who gets the benefit, as a first step. In that sense, last year we have been successful in reaching the media, academia, program directors. This year, we plan to get national through communications – we don't know by email, internet or how – but to get into the regions." (Latin America)*

- *We develop a standard methodology for analysis with GDN's material for PBA and BIA. It was very important because here we have the information but not a standard way to organize the information. We shared that with other think tanks and NGOs and with the [government agency]. As I know they're using it for specific analysis in other areas. About the CEA, it's not so commonly used here. But there are a couple of organizations that before they use that technique will try to improve it. At the end of the project, during the last two years, it was focused on policy simulation. There was no standard methodology; it was just each group picking up what kind of analysis they want to make. For the case of [our country] we decided to use standard methods we used in other analysis. If we did not share that it's because we didn't know how to do that. For the case of PBA and BIA, we produce an annual report for [our country] – it is additional to update that PEM requires every year. It is used by journalists, think tanks. The [government agency] was very interested about the numbers of how many users of health, education and water services." (Latin America)*

- *"One of the opportunities for using the research will come at the budget stage; June is when the budget is released. It will provide an opportunity for the media to take up any information that we have. In the first and second year we completed the PBA and BIA. One of the things we did is put on [a workshop], inviting civil society, government, media and others; we present any research that has implications for public expenditure…Every year [we make] proposals to Treasury on expenditure and revenue – based on findings from research. We did one last year on education, and tried to tie in the research – did not necessarily want to wait for research to be finalized. We wouldn't be able to say* *what changes have been made. On policy simulation, there has not been as much exposure, partly because there were delays on education and water parts, relating to things in-house and because we could not find data." (Africa)*

Across the various research products and sectors, two POs were uncertain as to whether *any* of their findings had informed policy change. An Asian PO communicated the results of the PBA, BIA, CEA and PS to government stakeholders, but mentioned that affecting policy is a long-term process; however, they were able to encourage members of parliament to consider making further budget allocations to social development programs. Despite the challenges in interactions with government, this PO was successful in forming collaborations with three domestic NGOs interested in PBA and BIA. A Latin American PO clarified that they may have indirectly influenced policy, as their research findings had been shared and acknowledged by policymakers, but they were uncertain as to the extent of policy impact.

- *"Our achievements in changing policy are limited. It will be very hard for us to influence policymakers. We are sure that they are aware of our project documents." (Asia)*
- *"The other things that we have done as a result of the PEM project have been known by the policymakers. Probably not a direct impact, but an indirect impact." (Latin America)*

The review of the "Most Significant Results List" and "Activity to Outcome" documents prepared by GDN yielded additional descriptions of utilization of the project-produced research:

- CIPPEC (Argentina) reported that their education sector research has informed dialogue between teachers' trade unions and provincial ministries of education, and provided a basis for negotiating power for the former party. The findings were highlighted by the President in a speech to media on addressing wage inequity. In the health sector, the Ministry of Health has altered legislation on catastrophic disease treatment based on CIPPEC's findings.
- FUNDESA (Guatemala) reported that the Government has set-up the Social Development Ministry in response to the organizations findings. FUNDESA is working with the Ministry to implement a public-private partnership that supports the implementation and evaluation of CCT programs. Additionally, FUNDESA has created a national database that analyzes budgetary information for the health, water and education sectors and introduced new indicators into policy discourse.

- Analysis conducted by AST (Armenia) on benefit incidence in the education sector has been utilized by the Ministry of Education to address inequities in higher education expenditure.
- Research findings from CSEA (Nigeria) in program budgeting analysis have informed the development of a new performance-based budgeting system that places an emphasis on reporting outcomes. The development of this system has been undertaken in conjunction with the Budget Office of the Federation.

Direct and Indirect Program Beneficiaries

The project's direct and primary beneficiaries are the 14 participating organizations. Over 100 analysts and communications staff have served as key staff at the POs over the duration of the project. In this section we present the effect of the project on POs' project portfolios and capacities as self-reported by the POs through the PO survey and interviews. We follow by a description of the indirect beneficiaries targeted by the project.

Effect of the project on direct beneficiaries. In the PO survey, we asked directly whether they thought the project had a direct impact on their organization in several areas. Overall, the project seems to have contributed to improving competence levels in communications and performing program budget analyses, benefit incidence analyses, cost effectiveness analyses, and policy simulations in POs. By project end the majority of POs felt they had mastered these analysis techniques, with the majority reporting normal competence levels and changes they made to their communications practices due to what they learned from the project.

In Table 3.2.12, we asked organizations to rate their competence in doing different types of analyses that were taught during the project: program budgeting analysis, benefit incidence analysis, cost effectiveness analysis, and (in the 2013 survey only) policy simulations. This question was not included in the 2009 survey, so in 2010 we asked about two time periods – once for before the project, and once for 2010 (i.e. current at the time of the survey). On program budgeting analysis, organizations rated their competence highest in 2010, when nearly every organization reported full competence. This then seemed to trail off until the endline, though 11 of 14 POs still rated their competence as "normal". Benefit incidence analysis also had a peak competence level, in 2011. Competence in cost effectiveness analysis was also highest in 2011, but only 43 percent reported full competence, while 50 percent reported normal competence – so its peak was lower than those of PBA and BIA. Finally, we asked about policy simulation exercises in the endline survey. Here, by the end of the project POs are less confident, as nearly 80 percent reported competence in the middle two categories. The trends seem to indicate that soon after learning the technique, POs felt they had full competence but after putting it in practice and over time, they were less confident and changed their rating to normal competence.

Table 3.2.12: Self-Rated Competence in Analyses

	Self-Rating of Competence Level			
	2010 (Before Project)	2010 (After Project)	2011	2013
Program budgeting analysis				
No competence	7%	0%	0%	0%
Little competence	29%	0%	0%	14%
Normal competence	21%	14%	29%	79%
Full competence	43%	86%	71%	7%
Benefit incidence analysis				
No competence	29%	0%	0%	7%
Little competence	36%	0%	14%	7%
Normal competence	21%	43%	4%	79%
Full competence	14%	57%	71%	7%
Cost effective analysis				
No competence	31%	0%	0%	7%
Little competence	46%	15%	7%	21%
Normal competence	15%	54%	50%	64%
Full competence	38%	31%	43%	7%
Policy Simulations				
No competence	--	--	--	14%
Little competence	--	--	--	43%
Normal competence	--	--	--	35%
Full competence	--	--	--	7%

Note: The question on policy simulations was only included in the 2013 survey. One organization did not rate themselves on cost effective analysis in 2010.

Since we did not include the question on policy simulations in earlier years of the PO survey, we asked an additional question on whether improvements in competence in policy simulations were a result of the project's capacity building activities on the subject. In Table 3.2.13 we see that half of the 14 organizations reported a fundamental increase and all reported at least a modest increase in competence in performing policy simulations. Combined with the results in Table 3.2.12, we may be able to infer that the POs had very little competence in policy simulations prior to the project given that half of the POs believed their competence had increased fundamentally yet less than half reported at least normal competence levels by project end.

In the bottom panel of Table 3.2.13, we asked POs to rate their increase in knowledge of advocacy and communications as a result of the GDN presentations. We found mixed results in each of the survey rounds, but POs were mostly optimistic by project end with 57 percent reporting a large increase and an additional 14 percent reporting a fundamental increase in knowledge.

Table 3.2.13: Self-Ratings of Improved Competency

	Fundamental Increase	Large Increase	Modest Increase	Negligible Increase
How would you rate the increase in your organization's competence in policy simulations as a result of GDN PEM presentations on these topics?				
2013	50	21	29	0
How would you rate the increase in your organization's knowledge of advocacy and communications abilities as a result of GDN presentations on these topics?				
2010	14	29	43	14
2011	7	29	57	7
2013	14	57	14	14

Conducted chi-squared tests for relationships between survey round and ratings - no significant relationship found.

Additionally, POs were asked as to whether they could name specific changes to their communications practices as a result of the project. The results are shown in Table 3.2.14. The percentage of POs responding that they had made communications changes as a direct result of the project rose steadily from 57 percent in 2010 to 77 percent in 2013. The descriptions of the changes also resonate well with best practices – especially better targeting of communications and the use of more appropriate communications tools. This result is indicative of the overall positive impact that the project seems to have had on strengthening the capacity of POs.

> "We have also managed to do disseminations in a decentralized setting, like at the local government, which we normally with other projects don't do because of lack of finances. I think during the mid-term evaluations where we were shown who are your policy communities…members of parliament, local government. We were indirectly forced to reach out to this particular category – members of local government – more than we normally do. We typically focus on members of parliament, people who relay at the center, like in the ministry. So through the GDN project we were able to go down to local government more than we previously managed to do." (Africa)

Table 3.2.14: Specific changes to communications practices

	2010	2011	2013
Can you name a major change your organization made to your communications practices due to what you learned at the GDN events?			
Yes	57	71	77
No	43	29	23
Descriptions of changes			
	Instituted an annual seminar where public expenditure issues are discussed, including those on the three sectors of education, health & water.	Develop a National contest among young economists to spread the use and knowledge of these techniques.	We started a new way of communicating our research…we have a dynamic relationship with citizens, especially with those college students and recent graduates that participated in [our program].
	There is a media consultant and he also reviews the papers prepared and presented at the workshop as well as conferences.	Realization of the need to work more closely with the communications unit/public agenda newspaper to get articles published.	Updated communication to the policy makers and concise write up distribution to media
	Well we are changing the website, hiring a new communication expert to work with us.	We received useful information on how to write press releases	Active interaction with policymakers
	Preparation of communications plan during development phase of projects.	Prepared policy brief on program based budgeting. Also organized seminar to discuss budget best practices and service delivery.	The launching of an online dashboard monitoring development outcome
	Engaged with greater understanding on budget expenditure and its processes among government officials in the sectors of health, education and water through workshops and conferences.	GDN's project is a little more flexible than other donor projects in terms of ability to adjust the topics/contents of proposed research activities. Such flexibility allowed us to start employing new mode of pre-research communication with beneficiaries	Writing down a communication plan.
	It has changed our view of how to communicate our findings and how important it is. Besides, it had helped us to understand better the cost-effectiveness analysis.	More Aggressive in inviting members of the house of representatives (congress).	Our organization now relies more on 'Policy Brief' than on 'Working Papers' to communicate study outputs to relevant stakeholders. 'Policy Briefs' are found to be more effective in making the target audience understand our points quickly and efficiently.
	Address issues systematically, convince with reasons	Policy briefs, two pagers	Our communication approach is now more systematic and grounded on research results than it used to be.
	Learnt about pitfalls to avoid when granting interviews to media networks	The importance of targeting clue agents.	Learned how to prepare actionable policy briefs
		A formal communication expert was recruited.	We started developing 2-pagers and policy briefs for our research findings.
		We gained greater competence in designing policy briefs.	

POs were also asked to comment on each of the analytic tools learned for the PBA, BIA, CEA, PS and the communications training of the program. These were open-ended questions, and organizations were not required to enter a response.

For two POs, PBA was a completely new tool. In addition to the tool itself, lessons learned among other POs included how to disaggregate analysis into wage and non-wage expenditure, handling budget data at the district level, and presenting analysis results. Exposure to BIA further improved understanding for two POs, as they learned how to link findings from PBA and BIA. Three POs developed the ability to triangulate BIA data with national survey data, which allowed for disaggregation of BIA analyses by income level and other demographic characteristics.

Four POs mentioned that although CEA was not a new tool, the project afforded an opportunity to apply CEA to evaluate social programs. One PO also used the tool to analyze government taxation practices. Another PO mentioned that while CEA was useful for the project, they did not anticipate utilizing it for others. This observation was also mentioned by one PO for the PS tool. Six other POs found the PS to be very useful. Reasons mentioned include addressing limitations in data, sensitivity analysis, and identifying program alternatives that match objectives.

Eleven POs commented that the communications training helped identify stakeholder engagement methods. Two POs mentioned that their skills in preparing policy briefs were honed, while others mentioned that they gained knowledge on how to identify policy audiences, write press releases, and utilize social media.

POs spoke of acquiring new capacities due to the project in their in-depth interviews. Five POs reported that they were largely successful in building their institutional research capacity. An African PO observed that even new and junior staff had become well-versed in the technical skills associated with the project. An African PO and two Latin American POs commented that this project had afforded them the opportunity to develop a strong understanding of the analytical tools. In contrast to these five POs, an Asian PO remarked that their objective of developing expertise in PEM had not been fully realized since their final research products were incomplete and they needed to further engage policy stakeholders.

POs also spoke of acquiring experience in new sectors and focusing on new topics within sectors of their current work. A Latin American PO remarked that the project allowed them to improve their health and education sector work, and provided an opportunity to work in the water sector for the first time. An Asian PO mentioned that they have been invited to join a working group to draft legislation for the water sector, as a result of engagement with policymakers.

- *"This was a first for the organization; although some staff had experience independently with PEM, this was the first encounter of [organization] with PEM. [Organization] did not have any experience or expertise before this project in public budgets and public spending… One objective was to build capacity through the project, to be able to serve to those counterparts as well. We have not achieved our objective 100 percent. Our recommendations are at the stage where they are yet to be more fully communicated and policymakers need to be more engaged, but so far, we have reached out to many stakeholders, in water particularly. We have had serious success in getting established as a reference in PEM." (Asia)*

- *One of the objectives of this project was capacity building, and we believe this has to a certain extent been realized. I was leading a team of around 4 people, many of whom had just joined. Over the past 2 years, I have been able to see that I give them instructions and they are able to just do it, because of working with GDN. So the capacity building has been very good…In research, there is usually an issue of free-riding. You have 3 names, you send in the report. But once we came to these regional workshops, we were meeting with the Technical Advisor, on a sector basis, one-on-one basis. Even if you had done this as a group, you as an individual had to know what you're going to say." (Africa)*

- *"The fiscal policy area in budget analysis was very strong, but the education and health areas were not so strong. We felt this project allows us to strengthen our capacity and see how these techniques worked in our area. In terms of water, it enables [us] to start a new research line that we had not done before. Yes, with different degrees, the education sector in particular has benefitted a lot." (Latin America)*

- *Water is a new sector as far as our group is concerned; but we have been doing some advocacy in health and education…We are more confident that we can go farther in health and education. What was very encouraging is we realized there is a very big gap in putting together water sector advocacy. When we were engaging policymaker, she expressed her desire to involve us in the technical working group in drafting [water legislation] that needed to be put in place. This is something we wanted to do and is an opportunity that is waiting for us. We would like to see our institution as a 'go–to' for these three sectors. We would like to build up expertise among team members and image.." (Asia)*

Members of the policy community were allowed the opportunity to provide additional feedback on the POs as part of the PCS survey. While policy stakeholders largely commended the POs on producing high quality research, they cited a number of challenges and ways in which the POs could improve. Challenges mentioned included insufficient funding, turnover in technical staff and lack of research visibility.

Indirect Beneficiaries of the project. The indirect beneficiaries are the participants of the government programs whose targeting of beneficiaries and operational efficiency may have been improved due at least in part to the project analyses and constructive engagement of the participating organizations. The extent of policy impact as well as the quality of project products has been discussed in the previous section. To identify the programs under analysis, we reviewed the program budgeting, benefit incidence, cost effectiveness and policy simulation studies prepared by POs and listed the type of programs and the populations they serve. These reports, however, are universally at the sector level for health, education, and water.

In its guidance on evaluations, DFID expressed interest in several groups of possible indirect beneficiaries. Responses from the interviews with the 14 POs, as well as the Activity to Outcomes Matrix prepared by GDN, indicate that these DFID groups are explicit targets of analysis for several of those interviewed. Table 3.2.14 shows the pattern for the fourteen who were interviewed.

Table 3.2.15: Population Groups of Interest to DFID Targeted for Analysis by Participating Organizations[a]

<table>
<tr><th rowspan="2">Organization</th><th colspan="3">GDN Focus Sector</th></tr>
<tr><th>Education</th><th>Health</th><th>Water</th></tr>
<tr><td>AST</td><td>Children, youth,
Poorer income quintiles</td><td>Pregnant women</td><td>---</td></tr>
<tr><td>CBPS</td><td>Children
Girls-gender equality</td><td>Lower class castes-quality of access</td><td>Lower class castes-quality of access</td></tr>
<tr><td>CEDS</td><td>Children from poor families</td><td>Patients with HIV/AIDS</td><td>Rural populations</td></tr>
<tr><td>CIPPEC</td><td>--</td><td>Patients with catastrophic (non-communicable) diseases</td><td>Rural populations</td></tr>
<tr><td>CIUP</td><td>Indigenous children</td><td>Low-income populations</td><td>Rural populations</td></tr>
<tr><td>CRC</td><td>Children, youth</td><td>Children, youth
Elderly (insurance access)
Those with disabilities</td><td>Children, youth
Elderly (rural, access)</td></tr>
<tr><td>CSEA</td><td>Children, girls</td><td>Low-income, rural populations</td><td>Low-income, rural populations</td></tr>
<tr><td>EGAP</td><td>Youth, low-income populations</td><td>--</td><td>Low-income populations</td></tr>
<tr><td>EPRC</td><td>Children, girls</td><td>Patients with HIV/AIDS</td><td>Rural populations</td></tr>
<tr><td>FUNDESA</td><td colspan="3">Indigenous population, poor families with children</td></tr>
<tr><td>IEA</td><td>Children, youth</td><td colspan="2">Low-income populations</td></tr>
<tr><td>ISODEC</td><td>Children, youth</td><td>Pregnant women</td><td>Low-income, rural populations</td></tr>
<tr><td>PRAD</td><td>Girls, children belonging to marginalized populations</td><td>Pregnant women</td><td>Low-income, marginalized populations</td></tr>
<tr><td>US</td><td>Children</td><td>Pregnant women</td><td>Marginalized populations</td></tr>
</table>

a. DFID's MTR guidelines list the following categories defined by personal attributes: gender, ethnicity (especially minorities), socio-economic group, age (especially children), disability, and persons living with HIV/AIDS.

3.3 Economy

Our examination of project costs finds that expenditures were appropriate to the context. Table 3.3.1 shows the actual allocations of project funds across major areas during the April 2008 – December 2010 period. Nearly half of all spending is accounted for by grants to participating organizations. If one adds costs of technical advisors, 59.58 percent of total costs are covered. Costs on regional and global workshops amounts to 10 percent of the total. Usually the standard for M&E is that 3 to 10 percent of the project budget be allocated to M&E; with monitoring included as part of the project's activities and independent evaluations being specifically budgeted. [22] GDN is just over the midpoint of this range at 5.89 percent. Project management – direct project staff and travel – costs are at 9 percent and indirect administrative and institutional support is at 11.4 percent, within DfID guidelines of being no more than 15 percent of the total budget.

Table 3.3.1: Allocation of Project Funds by Major Category, April 2008 – September 2013 (percent distribution)

Category	Percent
Grants to participating organizations[a]	44.64
Expenses for analytic workshops, (global and regional)	10.04
Dissemination[b]	4.11
Technical advisor costs[c]	14.94
Project management	8.98
Administrative and institutional support	11.40
Monitoring and evaluation	5.89

a. Grants to the participating organizations also includes cost for research communications and dissemination activities undertaken by the participating organizations.
b. Dissemination excludes cost of dissemination activities undertaken by the participating organizations - these costs are included within their grants (see note a). Total resources devoted to dissemination and project communications are greater than the amount given within the table because the project communications is supplemented by other GDN resources.
c. Technical Advisors costs includes: 1. Cost for Technical Advisors + 2. Cost for R4D + 3. Honorarium for external experts at global and regional workshops.

Other findings on expenses are summarized in turn.

- The evaluator examined project budgets which show the daily rates for various staff members and consultants. The daily rates are well within the range of rates paid for persons of similar experience and education. The level of effort assigned for tasks is reasonable.

  The grants to POs supported their project research, staff time for attending workshops, and

[22] Project/Programme Monitoring and Evaluation (M&E) Guide. International Federation of Red Cross and Red Crescent Socities, Geneva 2011. Page 32.

dissemination activities undertaken by the POs. The standard annual payment to a PO is GBP 31,000, with some variance in special circumstances. The research effort required is substantial, i.e., for each of several types of analyses conducting studies for each of three sectors—health, education, and water. Additionally, a strong communications program is required with the objective of getting the analytic results used in the policy process. Based on NORC's experience in conducting research with local organizations in a half dozen countries, including several where POs are located, the project is getting good value.

- Workshop costs are definitely reasonable, based on NORC's review of workshop expenditures and attendance at several. All workshop participants traveled economy class, with travel arrangements made and tickets purchased by GDN. Events were held at good quality but modest hotels.

3.4 Efficiency

Management and Coordination

The project, while led by GDN, is a partnership between GDN and R4D. At GDN, Ms. Ramona Angelescu Naqvi provides overall management and oversight. Savi Mull initially served as a Program Officer, assisting with overall management. After she took on the role of Monitoring and Evaluation Coordinator, Kaushik Ganguly stepped into the Program Officer position. Mr. Ganguly was brought on due to his experience in PEM research and participation in project sustainability initiatives, where he was responsible for finalizing training materials, publishing data in a public domain and providing training to policy stakeholders. The project management reflected that because Mr. Ganguly's relevant technical background allowed him to contribute significantly, it would have been helpful to have brought him onto the team at the start of the project. Pooja Sarin, a Project Associate, works with the Finance Office on developing and overseeing the budget, workplan, utilization of funds, absorption capacity and other related tasks. Ms. Sarin spends around 60 percent of her time on this project and Mr. Ganguly about 30 percent.

On the R4D side, Ms. Courtney Tolmie over the first three years of the project was solely responsible for managing technical support and providing technical review. During this time, Ms. Tolmie was assisted by an administrative staff that coordinated logistics and correspondence. After the first three years, the need for additional technical staff grew and two colleagues were brought onto the team to assist in technical oversight; each of these staff spent around 30 percent of their time on the project. Although Charlie was originally intended to serve as a technical lead, due to various constraints he ended up serving as a technical advisor.

Management challenges and suggestions for improvement. Project management staff from GDN and R4D were interviewed on the partnership between the two organizations, management structure, challenges and feedback for improvement. Over the five years of the project, the management team faced and overcame some challenges and learned some lessons for similar future projects.

- *"This is a very complex project – we went in being very ambitious, and possibly being quite naïve. It took up some time to work out the kinks and allow for the flexibility in terms of substance and management. Where we are is in a really good place. If I had a do-over I would design it somewhat differently, but we did work out some of the issues over the course of the five years." (Project management staff)*

The main initial challenge was in defining responsibilities between GDN and R4D, which was also commented on by one of the POs. It was difficult to make technical changes since it was unclear which partner would make final decisions. Roles and responsibilities were eventually sorted out so that R4D was responsible for the technical side, while GDN primarily oversaw financial and logistical activities. By project end, the partnership was viewed very positively.

- *"At the beginning, there was technical competition between R4D and GDN. So you will find an email from [R4D], and then an email from [GDN] saying the same thing. But as time passed, GDN has more presence and R4D are only participating in the technical part. (Latin American PO)*

- *"Sorting through some of these roles and responsibilities at the beginning is valuable, and we did a great job of that. There is a real tension in having one partner be a technical lead and the other be the overall lead, because there's a real question about if you want to make technical changes, who gets to make the call. (Project management staff)*

- *"It has been a tremendous partnership. It was R4D who started working on the proposal, from there we did selection of POs, identifying mentors, overall direction of the project. It has been mutually beneficial. The initial vision was R4D focusing more on technical side – the content of the research, workshops, identifying technical advisors and so forth, and also sharing in the management of the project. It has worked well. GDN obviously focused on management of the project – financial management, logistical arrangements for workshops, etc. We have had project management calls at least once a month, sometimes even more frequent in lead up to events. It worked out very well with R4D on all levels.*

- *"GDN was a really nice partnership. It functioned extremely well. We got the support that we needed from the GDN team. They really gave us leave to lead on the technical support. The relationship initially was a little bit tougher – both teams were understaffed and there were some questions about roles. Probably for the first year to year and a half it felt very tense, but where we ended up getting was a really good place."*

(Project management staff)

There were also some challenges with managing the POs. Among the specific issues cited were long distance communication challenges that should have been mitigated by more in-person visits to partner countries, a lack of adequate technical resource staff on the management team to interface with POs, and challenges with the peer learning system.

- *"The quality of the monthly management calls was horrible – we tried Skype, Webex, software where we could all view the same Excel document. More on-site visits with POs would have helped.*
- *"We underestimate the amount of time and resources it takes to lead something like this. If I were to do it over, we would have people assigned to different partners, so they could have a sense on a regular basis of what the challenges were. I would rethink some of the peer learning management; we tested a number of options, but we never found the perfect fit."*

(Project management staff)

The 14 POs were asked about GDN's project management, on whether GDN was clear about what was to be done and adequate resources provided. They were also interviewed on the quality of communications from GDN on responsibilities, project organization, and events, as well as opportunities to provide feedback and changes made as a result.

Financial Management. Discussions with GDN regarding financial management showed that they followed standard operating procedures regarding grant management. A part of the annual grant was provided upfront to the POs with the remaining tied to specific deliverables. GDN also examined fund utilization statements from the POs and in the case of low or deficit balance processed payment on priority. In most cases, however, POs had surplus funds with delays in the absorption and utilization of funds. Four POs provided specific feedback on GDN's financial management practices. A Latin American PO commented that the limited funding failed to take into account the significant time spent on the project by senior staff. Another Latin American PO remarked that the difference in fiscal year between themselves and DfID posed a challenge. Two African POs noted that transmission of funds sometimes took time.

- *"In terms of resources, there was little money to do a lot of work. For every research document we have two or three researchers working a lot, and very seriously, some of them are junior but some are not; it is very costly to do all this research. At [organization] we tried to cross-subsidize; education and health sector to do this research." (Latin America)*

- *"My other colleagues say that resources took time to come. Our side, we operate in a different environment. Once we have a signed contract, this is the amount of money that has to come; money can be spent even before funds arrive. It wasn't a constraint for us. In other places might be a little bit constrained." (Africa)*

- *"Our fiscal year is different from DfID's; it generated some problems. The management of the project tried to help us, but it has been an issue during the project." (Latin America)*

- *The transfers for us to carry out project; we had to give deliverables before they gave us money. It became difficult; somebody would go out to the field, and would say, there's no money, there's no way we're going to be able to get it. How they transferred funds to us was a challenge." (Africa)*

In an interview on the financial system and practices, the project management indicated that since the POs were provided training and standard templates for reporting, there were no major challenges that emerged. The project management also stated that attrition was not a factor since financial staff at POs largely stayed with the organizations throughout the project.

- *"In the beginning of the project we had worked through the financial reporting with them; we worked on a template that the partners used to disburse funds for this project. Earlier on we had clarified that the financial year would be the DfID financial year. Attrition was less of a deal for the financial system since staff did not change too much over the years; it was mainly technical staff who changed. The DfID-imposed templates were pretty straight-forward. DfID had templates for partners on programmatic issues; for that there wasn't a problem at all." (Project management staff)*

Communications. POs largely remarked that GDN had effectively communicated to POs on responsibilities, project organization and events. Seven POs commented that communications from GDN had been professional and provided adequate notice of upcoming events and deadlines. An African PO remarked on GDN's flexibility and patience in adapting to PO challenges and local realities, which the PO stated was "very rare in projects of this kind." A Latin American PO remarked that GDN promptly responses to requests and questions despite the considerable time difference. Although commenting that the project involved a lot of work, an African PO mentioned that GDN staff would often email helpful articles and other resources.

- *"Overall, the management has been good. In terms of linking institutions and technical assistance, deadlines and reports, all of this has been very clear from the word go… For the projects on water and education, we underestimated how much time it would take to acquire data, especially for CEA and PS, which are very data intense. In all that we saw extreme patience from Ramona and her colleagues to allow us to recover and work towards it. Deadlines are important as well, but there is the understanding that there are peculiar challenges by country and sector. Adjustments that need to be done regarding project work were made for us by significantly stretching deadlines to enable us to deliver work; that is very rare in projects of this kind." (Africa)*
- *"Communication has been fantastic, noting that you are communicating with many different people. It was always prompt, with adequate time for decisions to be made regarding workshops and stuff that needed to be done. The communication by GDN was superb." (Africa)*
- *"About the events, regional and partner workshops, we were informed in good time. They were giving us adequate time to prepare." (Africa)*
- "It was very good. It was at the beginning the difference of thirteen hours, but they have been very good communication between the management team and [organization]. They respond to emails, phone calls in a considerable period." (Latin America)
- *"They do very well. Sometimes they even send us other articles, research documents. Sometimes they send a whole lot of work; there are also other projects we are on, so the workload becomes too much. But to some extent they try to send a lot of information." (Africa)*

Five POs identified challenges that arose due to particular communications practices. An Asian PO mentioned that despite repeated explanations about cultural protocol, it was difficult to encourage GDN to extend a formal invitation to policymakers. Two African POs commented that they were a number of times where they were notified of deadlines with very short notice[23]. An additional African PO indicated that there was a gap in communication between GDN and itself, brought on by staff turnover in the middle of the project and a lack of clarity on new responsibilities among new staff. A Latin American PO commented that there were occasional communication challenges regarding deliverables– there were strong outcomes when clear guidelines were provided or when POs were free to determine parameters; however, there was ambiguity in some instances, as in the case of policy simulations, which presented a challenge.

- *"They do communicate. Many times communication comes up pieces; not comprehensive. Largely no complaints, but minor irritants. Most recent is [bureaucrat] coming here. [Here], senior bureaucrats behave very differently. There are processes. It took me four emails to convince them that they need a formal invitation; an email they'll not even process it. But on the whole they've been good; I would not undermine that." (Asia)*
- *"There have been a number of times where they have asked for particular information at a very short notice. They request by Friday or Thursday and ask us to send by Monday. And to me, that assumed this was the only project we're working on. In some cases it appeared they're also responding to requests from above." (Africa)*
- *"GDN was clear at the beginning, but not with policy simulations. When we had to do other stuff like the communications activity, we were effectively freer and productive. We decided we wanted to make two videos – that was very useful for our organization. We were not as free as we thought we were and that brought a lot of friction. When it was clear but constrained it worked out well, when it was freer it worked out well, but when we thought it was freer and it was not, that caused more trouble." (Latin America)*

Involvement of local partners. Many of the POs who spoke of opportunities for engagement in project design and management indicated that they were afforded the opportunity to provide feedback in the regional and annual workshops organized by GDN, and that changes were made to improve the project as a result. Two Asian POs as well as a Latin American PO had suggested having regional workshops, which were implemented by GDN; one of the POs had also requested that policymakers be invited to the workshops which GDN also arranged. Along with other POs, the Latin American PO had asked that further support of TAs be provided, in response to GDN soliciting feedback on what resources would be helpful in standardizing deliverables across POs. Another Asian PO mentioned that GDN had allowed them to choose an alternative analytical method instead of CEA. A Latin American PO commented that GDN had provided a grant extension to change the scope of research. An African PO mentioned that the technical advisors provided an opportunity for POs to offer suggestions on improvement. Another African PO noted that feedback from POs to GDN to consolidate reminder emails was taken into account.

- *"We have provided written feedback. Earlier we have suggested to have regional level workshop, with a focus on technical issues; earlier it was not there. We are glad that they have organized a number of regional workshops." (Asia)*

- *"With all the workshops they had slides in last sessions asking for suggestions. It was in the beginning of the project that we were talking about regional workshops and then they started doing those. My idea was to involve specific policy and decision makers from each country to participate, and it was picked up. It was very prompt feedback." (Asia)*

> There was a grant extension in which they let us expand our scope of analysis in the way we thought it was better. They were very receptive and it actually happened. In every workshop they have a session to get feedback about the work we are doing and how to work better." (Latin America)

- *"In annual conferences they always have space for feedback. They're quite open to listen to me… in [health sector], we're not really doing cost-effectiveness – we said it's a very significant thing we want to do. It has elements of cost-effectiveness, and we do not want to leave this issue. And they agreed on that." (Asia)*

- *"After we completed PBA, BIA and CEA, GDN said that we have to improve analysis in order to complete the policy simulation. But that was a little confusion on that because there was no standard methodology; there were a lot of options that each country could choose. In order to have kind of standardization among countries, GDN proposed some options. One was peer reviewing between countries in each region, the other option was hiring a specific expert in each country in order to have more experience in the field, the third was having an international TA supporting the institutions in order to have a minimum standard of quality of work. And we believe in education that it was a better option, the support of TA was okay. Most of the countries also decided that was the better option, so management team decided to hire the TA. One of the suggestions we made was the regional workshops. We believe that it was better for all to have feedback from peers in the same region. GDN considered that; the evidence is that we had a couple of workshops in Latin America. GDN was very open about that suggestion and decided that it was okay for the workshop." (Latin America)*

- *"At one point there were a number of emails from GDN on deadlines and things that needed to be done; people were becoming overwhelmed with emails. It was mentioned that we should try and be moderate with communication, and subsequently there was a change. Whenever GDN received feedback in areas in which they could adjust, we saw adjustments promptly." (Africa)*

Dissemination of skills and knowledge among PO staff

The project was focused on improving the capability of the PO institution, rather than individual researchers. In order to determine whether the training of one or two staff at a project workshop is transformed into organizational capacity, it is important to identify the degree to which workshop attendees share information obtained in the workshop with others at the organization. If there is little sharing, the institution-level effects will be small, even though the individuals may have learned a lot and are using the knowledge effectively.

In interviews conducted in early 2013 the 14 POs were asked about whether their organization's involvement in the program affected the whole organization or only those researchers who were part of the project. The interviews revealed that key findings and knowledge gained from the project were disseminated to non-project colleagues in nine POs, while an additional five mainly focused on improving capacity of staff working directly on the project.

Eight POs' involvement in the program significantly affected colleagues outside of the project. In addition to sharing results with other staff, an Asian PO engaged non-project colleagues in data collection efforts for the project, ensuring that these staff received training in project-related topics. For another Asian PO, staff is hired based on sectoral expertise rather than specific project needs, so knowledge is disseminated regardless of project involvement. In slight contrast, a third Asian PO mentioned that there are guidelines in place to ensure research is presented to the organization; this PO has also utilized the PEM tools and subject areas for curriculum development for a Master's program. Three POs, one based in Asia and two others in Africa, organized capacity building seminars to disseminate knowledge within the organization. Another Asian PO rotates staff sent to GDN workshops; due to the small size of the organization, all staff members have an opportunity to attend at least one workshop. Similarly, a small Latin American PO remarked that it has involved all staff in the project. An African PO has widely shared project results in organization-wide seminars with other staff. Another African PO solicits feedback from other staff on final deliverables, and provides progress updates through a reporting system used by colleagues across projects.

- *"In small parts. Our organization still remains a research institute; we have not been transformed into a think tank or a solely PEM-focused research house. We do applied social research, and half of our staff have different interests and are not really interested in public finance. What was important for the others to gain was we gained a lot of insight into health sector, education sector, water. ..General results have been disseminated to other staff…There were 20-25 people who got trained* *over the life of the project on policy issues in education, health and water. Not on budget terms but in the social sense." (Asia)*

- *"We have a strict rule to disseminate research. Each paper needs to be disseminated in open seminar in our school…Part of the subject we use our tools in GDN as the subject of two class lectures in Master's program…Most of the students are coming from government." (Asia)*

- *"We tried to do validation with our organization. Even before we send our reports to GDN, we try and share it among our colleagues. We have an in-house meeting where we share the findings, and what they think about the report. We get a lot of suggestions from them. We have a reporting system; every week, we sit down to report on that. So most of them know about GDN." (Africa)*

For five POs, dissemination of knowledge gained from the GDN project occurred to a limited extent. A Latin American PO remarked that other staff indirectly learned about the project tools and research. Another Latin American PO explained that while it was the organization's policy to present research to other colleagues, competing project demands and a busy work schedule make it difficult to adhere.. An Asian PO also engaged students as well as faculty members of the affiliated university, but to a limited extent; the PO did note that as a result of sharing knowledge, GDN has become well known in parts of the wider university community. For one African PO, knowledge gained has only been shared with project researchers.

- *"We had a minimum of two researchers from three different areas of [organization] involved in the project, at least, at any given moment – this does not count people from communications, administration, etc. Learnings were not directly shared; they were shared indirectly." (Latin America)*

- *"What we try to do in CIUP is once you do your research you share [it with] your colleagues, and its part of the capital of the research center. I would say that partially. We have several demands; you are working on several topics at the same time. Maybe we were too focused in complying with the model of this project; we didn't make sure that other people get benefit from that. The information is shared; everybody knows what other people are doing, and we discussed the results." (Latin America)*

- *"Researchers, plus a little more – including graduate students involved. We attempted to involve faculty members of the other school through workshops, and other institutes as well, to give orientation to their graduate students. A very limited engagement. We get published in inter-*

university communications. We have a webpage. GDN brand has been imprinted in the minds of people, especially in the [university] and in [organization]." (Asia)

Table 3.4.1: Knowledge Sharing by Workshop Attendees

	May-09			Jan-10			2011				2013			
	Yes, we did this	No, because staff know about this already	No, for other reasons	Yes, we did this	No, because staff know about this already	No, for other reasons	Yes, we did this	No, because staff know about this already	No, for other reasons	There was no communi-cation component	Yes, we did this	No, because staff know about this already	No, for other reasons	There was no communi-cation component
Analytic Presentations														
Distributed the materials obtained or a list of them to other researchers	86	7	7	92	0	8	86	7	7	--	93	0	7	--
Led an informal discussion about the event's topics	85	15	0	92	8	0	93	7	0	--	92	8	0	--
Gave a formal presentation on the event's topics	45	36	18	50	17	33	36	14	50	--	50	29	21	--
Led an informal discussion about the event's second topics	85	15	0	--	--	--	--	--	--	--	--	--	--	--
Gave a formal presentation on the event's second topics	58	33	8	--	--	--	--	--	--	--	--	--	--	--
Communications Presentations														
Gave materials on communications and impact strategies to your communications team	54	8	34	58	0	42	21	0	36	43	71	7	21	0
Met with the communications team to explain what the sessions on communications strategies were about	57	43	0	58	0	42	36	0	21	43	57	14	29	0
Other	100	0	0	100	0	0	100	0	0	--	100	0	0	--

NOTE: In May 2009, the first topic was program budgeting analysis and the second topic was benefit incidence analysis. In January 2010 the topic was cost effectiveness studies. In 2011, organizations responded about their most recently attended event. In 2013, the survey simply asked about "the GDN workshops and conferences".

Data from the monitoring surveys further highlights dissemination practices among POs. Table 3.4.2 summarizes the information on knowledge sharing within the organizations reported by POs in the yearly monitoring surveys. The questionnaire included the list of possible actions that could have been taken shown in the table and included pace for other action types to be included.

Substantial sharing of knowledge gained on research techniques is evident at nearly all POs. The incidence of sharing materials and leading informal discussions has consistently been in the 85-93 percent range throughout the project duration. In contrast, only 50 percent of organizations have given a formal presentation on workshop topics from November 2011 through the end of the project, a similar amount as between 2009 to 2010, but a substantial increase from 2011. The primary reason for a lack of formal presentation of workshop learnings from 2011-2013 was that staff were already knowledgeable about the topics.

The pattern for the content of communications events at the workshops widely differs. In 2009 and 2010, about half of the POs passed on to the communications staff materials from the "best practices in communications" sessions, and met with them to review lessons learned. There was a sharp decline in these practices in 2011, to 21 and 36 percent respectively. From November 2011 through the end of the project, there was a substantial increase to 71 and 57 percent respectively. In contrast to knowledge sharing from analytical sessions, the majority of respondents identified "other reasons" as preventing dissemination of communications strategies. As a note, in 2011, the primary reason was lack of a communications component.

Staff Turnover. Two African POs and one Latin American PO remarked that staff attrition presented setbacks in completing project deliverables. There was a complete turnover in staff involved at one African PO, creating a significant learning curve for replacement staff who were not properly initiated into the project or aware of specific responsibilities. For another African PO, staff who had attended workshops organized by GDN left the organization without sharing knowledge gained; although GDN tried to help, it was difficult for remaining staff to manage expectations. A Latin American PO commented that it was challenging to prepare for transfer of responsibilities when project staff moved on from the organization.

- *"I came onto this project in 2010, when previous people working on this had left this institution for other opportunities outside the country. Because of the inflexibility of not being aware of what GDN wanted from us, we were heavily penalized. First, we didn't know what kind of information was supposed to be presented or was required. Because of that, we took a lot of time in responding. They* *cut our budget in half in 2011. It was really a tough six months for us. It helped those of us working on the project for the first time to wake up. There should be some handing over, which was never done." (Africa)*

- *On [organization] part, there has been major attrition. People who came to the workshops left the organization, without knowledge transfer. We had to put our legs in their shoes and transfer responsibility. We did not participate in conferences so we did not know what had happened, who to go to, so it became very difficult. GDN was trying to help in a way, but it's about communication. How we understand it becomes different from how they understand – there's a gap in communication." (Africa)*

Level of satisfaction with POs. Interviews with the project management provided feedback on partner performance and detailed characteristics of the 'ideal partner'. While noting that the research produced has largely been of a high quality, the project management team mentioned various challenges, including the pace of submitting deliverables, the difficulties in developing training materials for 14 organizations with differing analytical and communications capacities, and the small number of partners who were not invested in the project.

- *"I have been overall a little bit disappointed with the speed at which they were able to deliver. I think they had data challenges in accessing budget, household and program data. It is partly reluctance because these are new methods they're exposed to, and staff turnover and other things. I am pleased with some of the work that has been done in this project – it is extremely useful for policymakers and is well articulated, documented and well done. Some of the policy simulations were quite innovative, for example the Mexico one on covering some of the health expenditure burden caused by smokers by additional taxation on smoker and producers of cigarettes. We had a good group; a mix of some organizations that fall at the academic end of the spectrum and NGOs with a natural disposition towards advocacy. Having a mixed group also meant that devising the workshop and training was difficult. The cross-learning between POs was really useful."*
- *"With a lot of partners I was extremely satisfied, they were very engaged. They ended up doing work that had policy impact in their countries, which is what we wanted to do. There were a smaller number of partners that really never got engaged in the work, they did the bare minimum and were hard to communicate with. That was maybe three partners. Most of the partners I was largely satisfied with." (Project management staff)*

In addition to varied challenges that arose with the 14 final POs, PEM management also provided the rationale for dropping ESRF in Tanzania from the project. Their reasons included delays in producing deliverables and refusal of technical support to address these challenges, lack of response to communications from management staff and failure to attend one of the workshops.

- *"Lack of performance basically. For example, there would be six months delay from everyone in the project in terms of getting deliverables. No responses to offers to provide additional technical support. At regional workshops, the Tanzanian team was there but even during the workshop would leave early and were supposed to present but wouldn't. Lack of communication with us. At some point they had one guy who was trying to do all three reports for all sectors. They also missed one workshop because they misread their tickets, so they missed some of their training."*

- *"It was a tough decision, but we connected a lot about this. We just felt that we kept giving them multiple opportunities, but they weren't rising to the challenge."*

- On the ideal partner, the project management stated that a well-suited organization would be experienced in managing similar grants, and would respond in a timely manner on financial reporting requests and communications sent by project management staff. In terms of technical background, PEM Management indicated that in addition to experience in the sectors, the PO would already have skills in communications and analysis with potential for capacity building in these areas. They also mentioned that in hindsight, the program could have been strengthened by ensuring that partners were investing in the PEM research through allowing greater flexibility in choosing topics that align with organizational objectives.

- *"One of the criteria is we would want to work with an organization that has a certain base capacity to do research – they weren't going in completely blind on the sorts of things they want to focus on. Ideally they would still have a ways to go to really strengthen their program, and maybe have some concrete ideas about how to do that. We would also have a partner that was really bought into the idea of strengthening their credibility and capacity …If doing over again, we could improve the partners by changing the structure to get them to buy-in even more. I think the constraint was we really dictated a lot of things…It make peer learning and training a lot easier. It also meant that partners weren't necessarily doing things that were valuable for their own research agenda."*

- *"The ideal grantee would be very professional in terms of grant management – funding, reporting, and communicating with us. Somebody at midpoint of spectrum between heavily academically oriented and heavily advocacy NGO, so that we can strengthen their analytical, project management, and communication skills but so that we don't start from scratch with anybody. Someone with exposure in all three sectors – education, health and water.*

Were the risks properly identified and well managed?

GDN has paid close attention to possible risks that could be faced by the project; defining risks annually and taking risk mitigation measures. Starting with the project concept note that identified 4 potential risks, each annual report rated the likelihood and impact of the risk, and a mitigation plan. Starting in the June 2010 annual report a distinction was made between external risks – risks beyond the direct control of the project, and internal risks – those that could be addressed by management and technical processes. The external risks identified by GDN have been discussed in the Relevance section. Table 3.4.2 shows the internal risks, rating and mitigation measures.

GDN also did individual partner risk assessments each year, however, more assistance could have been provided to work with partners regarding their individual goals and outcomes. As stated in one of the qualitative in-depth interviews: "At the beginning of the project, we thought through the risks, and did that together…..One of the things that we didn't work on as much as we needed to - we needed to have specific goals and outcome measures for particular partners. It definitely helped us in providing technical support." One unanticipated results due to this was the exit of ESRF, Tanzania from the project due to poor performance. This has been discussed in more detailed later under this section.

Internal Risks: The risks defined are clearly reasonable and the mitigation plans realistic. For example, to address the risk of not adequately monitoring 14 POs in as many countries (no.6) the Project Management Team had one conference call with each PO in two months, in addition to tracking their deliverables. These were implemented in response to POs request for more interaction with GDN-R4D. Importantly, all team members participated rather than just the two persons attending the workshops. This meant that all members could actively participate and help with staff continuity, helping manage the staff turnover risk.

GDN added a full time project associate in early 2009 that kept careful track of deliverables due from partner organizations, organized the bi-monthly phone calls and handled communication and other day-to-day operational matters with the partners.

Turnover of PO staff working on the project has been a significant issue. One GDN response was to fund two analysts from each PO to attend all workshops so that if one analyst departed continuity would be maintained. Later this was increased to three members from each PO. In some cases partners were advised to strengthen the capacity of the team; and some partners ensured that there was overlap between outgoing and incoming staff for a smoother transition.

Table 3.4.2: Evolution of Defined Risks and Mitigation Plans over the Project's Life

Risk	Rating	Mitigation Plan
1. Public expenditure analysis requires specialized skills, and partners may require support to rapidly develop these skills, particularly the ability to do timely analysis.	Likelihood – High Impact – High This risk is symptomatic of the need for the proposed Project. The fundamental purpose of the is to lower the likelihood that CSOs have not developed the skills necessary to monitor and influence the budget process.	By pairing financial support with implementation-oriented technical support, local partners will be able to develop skills to become more effective organizations.
2. Good, high quality information may be difficult to access in low- and middle-income countries.	Likelihood – Medium Impact – Medium While access to information is still an obstacle in many developing countries, other international actors (such as the IBP and TI) are making important progress in this area.	The lead organization will provide guidance to partners on accessing information, but will largely rely on existing partnerships with IBP and TI in situations where more openness is needed. We recognize that a minimum level of transparency is necessary for partners to have any impact, and environment for budgeting is part of the selection criteria for partner organizations.
3. In some countries, partisan divides may limit some public officials' willingness to interact with or listen to organizations supported by the program.	Likelihood – Medium Impact – Low Although this is a potential problem in some areas, organizations can still develop their capabilities to analyze and engage for improved public expenditure practices.	One of the main objectives of the program is developing rigorous policy analysis skills of the selected organizations, resulting in a set of members and local partners with increased credibility when engaging with public officials.
4. The work of partners in some regions could be opposed by the governments of those areas. **Dropped in 2010 - 2013 Annual Reports.**	Likelihood – High Impact – Medium In some countries, there is active government opposition to stronger organizations, particularly where partnerships exist between international and domestic organizations.	Resistance to civil society strengthening is diminishing in general, and the recent trend of democratization (particularly in sub-Saharan Africa) has led to a more tolerant and often encouraging environment for strong organizational involvement. Further, we do not currently expect to work in a country with a government actively opposed to such interventions.
Below Added: 2009 Annual Report		
5. Local partners have a relatively high turnover of staff.	Likelihood – High Impact – Medium This is a problem that partners have faced and potentially can limit the institutionalization of research capacity.	High turnover remains a problem. Instead of one, two members of project teams will be invited to the trainings to ensure continuity and integration of technical capacity in the partners. An intranet site provides all project related information to all project members. Some country partners in African are part of the Think Tank Initiative which has core staff in place. Partners will be encouraged to prepare for such exigencies by working with core staff and planning ahead of time.

3.5 Effectiveness

Achievement of objectives. The four project objectives are listed in Section 1 of the report. Our analysis presented in Section 3.2 on the impact of the project addresses the first two objectives and shows that there has been significant progress towards the achievement of the first and limited achievement of the second. The quality of report writing has been increased, communications practices have been somewhat improved, and there have been some impacts on policies in some countries. We address the remaining two project objectives below and show that the third objective has been only partially attained and the fourth has seen limited success based on the initiative of the POs rather than project implementation.

Development of internationally comparable information on public expenditures, benefit incidence, effectiveness, and policy alternatives that will begin to build benchmarks for the quality of public spending. The project partly achieved this objective. It developed benchmarks based on the benefit incidence analyses in Year 2 and these were presented and discussed in the global workshop in Prague in January 2010. While the project has produced a series of standard reports on program budgeting and benefit incidence in the health, education, and water sectors in each of the 14 countries, for the cost effectiveness analyses and policy simulations each PO focused on a very particular issue within each sector, making it challenging to develop international benchmarks. The project is also in the process of collecting data for an online Knowledge Portal in the public domain that will contain comparable data from across participating countries.

Creation of a strong network of institutions to share training materials, templates for analysis and communication. The project has been somewhat successful in creating a strong network among the POs. A strong network requires interaction among the members whereas there has been limited interaction among POs over the course of the project period as shown in Table 3.5.1. The majority of POs had contact with only one or two other POs and this contact was limited to only between one and five exchanges over the period of a year. Only one PO (7 percent) was proactive and interacted with more than half of the other 14 POs in the program.

Table 3.5.1: Interaction with other GDN PEM participants

Survey Year:		2010	2011	2013
Time Period:		Since beginning of program	April 2010 – November 2011	Since November 2011
During the time period, how many of the other GDN organizations did you have any type of contact with outside the organized workshops?				
0		14	14	14
1—2		57	57	71
3—5		21	29	7
5--10		7	0	0
10+		0	0	7
IF 0:	**Why did you not contact other GDN organizations?**			
	Did not have any questions	100	50	20
	Did not think it would be useful	0	50	18
	Used other resources instead	50	0	18
IF >=1:	**What is the total number of exchanges you had with others in the GDN network about the project?**			
	0	0	0	0
	1--5	92	85	86
	6--10	8	15	7
	11--15	0	0	0
	15+	0	0	7
	What is the total number of exchanges you had with others in the GDN on matters other than the project?			
	0	25	38	31
	1--5	75	62	54
	6--10	0	0	8
	11--15	0	0	0
	15+	0	0	8

The underutilization of other POs as resources is acknowledged by three of the POs. The reasons for this seem to be a lack of knowledge of other POs' work and subsequently a lack of knowledge of possible areas of interest for exchanging information. One PO pointed to the webspace, which did not succeed as a mechanism for sharing information, a second felt they did not create enough of a relationship with the other POs to share beyond the workshops , and a third felt it was due to the divergence in topics among the POs after the first year.

- *We don't know what other countries are doing. We just noticed in the recent workshop at Jakarta that there's a lot of information, interesting analysis from other countries that we didn't have before the workshop…One of the goals of the project was to encourage networking, collaboration among partners between the workshops and shared web space. The workshops did happen, but project web space had a little bit of low achievement level. It is missing a little bit of networking, sharing of information, making the analysis available in the public domain – it is something that is lacking right now (Latin America)*

- *We have underused our PEM partners – we could have exploited that more. There is a capital there that we have not exploited enough… Maybe we needed all these years to know each other. (Latin America)*

- *GDN formed a good platform using internet and shared web space, but we did only a couple of ad-hoc requests to specific countries where we knew that some of the things have happened… Only the first year of the project was quite uniform in terms of results. All countries did the same job. But then, after the first year, everyone started doing different things. For the benefit of our policy paper, I don't see any similarity to other countries. The exception was water, with tariffs, which is why we contacted Latin American countries." (Asia)*

Despite the limited level of interaction among the POs in the project, the POs did feel that the project had made some contribution to their international network. Table 3.5.2 shows that 36 percent of the POs felt they had greatly expanded their international networks and half felt that they had somewhat expanded their network.

Table 3.5.2 Impact of the project on international network

What was the impact of the GDN PEM program on your organization's network of peer research and/or advocacy organizations in other countries?	
I have greatly expanded my network	36
I have somewhat expanded my network	50
I have very modestly expanded my network	14
I have not expanded my network at all	0

In addition, POs talked about how the project has helped them expand their domestic networks through formal partnerships with other NGOs, dissemination events of their project research, and reaching out to organizations to obtain needed information. The majority of NGOs (79 percent), collaborated with domestic or international NGOs to complete their project research and only two NGOs reported they did not involve others.

- *"We usually do networking and collaboration with other NGOs, particularly at the local level. This project has facilitated that even further, mostly in the shape of shared seminar, more systematic peer reviews, more chance to discuss our research. While research is ongoing, to have opportunities to discuss viability, opportunities, difficulties and costs associated with proposals." (Latin America)*

- *"We had the chance to work with community organizations, like water organizations where they have some advocacy on water supply. They are sharing their experiences, and we are sharing our tools. We were able to better understand issues around accessibility.". (Asia)*

- *"We worked with World Vision on the education part of the analysis, because they have nationwide program rolled out in the districts. This was the major NGO we worked with. In the data collection, they have local staff, the ones that are based in the district. We used these for data gathering as well as trying to explain to the schools what our research was and why we were trying to collect this information." (Africa)*

- *"Especially the PBA and BIA dissemination, we invited NGOs working in health, education and water. We also sought their inputs. We realized that most of them were coming to us for more information, especially the education network." (African PO)*

Program Approach. The program approach was generally effective and appropriate to obtain the desired objectives. The majority (64 percent) felt that the project was mostly suited to their needs with 29 percent feeling it was perfectly suited as demonstrated in Table 3.5.3.

Table 3.5.3 Opinion of appropriateness of project approach

To what extent did you feel the GDN PEM program was tailored to fit your GDN PEM team's needs to be able to complete reports of the quality required and effectively disseminate the information?	
It was perfectly suited to our needs	29
It was mostly suited to our needs	64
It was somewhat suited to our needs	7
It was not well suited to our needs	0

If we look at each of the components of the program separately, the training workshops, mentoring program, and peer-to-peer reviews all received a rating of good or excellent as mechanisms for building capacity by the majority of POs whereas the shared project webspace was only rated as fair by the majority as shown in Table 3.5.4. This critique of the webspace is elaborated on by the Latin American PO quoted above regarding the fact that it did not facilitate networking or sharing of analysis among POs.

The peer-to-peer reviews received the highest percent of an excellent rating (43 percent), followed by training workshops (36 percent), and the mentoring program (29 percent). Looking at the specific capacity building providers in Table 3.5.5, there is a positive overall picture with about equal numbers of POs feeling they learned the most from those who led the training and workshop sessions, peers and the technical advisors – the exception here is the communications resource persons.

- *"The way these partner meetings have been structured, has been very very positive. First, in having your peers review, getting technical assistance and following up with technical advisors after we left the event that has been extremely useful… We benefitted intensely from peer review and sessions for technical assistance." (Africa)*

Table 3.5.4 Rating of capacity building mechanisms

How would you rate each of the mechanisms for their effectiveness in teaching new skills and building capacity?	Excellent	Good	Fair	Poor	No Response
Training workshops	36	50	7	0	7
Mentoring program	29	43	14	0	14
Peer-to-peer reviews	43	36	21	0	0
Shared project webspace	7	14	57	14	7

Table 3.5.5 Rating of capacity building providers

If you had to choose one, from whom do you feel you learned the most from the GDN PEM program?	
Peers (other grantees on the project)	29
Technical resource persons (those who led a session or presented at the workshops)	36
Communications resource persons	7
Technical advisors assigned to you by GDN	29

Specific feedback given on the program components is summarized below.

Training workshops: There was no consensus on recommendations for the training workshops. A few mentioned that the quality of the first training workshop was not very strong but that GDN learned and improved after every workshop. Most spoke of the communications training. Some felt too much time was devoted to communications training, others felt that they gained significantly from building their communications capacity. Overall it seems that each PO has different capabilities and training needs and would have preferred more on their areas of weakness and less on their areas of strength.

Peer-to-peer reviews: POs felt that they gained a lot from the peer reviews and they would have liked more structured interaction with other POs than what took place during the workshops. POs are highly valued, specifically in sharing experiences on common issues, exchanging research, and possible future collaboration on research projects. In terms of peer learning, some of the researchers said that reading the papers of the PEM partners was more useful sometimes than comments they received from PEM partners.

Mentoring program: There were mixed reviews on the mentoring program. There was mention of difficulties in addressing comments from three different reviewers with different opinions. Several mentioned that Technical Advisors should have been more available and there should have been consistency in their involvement from the beginning of the project. The quality of Technical Assistance provided was varied across sectors, some felt it was excellent whereas one PO complained about having PhD students as reviewers for some reports, one felt there was a bias towards econometric analysis, and two felt that the Technical Advisors did not have enough understanding of the local context - they would have preferred to have local experts as reviewers. The project did include local experts as technical experts at the regional workshops and linked a PO that was struggling with its analysis to a local Technical Advisor.

3.6 Equity

Participating organizations were the primary project beneficiaries. One of the objectives of their research was to conduct a benefit incidence analysis and determine whether vulnerable groups were benefitting equally from government programs. One way to obtain an idea of the likelihood of vulnerable groups benefitting from the PO research is to examine which programs the POs ultimately chose to target in their policy simulation work. Table 3.2.14 in Section 3.2 provides a listing for each PO of DFID priority population groups served by programs being analyzed. The extent they have been able to influence and improve programs and policies that effect women, children, the elderly, minorities and other vulnerable groups by impacting policies is discussed in Section 3.2 on Impact.

3.7 Value for Money

Value for money (VFM) is defined as the optimal use of resources to achieve intended outcomes. It does not necessarily involve the use of the cheapest option; its focus is on obtaining the maximum benefit from available resources. VFM is therefore the sum of three components: the economy, efficiency and effectiveness of the project in achieving results. The table below explains how we have used these terms in our VFM assessment. The breakdown and analysis of indicators under each of these components helps us answer the three VFM questions in the review criteria matrix.

Table 3.7.1 Rating of capacity building providers

Success Factors	Indicators
Economy: Cost of inputs to achieve both outputs and outcomes	
Appropriate allocation of project funds across various categories	-project funds spent on M&E -project funds spent on management and administration
Efficiency: Value from combining inputs into outputs and outcomes **Operational efficiency: Amount spent on activities and outputs**	
-Regional breakdown (Asia, Africa, Latin America) of costs across grants and technical workshops and TA for POs -Timeliness and appropriateness of activities	-Regional breakdown of project funds for regional and global workshops, grants, and presentations at external events -Willingness of policy community to accept input from think tanks -Frequency of use of data/research and analysis by policy community -PO perception on if their objectives in participating in the project have been achieved (qualitative)
Allocative efficiency: Amount spent on achieving intermediate outcomes	
-Improvement in the skills and capacity of POs -Improvement in the quality of reports produced by POs	- Improvement in PEM skills and tools by POs -Percent of POs responding that the project was perfectly suited to their needs to complete high quality report and effectively disseminate the information -PO ratings of PEM mechanisms for their effectiveness in teaching new skills and building capacity -Change in report scores over the project period
Effectiveness: Achievement of outcomes	

Success Factors	Indicators
-Changes in policy: both demanded / requested by policy community; and supplied by the partner organizations	Demand side -Consultation by the policy community with research or advocacy organizations on policy issues -Receptiveness of the policy community to information; reaching out to the PO for a general policy discussion or to learn of the results of their work -Use of information in the research and analysis reports by the PC -PC perception on research and analysis done by the PO Supply side -PO perception on if their work has been used by the government or parliament -Continued research on transparency and accountability projects by the POs after the project finishes

Economy: As discussed in the section on Economy, Table 3.3.1 shows the actual allocation of project funds by major categories which are all reasonable.

Efficiency: Both operational (amount spent on activities and outputs) and allocative (amount spent on achieving intermediate outcomes) efficiency are examined. Table 3.7.2 shows the regional breakdown of project funds for regional and global workshops, grants, and presentations at external events. Keeping in mind that the Asian countries number 6, while the African and Latin American countries are 4 each there are some variances in the amount spent in each region across various categories compared to the total regional breakup which is more or less uniform. However, since our evaluation design is not able to attribute the outcomes measured directly to the project it is difficult to ascertain whether these variances affected outcomes among the POs. Spending on regional workshops and events for African partners is practically double that spent for Latin American partners. Technical advisor visits only took place in Asia (for one PO that was struggling with the analysis) and not at all in Africa or Latin America. Finally, only the Asian and Latin American partners requested and received additional funds for presentations at external events. These costs are linked to project results and discussed under the evaluation question: “Overall did the program represent good value for money?”

Table 3.7.2: Allocation of Project Funds by Technical Categories across Regions, April 2008 – September 2013 (percent distribution)

Category	Asia[a]	Africa	Latin America
Regional workshops and events	33	44	24
Global workshops and events	41	30	29
Technical Advisor visits	100	0	0
Grants to partners	43	27	31
Presentations at external events	58	0[b]	42
TOTAL	42	28	30

[a] Asia includes 6 countries: Armenia, Bangladesh, India, Indonesia, Nepal and the Philippines; while Africa and Latin America include 4 countries each.

b There was a joint regional workshop and presentation at the AERC Bi-annual workshop for African partners in June 2010. However, it is difficult to isolate cost of the GDN regional workshop from cost incurred to make the presentations.

To examine the timeliness and appropriateness of the project we look at three intermediate outcome indicators: (i) Willingness of policy community to accept input from think tanks (Table 3.1.2) – i.e if the country was receptive to this, (ii) Frequency of use of data/research and analysis by policy community (Table 3.1.2), and (iii) PO perception on if their objectives in participating in the project have been achieved. The first two indicators have been discussed in detail in the sections on Relevance and the third under Impact. As seen in Table 3.1.2 in the Relevance section, overall we see mixed results across the POs in the willingness of the policy community to accept input from the POs and use data / research / analysis. Statistically significant changes in average scores between the midline and endline survey are seen in only a few countries. An analysis of the above two tables to see if there is an improvement in scores between the baseline and endline as well as between the midline and endline are presented below.

Table 3.7.3: PO countries that show Improvement in Willingness by Policy Community to Accept Input from Think Tanks and Research-Advocacy NGOs and to Use Data/Research Analysis

Type of Improvement	Among Govt. Officials	Among Parliament Members
Improvement in Willingness to Accept Input		
Improvement between baseline and midline + Improvement between midline and endline	Kenya Mexico	Bangladesh India Indonesia Mexico Uganda*
Improvement in Use of Data / Research / Analysis		
Improvement between the baseline and midline + Improvement between the midline and endline	India* Indonesia Nepal Nigeria Peru* Philippines	Armenia Guatemala Indonesia Kenya Mexico

(a) T-tests have been performed to test whether the average score at the endline was significantly different from the average of the previous / midline survey round. * = P<0.10, ** = P<0.05, and *** = P<0.01

PO perception on if their objectives in participating in the project have been achieved. Discussed in the Impact section, a look at qualitative data from in-depth interviews of the partner organizations reveals that POs objectives for joining the project included increasing capacity in research in the water, health and education sectors; affecting policy; building a network of project participants to build collaboration and knowledge sharing; and obtaining material to train students in technical skills. While most were satisfied with the capacity building and getting technical material to train students, others felt that their objective of affecting policy and building a network of research organizations had not been realized.

Other intermediate outcomes examined under the efficiency criteria include improvement in PEM skills and tools; PO ratings of PEM mechanisms for their effectiveness in teaching new skills and building capacity; percent of POs responding that the project was perfectly suited to their needs; and change in technical report scores over the project period. Table 3.2.13 shows PO ratings on the increase in their organizations competence in running policy simulations and knowledge of advocacy and communications. While 50 percent of POs rated the increase in policy simulation skills as being fundamental, there was a more modest response to increase in advocacy and communication skills. In 2010 and 2011, 43 and 57 percent of POs stated that the increase in advocacy and communications skills was “modest,” but this changed to “large” increase in 2013.

In the 2013 survey POs were asked “How would you rate each of the mechanisms for their effectiveness in teaching new skills and building capacity?” (Table 3.5.4). Training workshops and the mentoring program were each rated “good” by 50 and 43 percent of POs. “Peer-to-peer reviews” was rated “excellent” by 43 percent, and shared project space was rated “fair” by 57 percent. Only 29 percent of POs said that the project was tailored to be “perfectly suited to our needs”, while 64 percent said that it was “mostly suited to our needs.”

The indicator change in report scores, discussed in the Impact section shows that there was an 11 percent significant improvement in overall report scores over the 5 years. Reports were scored for organization and analysis and research conclusions and both GDN produced reports as well as other project reports were scored.

Effectiveness: For the VFM analysis effectiveness is viewed as the final outcome or impact of the project rather than the achievement of output and outcome targets in the original logical framework or the effectiveness and appropriateness of the program approach. The goal and purpose of the project is to affect both the demand and supply side of policy changes; that is, generate demand within the policy community for research conducted by the POs; and strengthen the capacity of partner organizations to engage with policy makers by providing evidence based research (supply side).

For the demand side the following indicators are examined:

- Consultation by the policy community with research or advocacy organizations on policy issues (Table 3.2.5)
- Receptiveness of the policy community to research information: reaching out to the PO for a general policy discussion or to learn the results of the organizations work (Table 3.2.4).
- Use of information in the research and analysis report by the policy community (Table 3.2.8).

Table 3.2.5 indicates the average number of times the policy community consulted with local think tanks or policy research NGOs and local advocacy NGOs over the life of the project. The change is marginal and shows limited improvement in the policy community using the think tank / policy research / advocacy NGOs as a source of information. Searching the internet as a source of information saw a large spike in 2010-11 and in 2013 getting information from the relevant government ministry or agency also saw some gain. This result is confirmed by Table 3.2.4, where across all countries and for 10 of the 14 POs the largest percent of the policy community responded that in the previous year they met with the organization "to get information for a general policy discussion or to learn of the results of the organization's work" only "1-2 times a year" or "never." Positive results are seen in Indonesia where 45 percent PC respondents said that they meet with CEDS "once a month or more often" and in the African countries – Ghana, Kenya, Nigeria, and Uganda and in one Latin American country, Argentina where the largest percent of policy community responded that they meet with the POs "every couple of months."

Table 3.2.8 also shows that when information from research and analysis reports are used -- and 78 percent stated that they do use the information -- it is mostly to raise public awareness of an issue, 60 percent; for evidence in policy debates, 58 percent; and to provide access to information, 55 percent. Overall, 42 to 56 percent of policy community respondents said that they use the information "mostly" or "somewhat" to make any changes.

Finally, Table 3.2.6 examines the general effectiveness of the research undertaken by the partner organizations and the policy community's perception on the quality of their research. It shows that even though the policy community rated the following questions: (i) Are the organization's policy recommendations helpful? (ii) Is the organization a valuable source of research, including data and statistics, (iii) Does the organization focus on issues that are of high priority as "very much" (and the percent giving this rating has increase over the duration of the project); they rate the question "does the organization's work positively impact public policy or administration as "somewhat." Thus even though the research is being considered as useful it is not being used to the same extent.

For the supply side the following indicators are examined:

- PO perception on if their work has been used by the government or parliament
- Continued research on transparency and accountability projects by the POs after the project finishes

Information for both of the above indicators were obtained from in-depth interviews with the partner organizations; the first has been discussed in the Impact section and the second under Sustainability. POs stated that work in the education and health sector was more useful for the government than work in the water sector which had several problems with data availability. Program budget analysis, BIA, and CEA were also cited as being tools that were more accessible and useful for the government compared to policy simulations that were done towards the end of the project. Lastly, there were very few examples of concrete changes in policy due to the research of partner organizations, the exceptions being seen in Armenia, Nigeria, Guatemala and Argentina.

All the POs indicated that they will continue working on transparency and accountability issues after the project finishes. However, of them, six indicated that they lacked sufficient resources to continue research in this area and are seeking additional funding.

Evaluation Questions: There are three specific questions related to VFM for the final evaluation.

Is there an optimum balance between Economy, Efficiency and Effectiveness? Overall, the project is mostly well balanced in economy and effort/focus on efficiency. With respect to effectiveness, it was realized before the mid-term review that there needs to be more focus and support to the POs on communication and dissemination of their results else they would not see results in effectiveness. While some results have been achieved in effectiveness, more could have been done to set the stage at the start of the project.

What are the costs and benefits of this program? In addition to the direct projects costs (grants, workshops and technical advisors, dissemination, management, administration and M&E), other costs associated with the program include the time spent by the management and technical staff on this project and time spent by the partner organization staff – both directly in project research and write up as well as travel time to conferences, regional workshops and external presentations. The benefits include increased skill and capacity of the partner organization, improvements in the quality of the research and reports, and their ability to inform, influence and change policy in their countries. While it is difficult to put a monetary value on the benefits of the project on the partner organizations (and compare it to the costs) it is clear that the increase in capacity and skills is not a one-time fixed value, since it will have a multiplier effect on future research and dissemination undertaken by the research organizations. The visibility that the project has given them in their countries is also valuable, since it has increased awareness of their work and in some cases made them an authority on certain topics. Thus in the evaluators opinion the benefits have outweighed the costs.

Overall, did the program represent good value for money? Yes, especially with respect to economy and efficiency, and to a limited extent with effectiveness. One shortcoming of the project was a lack of sustained engagement with the policy community since inception to make them recipients of the policy options. An increased focus on communication and dissemination mid-way into the project certainly helped, but it only affected the supply side of quality research – ensuring that the partner organizations reached out and distributed their research to a wide audience. This was not enough to generate demand for quality research by the policy community. Suggestions on how to structure the project differently are provided in the recommendations section.

3.8 Sustainability

The sustainability of the project is based on whether the POs have developed the capacity to produce research of a high enough quality and the abilities to disseminate it effectively to the policy community so that it is used to improve the quality of government budgeting and expenditures, ultimately improving service delivery and the lives of their countries' citizens. Looking at the results of our analysis of project impact and effectiveness presented in Sections 3.2 and 3.5, it seems that POs have mastered these skills sufficiently in order to continue these types of analyses in the future.

Retention of these capabilities is a key issue and we see that while there has been turnover during the course of the project, the POs have still managed to improve their capabilities based on our analysis. As evident in Table 3.8.1, we see that there are only four organizations in which the original project team remained in place for the duration. This means that a fair amount of gained knowledge through the workshops has been lost to the POs over the course of the project. Our analysis of the project outcomes takes into account this staff turnover that has already taken place. At the same time, all those that are currently working at the PO intend to stay beyond the end of the project, with the exception of the Nigerian PO.

Table 3.8.1 Staff Turnover and Retention After Project End

	Number of staff that worked on the program	% still working at org	% planning to stay past the project end
All countries	**103**	**63**	**64**
Argentina	10	30	30
Armenia	5	40	40
Bangladesh	11	55	55
Ghana (a)	5	0	20
Guatemala	8	88	88
India	15	60	60
Indonesia	9	100	100
Kenya	5	80	80
Mexico	7	86	86
Nepal	5	100	100
Nigeria	3	100	67
Peru (a)	5	-	20
Philippines	7	100	100
Uganda	8	50	50

(a) Did not provide an answer regarding the number of project staff still working at the organization.

The continuation of transparency and accountability work also depends on the interest of the PO researchers in the work and funding levels. All PO organizations indicated that they would like to continue research on transparency and accountability issues although four Asian POs and two Latin American POs mentioned that funding was a limiting factor for this continued research. While the number of POs working on expenditure-accountability focused projects has not changed much over the years, the average number of these projects increased over the course of the program as seen in Table 3.8.2, a promising sign.

Table 3.8.2 Work on Transparency and Accountability

		Projects with public expenditure-accountability focus	
		Number of projects	*Percent of revenue/turnover*
2009	# of POs with the activity	11	8
	Mean of nonzero responses	2.8	41.5
2010	# of POs with the activity	13	13
	Mean of nonzero responses	2.8	61.8
2011	# of POs with the activity	14	14
	Mean of nonzero responses	3.2	34.2
2013	# of POs with the activity	12	10
	Mean of nonzero responses	3.8	52.7

Lastly, we look at the sustainability of the PO network. Although we determined that the creation of a network of POs involved in the project was only partially successful in our analysis in Section 3.5, all POs indicated they thought they would contact other POs in the project after project end, signally that the network would not disappear.

3.9 Replicability

The project has several innovative and practical factors (discussed in the next section). The project can and should be replicated subject to three conditions stated below.

1. All five components outlined in the next section must be included in the project. The discipline of the training and the multiple opportunities for implementing and discussing the new analytic techniques are extremely important for included organizations to master this type of analysis.
2. Constant support from the project organizers to participants for the analytic work is essential. This includes frequent monitoring of progress and feedback on progress.
3. The five-year project time period is very important for both mastering the analytic skills and getting the research used in the policy process. It clearly takes time for policymakers and their advisors to become comfortable with new types of analyses and to accept findings derived with them. Even more, while policy action in the form of a law or regulations being passed may occur in 2-3 years, successful *implementation* of the laws or regulations may well require additional outside advice. Hence, the credibility of the analytic results as evident in operational project success can easily take the full project term.

4. Innovation and Learning

4.1 Innovation

The Evaluator's view is that the project is a genuinely innovative project. Its implicit paradigm can be defined as consisting of the five elements discussed below with some suggestions for future change. While none of these is unique to this project, the project has been successful in employing them effectively together.

Solid Conceptual Framework. In contrast to many technical assistance interventions, the project was underpinned by a solid conceptual framework, supported with rigorous empirical evidence, that makes a convincing case for not only for citizens and consumers of government services making their voices heard by politicians and government officials but also for evidence-based advocacy.[24]

Recruit Organizations with a Clear Interest in the Approach. The Call for Proposals stated a "preference for institutions with experience in analyzing budgets" and also said that "institutions [must] have the skills necessary to fulfill the tasks," which were detailed in the Call. In short, applicants had to demonstrate prior real interest in this accountability field and an ability to do more demanding analyses that are key to the Project.

Caveat: However, some organizations already had skills in PBA and communication (Latin American) and felt that there was limited new learning when these technical areas were covered. Others felt that each topic contributed to their skill set (Africa and Asia). Greater screening of partner abilities and skill set during selection will help ensure that organizations have a similar skill base. Another option is to manage partner expectations regarding new skills that are applicable to them.

Highly Structured Technical Assistance that Is Immediately Applied to Local Issues. The project instructed POs in four key types of PEM analysis: program budgeting, benefit incidence, cost effectiveness analysis and running policy simulations. These were introduced in sequence at global and regional workshops, with each workshop having the objective of equipping partners with the ability to apply the technique when they returned home. Each PO was required to apply each analytic technique or

[24] See C. Griffin et al., *Lives in the Balance: Improving Accountability for Public Spending in Developing Countries* (Washington, DC: Brookings Institution, 2010).

tool to three sectors: health, education, and water.[25] Reports were produced on a firm schedule. Draft reports were reviewed by sector specialist technical advisors and project management team and detailed comments sent to the PO. During the implementation period, support was available from staff and peer organizations. At the next workshop, the studies produced were critically reviewed by peer organizations. The one-two punch of instruction-application reinforced by actively reviewing other applications proved to be a solid method for transferring knowledge and building confidence.

Caveat: The water sector was the most difficult for POs to analyze due to problems with data; in some cases policy simulations were run on unrealistic scenarios, or issues not a priority for the government. Since simulations were the natural end to a progression of policy related skills, more time could have been spent on this topic.

Transforming Advocacy, Engaging Constructively. Within the project the accent was squarely on evidence-based policy making and the firm belief that if advocate-analysts bring solid empirical information to policymakers the chances of constructive engagement are much higher than employing a more "rights-based" approach, in which change is demanded because it is the right of citizens to have better services or more influence on budget decisions. The project provided technical assistance for communicating results effectively and this TA is viewed as important as its analytic counterpart.

Caveat: The project was heavily focused on the supply side of policy research, and it was hoped that effective communication and advocacy of research results would be sufficient to generate demand for quality research. Additional funding was also utilized every year to train MPs and journalists on the budget, and engage citizens in the policy debate. While this helped in some countries, it did not generate sufficient demand from the policy community.

Peer Learning. The principle of peer learning was important to the project for two reasons. First, it was a way to provide a greater volume of technical expertise—stronger POs helped those with less experience. Second, the project's peer review system reinforced the training in analytic and communications techniques that are fundamental to project success. Each review session benefited both the reviewers and the analysts whose work was under review.

[25] A question in the MTR survey of five POs asked if they thought this process provided them with sufficiently space to address important questions while also learning the techniques. All thought it did.

4.2 Support of project M&E process to the evaluation

In this section we discuss the extent to which the project logframe and M&E processes supported the evaluation. The project logframe and the evaluation were well integrated. The key to this success is that the evaluator was integrated at the very beginning of the project and was involved in the design of the logframe and its corresponding indicators. As such, the evaluator was able to provide the necessary input to develop measurable and useful indicators for evaluation purposes.

Where the evaluation could have been improved was in the availability of resources to collect higher quality survey data. Due to budget limitations, the evaluator chose to field the surveys online or through emailed or mailed paper versions rather than in-person. Self-administered surveys suffer from low response rates (it is easier to ignore an email or a letter than someone calling you or knocking on your door) and lower quality data (respondents can skip questions or may answer them in-correctly when using an emailed or mail survey). On the other hand, in-person surveys require significant additional resources to train interviewers and pay for their time to administer the surveys. Additional resources would have also allowed for a more sophisticated evaluation design that could have integrated a comparison group, allowing the evaluation to attribute the results to the project with greater certainty.

5. Summary of Recommendations

5.1 Program design recommendations

- *Have a less complex project.* The project in its current design is very ambitious in its scale and scope -- 14 POs across the world, 3 sectors, and 4 technical policy analysis tools. While the tools are important and have a natural progression in developing skills, culminating in examining policy alternatives via simulations it may be easier and more effective to focus on just 1 or 2 sectors. In addition, the project hoped to use several mechanisms such as the Steering Committee and Regional Hubs, and establish / institutionalize new structures such as a strong network of institutions and knowledge management material. Of the four key outcomes listed for the project, the first building capacity has definitely been achieved and change in policy to some extent. However, benchmarking was difficult due to different country contexts and focus of research; and there was insufficient time and focus on establishing a strong network of institutions.
- *More flexible program design.* The project was prescriptive in the sectors and methodology that each PO needed to work on and build skills in. While this made training and peer learning easier to manage, it constrained the POs from doing research in areas that were valuable for their own research agenda or country priorities.
- *Engage the policy community since the beginning of the project*. Finding a champion in the policy community who is vested in the research or a key client for the research will increase its use and culminate in policy changes. In many countries frequent transfers of senior government officials creates challenges, but working with department officials from the beginning – both technical and management may help to maintain interest in and use of the research.
- *Engage technical advisors from the beginning and maintain continuity*. It is important to have continuity in technical advisors and have them be on the same page with respect to the rigor of research, sensitivity to the cultural context, and availability of data. Given the scale and scope of the project, having 3 main TAs was insufficient, and getting shorter-term TAs up to speed on program objectives and status of the various partners was difficult.
- *Include a communications and dissemination focus from project start up.* This was added to the project only mid-way. Outreach and sharing research results from the use of each tool – PBA, BIA, CEA and PS with the policy community could have created interest in the research and increased its use in making policy changes.

- *Have partners that are uniform in their skills base*. Partners included organizations at the academic end of the spectrum as well as NGOs that were more inclined towards advocacy. A mixed group makes the design of workshops and trainings more difficult and resulted in one organization being dropped from the program and a few that only put in minimum effort into the project.
- *Explore pairing or partnering organizations from inception*. Peer learning was considered by most to be very valuable, perhaps partnering two countries in a region to work together could have created synergies and a natural competition.

5.2 Program management recommendations

- *Ensure there are sufficient technical and management staff for project implementation*. Both the technical (R4D) and management (GDN) side required additional staff than originally envisioned to implement the project. Both sides underestimated the time and resources needed to lead a project with this scale and scope, and rectified this omission after the first year or so. Assigning staff to work with different organizations to maintain continuity and understand their challenges is also important.
- *Combine technical tasks and management tasks as being the responsibility of one organization*. Having two separate organizations in charge of decision making -- one for management and the second for technical content creates time-lags in resolving issues that include a program design change.
- *Include additional resources for on-site visits to POs as part of the monitoring and evaluation*. Telephones and Skype are effective only up to a point. The quality of calls was not always good and caused frustrations. On-site visits were done for a few partners and considered very helpful.
- *Invest more resources in developing a more user-friendly webspace dedicated to encouraging sharing of information among POs.* There was a lack of knowledge among POs regarding others' research topics and analysis methodologies prior to the workshops. Development of a platform that allows the PO to very quickly understand the topics and analysis methodologies being used by the other POs during the development of reports would facilitate greater peer learning.