MULTILATERAL BANKS AND THE DEVELOPMENT PROCESS

VITAL LINKS IN THE RESULTS CHAIN

VINOD THOMAS AND XUBEI LUO

MULTILATERAL BANKS AND THE DEVELOPMENT PROCESS

MULTILATERAL BANKS AND THE DEVELOPMENT PROCESS

VITAL LINKS IN THE RESULTS CHAIN

VINOD THOMAS AND XUBEI LUO

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2012 by Transaction Publishers

Published 2017 by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN
711 Third Avenue, New York, NY 10017, USA

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Catalog Number: 2012005689

Library of Congress Cataloging-in-Publication Data

Thomas, Vinod, 1949-
Multilateral banks and the development process: vital links in the results chain / Vinod Thomas and Xubei Luo.
p. cm.
Includes bibliographical references and index.
1. Economic development projects—Developing countries—Evaluation. 2. Economic assistance—Developing countries—Evaluation. 3. Economic development—Developing countries. 4. Development banks. I. Luo, Xubei. II. Title.
HC59.72.E44T56 2012
332.1'53091724—dc23

2012005689

ISBN 13: 978-1-4128-4767-4 (hbk)

Contents

List of Tables and Figures

Figures

Tables

Boxes

Foreword

A chain is only as strong as its weakest link. This metaphor is highly applicable to the work of countries, the multilateral development banks, and the development process itself. If each link represents a development project, when one project misses its goals, the development program's effectiveness is compromised; if the links are strong, development effectiveness is maximized for the sector, country, and region.

Crucial to ensuring the strength of the process is research, particularly evaluative research. It informs the quality and impact of development operations on the ground and helps policymakers make development decisions. Indeed, economic gains are greatest when developing countries exploit cross-sectoral links, when innovative approaches improve incentive structures, and when thoughtful policies can scale-up projects and programs with favorable outcomes to promote economic growth and create jobs.

As this book emphasizes, two of the most important overlooked links in the development chain are relevant data and farsighted planning from policymakers and their partners.

The commitment of the multilateral development banks, including the World Bank, and other institutions to evaluating project impacts has produced a wealth of new data on specific impacts on certain populations. This work has vastly expanded our knowledge of what works in development and what does not. Despite these advances, poor or missing data continue to hinder further research and analysis in some sectors—including infrastructure, gender, and climate change—and for some target populations.

And parts of government have failed to address important, longstanding, but less politically glamorous issues because of the nearsightedness of some policymakers. Investing in growth-enhancing infrastructure such as roads, ports, and electricity grids, for example, could expand trade in goods and services and increase movements

of people. Investments of this nature are critical to sustaining global growth and reducing poverty.

A lack of planning is not limited to the developing world. With global markets in upheaval over the uncertain outlook for the US and European economies, both advanced and developing economies must attend to long-term goals. Policymakers should avoid politically expedient decisions when more comprehensive policies can be put in place to spur job creation and economic growth. One approach to promoting growth is to develop national industrial policy along the lines of a country's comparative advantage. East Asian countries present good examples of the huge economic gains of exploiting comparative advantage and adapting to the dynamic global economy.

Indeed, the global economic landscape is changing at a dramatic pace. In a paradigm shift to a new multipolar world, the most powerful engines of growth are coming from the developing world, particularly from China, India, and Brazil. Characterizing this multipolar world are stronger South–South links for trade, finance, technology, and knowledge. Trade is increasing between East Asia and emerging Latin American countries like Chile and Peru. And sub-Saharan Africa is experiencing unprecedented growth.

The developing and newly industrialized nations present many lessons. China has lifted 600 million people out of poverty since 1980, through agricultural reforms and labor-intensive manufacturing development. India has followed its comparative advantage in services to grow remarkably in the last two decades. And the Republic of Korea has increased its GDP per capita ninefold since the 1970s. But not all attempts to facilitate growth and development meet with success. It may seem obvious that development practitioners should apply policies that work and avoid those that do not, ensuring that the links in the chain are strong. But too often they overlook those links—because of complexity, because of tradition, because of strong vested interests.

Informing these links are lessons from World Bank research, from Bank-supported projects, and from rigorous evaluations by the World Bank's Independent Evaluation Group and Development Impact Evaluation initiative. The authors, Vinod Thomas and Xubei Luo, present many of the lessons in this book. These contributions can help policymakers and practitioners focus on the right results, measure the

progress, apply the lessons, and thereby strengthen the development chain for greater effectiveness.

Justin Yifu Lin
Senior Vice President and Chief Economist
The World Bank

Preface

With the spread of the global financial crisis, the world economy shrank in 2009 for the first time since World War II. Forty million more people went hungry that year, while 64 million more found themselves living in extreme poverty by the end of 2010.

If the benefits of sound financial regulation had been recognized and the needed regulatory measures had been put in place, billions of dollars might not have evaporated and millions of people might not have suffered huge setbacks in life. If the rosy assumptions of prestigious institutions and leading academics had been challenged, the falling dominoes might not have wrought such damage.[1]

Meanwhile, the value of global GDP exposed to tropical storms soared to more than $1.5 trillion a year, triple the $525 billion of forty years ago (in 2000 dollars).[2] Acts of prevention and mitigation of natural disasters could have reduced the damage: $1 spent on loss prevention, by some estimates, could avoid an average of $4 in future losses.[3] But in the wake of natural catastrophes, attention seems to focus only on relief and rehabilitation.

Where prevention is emphasized, the investment pays off. Bangladesh, subject to annual flooding and massive losses of life, invested in early-warning systems and hurricane shelters and evacuating areas most at risk. While the cyclone and floods of November 1970 took the lives of 300,000 people, a similar storm in May 1997 claimed 188.

These lessons from the financial crisis and climate-related hazards are two of several from recent evaluations that shed light on development effectiveness in the context of the support from multilateral agencies such as the World Bank. They illuminate the links in the chain connecting development efforts to development results. The results chain comprises inputs such as financing for schools, outputs such as student enrollment and graduation rates, outcomes such as student learning, and impacts such as competitiveness and social progress. Evaluations show that the connectivity among actions along the chain,

by countries and their development partners, differentiates the big from the small impacts.

We are familiar with how results chains work. In the food chain, for example, plants, animals, and human beings are linked. If crucial links are threatened, such as by climate change, the entire chain could break.

Many of the evaluative lessons are not new, and some are even routine; but they could have large impacts. Yet, they receive scant attention. The reasons for such neglect vary—from the inconvenience of tracking some of the more complex links to a short-term focus driven by pressures from vested interests.

Evaluation can provide timely information on what has worked, in what contexts, and why. It can help decision makers understand what has happened and identify ways to achieve better outcomes. Often it confirms existing assumptions, but it can also encourage fresh thinking that improves development effectiveness. Publicizing findings, having debates, strengthening methods, and focusing on checks and balances—all can give overlooked evidence greater traction and impact.

This book looks at several examples that show how learning from the lessons of experience improves development effectiveness. It marshals important lessons from recent evaluations, especially those that are underemphasized or unexpected, with the main purpose of informing policymakers and practitioners about issues vital to development.

Its basic message: To meet emerging challenges, development professionals need to focus on results, measure progress, and apply lessons. This may be common sense, but the chapters show that concerted effort is often needed to overcome myopia and inertia.

Vinod Thomas and Xubei Luo

Notes

1. Independent Evaluation Office (2011).
2. UN (2011).
3. Multi-hazard Mitigation Council for the Federal Emergency Management Agency (2010).

Acknowledgments

We would like to thank Patrick G. Grasso for his special contribution to the writing of the book during all stages, and for writing the chapter on the timing of evaluations. We also would like to thank Silke Heuser for valuable inputs on natural disasters and infrastructure.

We would like to pointedly acknowledge the evaluation work of team leaders, staff, and the management team at the World Bank Group's Independent Evaluation Group (IEG), on which much of this volume is based. We are grateful to their generous support, advice, and inputs during the various phases of preparing the manuscript.

Special thanks go to Daniela Gressani and Jiro Tominaga for taking the time to go through the different drafts and for fruitful discussions, and to Kalpana Kochhar and Hans-Martin Boehmer for peer reviewing the manuscript. Valuable comments from the following individuals are gratefully acknowledged: Martha Ainsworth, Amitava Banerjee, Geeta Batra, Jaime Biderman, Soniya Carvalho, Kenneth Chomitz, Victoria Elliot, Ariel Fiszbein, Nils Fostvedt, Monika Huppi, Jaime Jaramillo-Vallejo, Ali Khadr, Linda Morra Imas, Ray Rist, Mark Sundberg, Helena Tang, Marvin Taylor-Dormond, Stoyan Tenev, and Christine Wallich.

Our sincere appreciation goes to Sanjay Pradhan for chairing a workshop at the World Bank, and to Christopher W. MacCormac for moderating a learning event at the Asian Development Bank to discuss a shorter version of this work; to Alan Gelb, Arianna Legovini, Samuel Otoo, Alejandro Soriano, and Xianbin Yao for their detailed comments and suggestions; and to Sidney Edelmann and Njoman George Bestari for organizing the events. We are most grateful to Shahrokh Fardoust and Shahid Yusuf for inspiring brainstorming at the early stage when the book was taking shape and for continued comments, and to Nancy Birdsall, Partha Dasgupta, Patrick Guillaumont, Oh-Seok Hyun, Naoko Ishii, Ravi Kanbur, Justin Yifu Lin, and Robert Picciotto for valuable

comments and suggestions. We wish to thank Bruce Ross-Larson and his team at Communications Development Incorporated for editing this book for publication, and William B. Hurlbut, Heather Dittbrenner, Helen Chin, and Eric Van Zant for editorial support at various stages.

Abbreviations and Acronyms

CPIA	Country Policy and Institutional Assessment
GDP	Gross domestic product
HDI	Human Development Index
IDA	International Development Association
IEG	Independent Evaluation Group
IFC	International Finance Corporation
MIGA	Multilateral Investment Guarantee Agency
REDD	Reduced Emissions from Deforestation and Degradation

Eight Lessons

Lesson One: Let not the urgent divert the important

The needed attention to the immediate ought not to crowd out what is important for the long term.

Lesson Two: Connect links that strengthen results

Gains are impressive when vital links among related areas are capitalized on.

Lesson Three: Composite indicators can mislead

Summary measures are catchy, but if not well grounded, they can misdirect policy.

Lesson Four: Go from averages to targeted segments

Measures of the average are not enough as a basis to guide interventions that are intended to target a subset of the population.

Lesson Five: Align intermediate and final goals

An exclusive focus on an intermediate part of the solution can come at the expense of achieving the ultimate objectives.

Lesson Six: New challenges call for shifts in direction

Past success of projects in an area notwithstanding, direction may need to shift in view of new challenges.

Lesson Seven: Capture opportunities

It may not be enough to do things right when changes can enable us to do the right things.

Lesson Eight: Tarry not—for timing is (almost) everything

Follow-up is more likely when evaluative lessons are available at the right time.

1

The Results Chain

One of the great mistakes is to judge policies and programs by their intentions rather than their results

—Milton Friedman

Taking lessons to heart can be a matter of success and failure. While keeping kids in school has its benefits, the objective of education is learning. Several countries illustrate the limits of an exclusive reliance on raising inputs to achieve education results.[1] Tanzania, for instance, rightly won a United Nations award for the rapid expansion of school enrollments in recent years, and yet learning outcomes in secondary education declined. When students are pushed into the next grade and qualified teachers are in short supply, results suffer. The forgone impacts of ignoring learning outcomes are estimated to be sizable losses in GDP, and hence employment and poverty reduction as well.

Today, we know the enormous costs that climate change can impose on society. Cutting and burning forests produces a fifth of the greenhouse gases in the atmosphere. But when the choice is posed as complete protection or none, the pressure from vested interests to cut down trees usually wins out—at great cost to society. It turns out that a balance—encouraging forest protection in tandem with economic activities in some areas—can deliver better results than either extreme.

This book looks at vital but often overlooked links in the results chain—links that the evaluative experiences of countries and their development partners such as multilateral banks can illuminate. When critical links are broken, big societal or opportunity costs can result. The objective is to signal to policymakers and development practitioners the critical elements connecting actions to results that could change decision making, especially ones that tend to be overlooked. Such a focus is not intended to suggest that our knowledge has to be complete or that action has to be comprehensive to make progress.

Rather, it is meant to put the spotlight on knowledge of what works and in what context, and of what elements can block or facilitate the translation of that knowledge into actions and results.

The stakes in deriving such lessons and strengthening the links are high. Despite progress in reducing poverty, some 1.4 billion people were estimated to be living in extreme poverty (on less than $1.25 a day) in 2005, with the absolute numbers affected by increases in population—and subsequently by the impact of multiple global crises. Ineffective resource use for development misses opportunities both for potential beneficiaries and for those supporting development programs.

In an era when resources for investment are scarce, evaluating the effects of investments can provide considerable social value. Countries and agencies find that independent evaluation can help improve development effectiveness. Evaluating public programs, projects, and policies is now accepted as a way to hold authorities accountable for their use of resources while learning which approaches work and which do not, and under what circumstances.

Opinion drives policy. Opinion driven by vested interests can induce bad policies. But evaluation can enrich opinion with knowledge and enhance the accountability of decision makers for results.

Evaluation can promote accountability relating to actions taken by countries and international financial institutions, and contribute to learning about development effectiveness. It can influence the change of process in policy and institutional development.[2] It can especially add value when it identifies overlooked links in the results chain, challenges conventional wisdom, and shines new light to shift behavior or even ways of doing business. Some of the most important findings are those that shed light on important but often ignored lessons, providing a basis to reconsider and guide the actions of organizations charged with implementing public policy.

The chapters here provide examples of how both evaluators and users of evaluation can look afresh at some of the vital, yet underappreciated links in development activities. The illustrations are in good measure drawn from the work of the Independent Evaluation Group (IEG) of the World Bank Group (comprising the World Bank, the public sector wing, and the International Finance Corporation, IFC, and the Multilateral Investment Guarantee Agency, MIGA, the private sector wings) but also from the work of other development evaluation organizations and researchers.[3]

How Evaluation Can Inform Development

Evaluation has a long history in development. Bilateral and United Nations aid agencies have funded evaluations of their work over the years, and most now have evaluation offices.[4] And since the World Bank established an Operations Evaluation Department in 1973, other international financial institutions have set up independent evaluation units.[5] Countries also find that evaluation, especially when independent, can make development programs more effective. Colombia, Mexico,[6] South Africa, the United Kingdom,[7] and the United States are among the many countries that have recently adopted or strengthened evaluation policies and capacities.[8]

When competition for scarce resources increases, it seems reasonable to know which programs are achieving intended results with a high cost–benefit ratio and which are not. In turbulent times after crises, there is more demand for information on how government-funded programs are performing.[9]

Evaluation can add real value in assessing whether interventions produce desired outcomes, in what contexts, and through what channels. Success and failure lie at the intersection of individual and micro impacts with social and aggregate impacts, but the interactions are difficult to discern.[10] Good evaluation is associated with good project outcomes.[11] But it has costs. It requires funding, staff time, and other resources to develop the capacity of aid agency and developing country staff to conduct evaluation studies, oversee them, and use their results. Even so, the fragmentation of evaluation products and the diversity of evaluation methods can contribute to "evaluation bombardment."[12] That lack of coherence raises the cost of development assistance, including the cost of evaluating results. So, it is worth asking: When can evaluation be uniquely helpful? And how can the resources spent on carrying out evaluations and applying the lessons have the most impact?

How an organization (or initiative) is expected to achieve results depends on the underlying assumptions—on the validity of the theory of change.[13] The theory of change could lay out a chain linking inputs to outputs, outcomes, and impacts. To answer the question of why an intervention worked or did not work, mapping out the results chain to test the underlying assumptions is key.[14] But many of the events or conditions that are assumed to produce the desired outcomes might not be in place. Nor might the interventions function as expected,

particularly given the complexity and interrelatedness of development programs. These assumptions need to be identified and tested based on the prevailing macroeconomic and political environments and on local policymakers and international donors. Evaluation can unbundle the theory of change to suggest how an intervention converts inputs and outputs into outcomes and impacts.

Through a series of events—such as the 2002 International Conference on Financing for Development in Monterrey, Mexico, which established the Millennium Development Goals, and the 2008 Forum on Aid Effectiveness in Accra, Ghana—the development community has tried to move from a focus on inputs and outputs to a consideration of outcomes and impacts.[15] But there is still some distance to go in focusing on results.

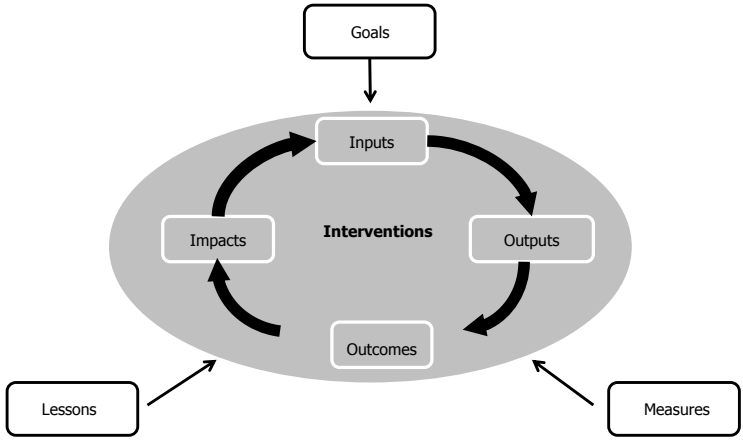
Independent evaluation has a valuable, if not always explicit, role throughout the development process. Its obvious and traditional focus is the link between results monitoring and organizational learning: an independent assessment of results and lessons is generally a key part of the project or program cycle. But independent evaluation does more than look retrospectively for achievements and lessons. And timing can make a crucial difference in influencing decision making and results.

Evaluation considers the goals of interventions, measures their effects, and draws lessons (Figure 1.1); these make up the three main parts of this book. The findings of evaluation refer to and intersect with the full cycle, from inputs, to outputs, to outcomes, and to impacts relating to the interventions. Including the cycle's design implies that evaluation can bring value not only retrospectively, but also in real time and prospectively. We will see that many factors influence results, including conditions outside the domain of the interventions (see Figure 3.1 in Chapter 3).

For evaluators, the shift from tracking inputs and outputs to focusing on outcomes and impacts was a natural outgrowth of evaluations that found discord between plans and results. The focus on results also draws attention to the vital links in the results chain and to the complexity of attributing outcomes to particular inputs.

To be effective, evaluation needs to consider the links connecting inputs to outputs—and to outcomes and impacts. This requires focusing on what might be the right results, getting the appropriate measures, and providing lessons to enhance development effectiveness.

Figure 1.1
Development Process and the Results Chain



What Are the Results to Focus On?

It seems straightforward to focus on results. Less obvious, however, are the level that an intervention should target and how results interlink over time and across microeconomic and macroeconomic dimensions.

It is often assumed that addressing immediate needs will produce desired outcomes over time and that doing projects right will take care of problems in a sector or country. In each of these instances, the intermediate steps may be necessary for the longer term or more full results, but they may be far from sufficient. Focusing on results and being aware of the links across projects, sectors, and country programs are essential.

We might think of a three-dimensional space, with one axis focusing on time, the second on the unit of intervention (project, program, country), and the third on the area or sector within which desired outcomes can be identified. All three aspects come into play in varying measure in generating the results. The findings assembled here point to two sets of lessons.

The needed attention to the immediate ought not to crowd out what's important for the long term. Sometimes a focus on immediate requirements, not aligned with long-term goals, may do more harm than good because development is path-dependent. For example, responses to

natural disasters usually target reconstruction, not mitigation or prevention. But rebuilding substandard homes in haste can lead to more deaths and losses if disasters hit again. Disasters should be treated as risks to development, not interruptions. Among countries that received World Bank support to deal with disasters, fewer than half even discussed disaster responses in their development plans.[16] The opportunity cost is high if the resources are allocated exclusively to immediate needs after a disaster, because in the long run one dollar in prevention can yield as much as four in benefits.[17]

Gains are impressive when vital links among related areas are capitalized on. Country and sector results may differ from project results—for two reasons. The country and sector objectives are not the aggregate of project objectives. And the factors affecting country and sector outcomes can be far beyond those affecting project outcomes. For example, the Chad–Cameroon oil pipeline project was well implemented and a financial success, but the main country's objectives—capacity building in the sector, improved governance, and reduced poverty—were not met. And sector outcomes are often linked to results in other sectors in ways that are not always obvious. Capitalizing on cross-sectoral links can lead to better outcomes and impacts. For example, health outcomes can be better when focusing not only on health projects, but also on infrastructure facilities to distribute medical supplies to health clinics and community outreach activities to inform people about having their children vaccinated.

These lessons are explored in Part I.

How Might Results Be Measured?

What gets measured gets done, so it is crucial that the right things be measured in the right way to promote development effectiveness. Focusing on results is of limited value if they are not measured properly. Poor monitoring and evaluation can risk achieving the desired outcomes and harm development effectiveness by misallocating scarce resources from higher-value activities. Developing sound measures is not entirely a technical matter. To be useful, measures need to be meaningful to those who need the information to address issues of interest. This implies a collaborative approach to designing indicators, and engaging executing agencies, sector experts, and measurement professionals. Moreover, the quality of the indicators depends on their intended use.[18]

Technical issues are important in this context. Consider four technical criteria for indicators: defining results targets, identifying the sources and frequencies of data collection, programming funds for data collection and analysis, and ensuring capacity for achieving the agreed results.[19] Each of these aspects makes a difference to effective measurement.

Among the lessons explored in this book, several involve measurement issues.

Summary measures are catchy, but if not well grounded, they can misdirect policy. Composite measures can be appealing because of their apparent simplicity, and they usefully focus attention on a specific issue. But they sometimes send wrong signals and lead to misdirected development activities. The risks are particularly high when the assumptions and logic behind the construction of the indicators are not clearly documented and when what the indicator actually measures differs from what it claims to measure. For example, the Doing Business Indicators posit that lighter regulation and lower taxes that lower the cost to private business signal a better regulatory regime, spurring better socioeconomic results. A mechanical interpretation—that fewer regulations would be better, without looking at which aspects of regulation and up to what point—can lead policy in the wrong direction.

Measures of the average are not enough to guide interventions intended to target a subset of the population. A project or program may succeed on average but fail to address the right constraint or reach the targeted beneficiary groups. If only the averages are measured, the target population can be left out or wind up worse off when the distribution of benefits is skewed toward the better off. Targeting the areas where the poor concentrate, the issues that the poor commonly face, and the constraints especially relevant for the poor do not ensure that the poor gain disproportionately. Even changing the structure of aid delivery to empower the poor may empower others even more. In Benin, a community contribution requirement created hardship for the poor. Because it is very difficult for the poor to make cash contributions, they usually contribute time and labor, taking them away from income-earning activities.

An exclusive focus on an intermediate part of the solution can come at the expense of achieving the ultimate objectives. Intermediate outcomes are easier to monitor, and we often need to target them to make things manageable. But they may not produce the desired final

outcomes if vital links in the results chain are missing. If the desired results are not spelled out, program achievement will remain unknown: success cannot be distinguished from failure and thus cannot be rewarded. For example, Tanzania rightly won a United Nations award for its progress in attaining universal education ahead of the 2015 Millennium Development Goals, but the gains in enrollment were accompanied by a decline in learning outcomes in secondary education. With resources devoted exclusively to expanding schools, other factors, such as competent teachers, still lagged.

These lessons are discussed in Part II.

How Might Evaluative Information Be Used to Boost Results?

Focusing the results at the right level and measuring them the right way can still be of limited value if the evaluation findings are not used properly. For evaluation findings to have the desired results, it pays to present them at the right time, in the right format, and to the right audience—and to apply them in the right context.

It is often assumed that if evaluative information is useful, it will automatically be used the way it should. This may not be true for many reasons. First, the future may not resemble the past. In a rapidly changing context, simply replicating what worked in the past may not help in the future. More investment is beneficial only if it considers the changes. Second, it is important to look not only at what happened, but also at what could have happened. Getting beyond existing assumptions and received wisdom and looking with a fresh eye could shift behavior or even lead to new ways of doing business. Third, the timing and process often determine to what extent, if any, evaluative lessons can be useful.

This book examines several cases of using evaluation effectively—and the limitations on its use.

Past success of projects in an area notwithstanding, direction may need to shift in view of new challenges. The uncertainties and complexities of development mean that changing environments and emerging challenges can make it inappropriate to mechanically apply the findings from past work to future efforts. For example, large irrigation projects have improved agricultural productivity. But with water scarcity a growing concern, they may not be sustainable. And road investment has increased accessibility and stimulated growth. But rapid urbanization and growing congestion can erode those benefits. Innovative, sustainable strategies are required to meet future needs.

For water, the strategy will include coastal zone management, pollution reduction, and groundwater conservation. For transport, it will include cross-cutting approaches to address the links with energy, land use, urbanization, the environment, and climate change.

It may not be enough to do things right when changes can enable us to do the right things. The role of evaluation in identifying opportunities is often underappreciated. For example, if we could overcome the apparent conflicts among policies, strategies could be built around reducing energy subsidies and targeting them to the poor, thus improving energy efficiency, saving money, and contributing to growth. For another example, if we could overcome internal institutional constraints and limitations in governing charters and policies, the ability of organizations to provide services to clients could be enhanced. Or if we could overcome information constraints, we could make more informed decisions.

Follow-up is more likely when evaluative lessons are available at the right time. Evaluative information can be effective only if it is delivered when it can affect key decisions. For high impact, it is crucial to learn faster what works and what does not, to focus on results at the right time, and to link evaluation findings to development actions. Real-time and prospective evaluations could overcome some limits of ex post evaluation and provide timely feedback. In Mexico, the early results of the evaluation of *Progresa-Oportunidades* helped persuade a new administration not only to maintain it, but also to scale it up. In the Philippines, the evaluation of an early childhood development program affirmed strong results and supported the decision to expand it.

These lessons are covered in Part III.

How Might Evaluation Lessons Be Used to Improve Development Effectiveness?

For actions to translate into the desired results, the intended and unintended consequences of development activities must both be considered. It is not enough to measure only the intended results because the unintended ones may provide unexpected benefits or costs. Unintended results can provide a rich source of learning for future activities and checks on current ones.

Many evaluation findings confirm what is known. Their value lies mainly in summarizing lessons and, perhaps, in suggesting improvements for future interventions. But some evaluations generate

unexpected results that question the expected connections between actions and desired outcomes, including the critical assumptions and context for the underlying theory of change implicit in the activity.

Some of these important elements are neglected or simply taken for granted. By pointing out crucial but neglected areas and providing timely information to change development thinking and guide policy decisions, evaluations can push policy interventions from a generally accepted but perhaps ineffective (or even harmful) state of inertia to a more beneficial course.

The rest of this book elaborates on eight major lessons for development effectiveness. The lessons address three questions: What results are to be measured? How might they be measured? And how can the evaluative information be both useful and used?

In addressing these questions, the chapters that follow give examples of how these issues seem to have played out in various development circumstances. The aim is to bring out some of the unexpected or less than obvious lessons from past evaluations and to suggest how future work might identify new lessons.

We might imagine some of the reasons why these links often get short shrift. They range from the lack of data and information, to the inconvenience of dealing with their implications, to the vested interests that block their pursuit. For example, since data on averages are easier to compile than those for particular segments, it is the averages that often get examined even if the goal is to reach target groups. It might also be inconvenient to deal with critical links such as regulation, when it takes time and attention away from the convenient, such as liberalization.

Incentives in an organization or the development process are often stacked in favor of the immediate: for example, emphasizing relief efforts in a natural disaster, vital as they are, rather than prevention, which is equally critical but takes time to render results. Or there may be emphasis on what is visible, such as the opening of a school building and getting kids to school, rather than the learning outcomes, which not only come later but might also be more intangible in the first instance. Furthermore, vested interests can block reforms that could take away the privileges of special groups even as societal gains might outweigh their losses. Ignoring the link between climate change and natural disasters or agricultural productivity is a case in point.

Notes

1. IEG (2010g).
2. Rist et al. (2011).
3. For another effort to look at how evaluation can unearth the unexpected, see Banerjee and Duflo (2011).
4. OECD (2010); UNEG (2009).
5. ECG (2010). Independent evaluation offices later were established at the IFC (1995) and MIGA (2002). All three units have been merged into the IEG; African Development Bank Group (2011); Asian Development Bank (2011) Inter-American Development Bank (2010).
6. Castro et al. (2009).
7. UK House of Commons, International Development Committee (2010).
8. USAID (2010).
9. Morra-Imas (2011).
10. Dasgupta (2007).
11. Thomas and Tominaga (2010).
12. Picciotto (2007).
13. Morra-Imas and Rist (2009).
14. White (2009).
15. OECD (2009).
16. IEG (2006b).
17. Multi-hazard Mitigation Council for the Federal Emergency Management Agency (2010).
18. Mackay (2006).
19. Castro (2008).

Part I

Aim to Get Results

Focusing on results seems straightforward. But it is often not practiced because of the challenges in carrying out the required analysis and because of the political or bureaucratic pressures to ignore them. Indeed, interventions can translate to results at different levels, which are interdependent and do not necessarily flow in a linear fashion. Focusing on results over time, across sectors, and between the project portfolio and country program can produce longer-lasting results. For natural disasters the need is to invest in preparedness and mitigation, not just in relief and reconstruction. In this instance and others, we see the disadvantages of remaining permanently reactive.

Going from project to country outcomes draws attention to the value of analytical work and policy changes that can affect individual investments. Factors beyond the control of the government and donors often influence country outcomes. Assuming a simple translation or aggregation of results from project to sector or country can risk not pursuing, and thus not achieving, the desired results. Examples, even if limited, suggest the value of cross-sectoral approaches. Interventions that connect the dots—rebuilding schools, providing better health, and improving infrastructure—are more effective when the pace and sequence of actions in different sectors are considered together, because the actions interact and their results are often interdependent.

2

Let Not the Urgent Divert the Important

We can't control when or where a terrible storm may strike, but we can control how we respond to it.

—Barack Obama

It has become increasingly necessary to respond quickly to urgent needs in development work. But one of the drawbacks in such responses is the temptation to focus exclusively on the immediate needs, even when doing so undermines long-term results. Attending only to immediate objectives may not ensure attention to continuing concerns and at worst might impede improving sustained outcomes.

This often happens when the immediate needs are clearly visible and compelling. To be sure, a quick response to urgent needs is essential in the wake of a calamity, such as a natural disaster. The top priority must be to save lives, and basic human needs must be addressed quickly. Reconstruction and rehabilitation are vital to resume life and restore livelihoods.

But when things are rushed, quality is often traded for speed, as in many postdisaster situations. Immediate steps need to feed into longer-term solutions—just as the responses and the protocols in the emergency room are essential to a patient's long-term care. Rebuilding substandard homes in disaster-prone areas may provide housing to people in need, but in future disasters it could lead to more severe losses and casualties.

The uncertainties of future and intergenerational well-being could also lead to paradoxical behavior. People may be anxious about climate change but not sure how to think about it or take action.[1] They understand the importance of controlling carbon emissions, enhancing sequestration possibilities, and investing in alternative energy technologies to mitigate the negative effects of climate change. But

the investment needed now seems large while the benefits will accrue only in the years to come.

Of course, these considerations apply to many other situations, including health, education, and finance. The short time horizon of politics can be another factor that leads to a myopic focus on the urgent over the potent.

Rather than seeing the short-term and long-term objectives as tradeoffs, the development community needs to understand how the two can complement each other. Evaluative findings suggest that actions required now that address root causes but also have an eye on the future can solve the dilemma of whether to deal with symptoms quickly or build toward long-term results.

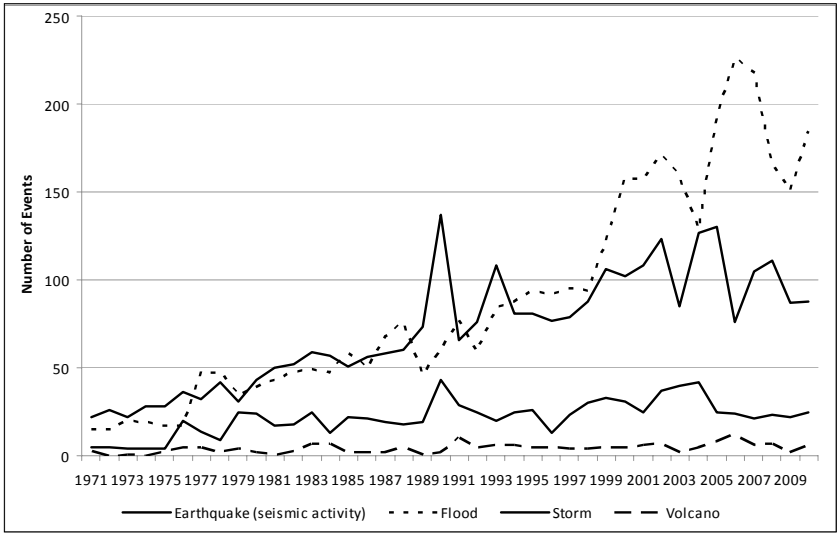
The Incidence and Severity of Disasters Are on the Rise

Disasters can wipe out development gains and negate years of development investment. The 2005 Kashmir earthquake in Pakistan caused an estimated $5 billion in damage, roughly the total official development assistance for the preceding three years and what the World Bank had lent the country over the preceding ten.[2] In 2010 alone 373 natural disasters killed more than 296,800 people, affected 208 million others, and cost nearly $110 billion. And with the full impact of the last 2011 quake and tsunami in Japan still unfolding, estimated damages range from $122 billion to $235 billion or 2.5–4.0 percent of GDP.[3]

Some countries are at more risk given their locations.[4] About 85 percent of deaths from tropical cyclones occur in just two Asian countries, Bangladesh and India, with Bangladesh alone accounting for more than three-quarters of deaths from the cyclones. And 85 percent of the global earthquake risk is concentrated on just 12 percent of the Earth.[5] Small island states with low-lying coastlines are more likely to be hit frequently and to suffer greater losses from storms. Because their people tend to be concentrated along the coasts, they are exposed to heavy winds, storm surges, flooding, and coastal erosion—raising the risk of casualties.

The number of natural disasters rose from 63 in 1971 to more than 400 in 2010 (Figure 2.1), when the world suffered natural disasters of extraordinary magnitude and impact. Devastating earthquakes struck Haiti and Chile, and floods hit Pakistan, West Africa, Sri Lanka, Brazil, and Australia. Disaster-related damages increase as population pressures mount. Some 2.6 billion people have been affected by natural

Figure 2.1
Natural Disasters Are on the Rise



Source: Center for Research on Epidemiology of Disasters (2011)

catastrophes over the past ten years, up 1.6 billion from the previous decade.[6] The cost of such damages is now severalfold higher than in the 1950s.

The United Nations reports that weather-related disaster risk is affecting an ever-growing area, endangering decades of hard-fought economic growth and poverty reduction.[7] The frequency of hydrometeorological disasters, especially floods and droughts, has dramatically increased over the last two decades. Up from 150 disasters a year in the 1980s, there were more than 370 in the late 2000s.

The increasing frequency and intensity of flooding and windstorms are linked to climate change, with the number of disastrous floods and storms globally tripling over the past three decades. Very heavy precipitation increased sharply in the last half century across the globe and in the United States, especially the Northeast and Midwest.[8] Facing the floods of an altered climate, warming will likely fuel more water-related disasters.

Climate change, urbanization, and environmental degradation have often exacerbated the adverse impacts of natural disasters. More than 130 million people in China and roughly 40 million in Vietnam live along the coast and on low-lying islands. In Vietnam's

Mekong River Delta, an expected sea level rise of 30 centimeters by 2050 would increase seasonal inundations and salinity intrusions on more than 300,000 hectares of paddy fields, reducing rice production by 13 percent.[9]

Overcrowded, rapidly growing cities aggravate environmental insults and increase the risks to many of their inhabitants, especially the poor. Of the world's twenty-five megacities, fourteen are on the coast and seven are within a few hours' drive. The number of people exposed to storms and earthquakes in large cities could double to 1.5 billion by 2050.[10] Many less densely settled agricultural zones along the coast are vulnerable to tidal surges and the runoff from swollen rivers.

Environmental fragility caused by rising populations and changing land use over the past fifty years has greatly increased vulnerability worldwide, so that even small-scale hazards can produce large losses. A striking example is the loss of wetlands, particularly along coasts, where storms and storm surges are no longer met by energy-absorbing mangrove buffers. Not only do the mangroves offer this natural protection, but they also trap runoff long enough for it to deposit sediments, provide nesting for fish and crustaceans, and prevent coastal erosion. The world lost almost 5 million hectares of mangroves over 1980–2005—from 19.8 million to 15 million—as growing populations converted them to settlements, rice fields, tourist resorts, and aquaculture ponds. South and Southeast Asia have experienced particularly dramatic losses due to aqua-farming and other causes.[11]

Disaster Responses Should Not Be a One-Off Exercise

Natural disasters are not just interruptions to development—they are persistent risks. In many countries it is more a question of when and where exactly a disaster will happen, rather than whether it will happen at all.[12] Small island states in the Caribbean are hit repeatedly by hurricanes, Pacific Rim states in the ring of fire, by earthquakes and volcanic eruptions, and low-lying coastal areas on the Bay of Bengal, by flooding, and so on. Some flooding, storms, and fires are annual events, the only variation being the magnitude of the disaster. Preparing a strategy or action plan for natural disaster assistance that spells out the real long-term objectives of disaster prevention and mitigation—and includes an assessment of eachcountry's disaster risk—can make postdisaster interventions more effective.

Despite the recurrence of natural disasters, governments and international aid organizations do not systematically plan for preventing

Table 2.1
Too Few Country Assistance Strategies Discuss Disaster Prevention

Number of disaster response projects in a country (1984–2005)	Number of countries with this count	Number of strategies with no discussion of disaster prevention	Percent
More than 8	16	5	31
4–7	24	8	33
2–3	33	15	45
1	24	15	62
Total	97	43	44

Source: IEG (2006b).

them or mitigating their effects. Even countries regularly hit by natural catastrophes seldom consider the effects of their rising incidence, damage, and cost. During 1984–2005, ninety-seven countries borrowed from the World Bank for natural disaster response projects, but forty-three of them did not mention disaster prevention in their most recent development plans (Table 2.1). For the twenty-four countries with one disaster response project, a startling two-thirds of their strategies did not mention disaster prevention. Even in the forty countries with four or more disaster response projects, a third of their strategies did not mention prevention.

This one-off response thinking has large costs. Postdisaster reconstruction projects have not taken enough precautions to address the risk in some areas hit repeatedly by natural disasters. In Honduras, 56.5 percent of its GDP comes from an area at risk of two or more natural hazards. Estimates of the damage to roads from 1974's Hurricane Fifi were some $454 million.[13] But the design of infrastructure did not pay enough attention to prevention. Hurricane Mitch destroyed 6,000 km of the better roads in 1998, 60 percent of the total, and damaged or destroyed more than 163 bridges.[14]

In Haryana, India, the Water Resource Consolidation project[15] aimed to improve water distribution and drainage, but the project design did not sufficiently acknowledge that Haryana is flood-prone. Soon after the project started in 1994, rainfall and flooding from July to September 1995 severely disrupted initial project implementation.

All project staff and resources were diverted to flood damage repair and relief works, and project objectives were not met.

Distributing supplies by helicopter and building temporary homes can make headlines in postdisaster situations. Useful as they are in addressing urgent needs, however, their contribution to long-term results is limited. Evaluations of the responses to natural disasters show the tradeoffs between focusing on short-term needs and pursuing long-term objectives. Even in emergencies, actions should have an eye on the future.

Immediately after a disaster, most efforts rightly target reconstruction. The World Bank has demonstrated considerable agility in its approach to natural disasters, and Bank-financed natural disaster projects have had higher ratings for outcomes and sustainability than the Bank's overall portfolio. But past disaster assistance was most often reactive. Mitigating or even preventing future disasters was seldom among the objectives, with efforts providing short-term fixes and rarely addressing root causes.

Such one-off natural disaster response limits the capacity to react in future crises. Too often it is impossible to provide urgent postdisaster care because critical-care facilities are no longer functioning or people cannot reach service facilities. Without prevention systems, lifelines for potable water and first aid during calamities cannot be ensured, adding to desperation and breakdowns in order, even in well-off countries. The lack of sanitation and sewage systems can create conditions for water-borne diseases to spread and cholera to become epidemic. When floods hit West Africa in 2010, the humanitarian and health situation caused worldwide concern.[16] Inadequate institutional capacity, often overlooked in an emergency, can repeatedly constrain the effectiveness of hazard responses.

Lack of maintenance, also a consequence of a myopic view, has reduced the sustainability of structures rebuilt by postdisaster projects. Whenever massive reconstruction is needed following a disaster, the pressure for haste is high. But haste can result in incomplete reconstruction and account for much of the longer-term GDP cost of a disaster.[17] There also has been neglect—of fifty-nine completed emergency projects in disaster-prone settings, only ten have had follow-on projects.[18] Maintenance, follow-up, and preventive investments demand greater attention.

Prevention and mitigation now figure more prominently, with the lessons identified here featured in the new directions for natural

disasters at the World Bank.[19] The operational policy implemented in 2007 recognizes the importance of integrating risk reduction and crisis prevention into the development strategies of countries at high risk of disaster, with their country's assistance strategies now expected to include prevention and mitigation. This shift in emphasis signals important learning from experience.

This more proactive, strategic approach should produce longer-term benefits. Strengthening prevention and mitigation in noncrisis times is essential to reducing the damages of natural disasters, to improving preparedness, and to balancing the tradeoffs over time when reacting to immediate needs after a disaster hits.

Prevention and Mitigation Show Great Potential

Natural hazards and extreme events should not automatically lead to catastrophic loss of life and property. Prevention and mitigation can lessen damage. Intelligent preparation can facilitate an effective, immediate response that will make a vital difference to recovery.

The tendency to treat disasters as one-off, random events needs to be rethought.[20] Evaluative lessons confirm the urgent need to invest in climate change mitigation, disaster preparedness, early response, and postdisaster reconstruction.

Recent research in *Nature* established the connection between human action and the increasing frequency and magnitude of weather-related disasters. One paper identifies a human contribution to the observed intensification of extreme precipitation by examining the increase in heavy rainfall globally from 1951 to 1999.[21] Another demonstrates a link between climate change and extreme floods in Britain, suggesting that twentieth-century anthropogenic greenhouse gas emissions increased the risk of floods in England and Wales in autumn 2000 by more than 20 percent, and in two of three cases by more than 90 percent.[22] Yet another finds that human influence has at least doubled the risk of a heat wave as bad as the deadly European heat wave of 2003, then the hottest summer in at least 500 years.[23] To the extent that climate change is playing a role, its mitigation is critical to preventing and mitigating manmade natural disasters.

For far too long disaster-response has been focused on the R's of disaster management—relief, recovery, reconstruction—and not enough on the P's—prevention and preparedness. In 2002, faced with impending floods, Mozambique requested $3–4 million from donor countries to help it prepare. It received about half that amount.

After the floods struck, the same donors gave Mozambique more than $100 million in relief and pledged more than $450 million for recovery and reconstruction.[24]

Disasters are hazards of nature, but the damages are not completely exogenous. Human action in prevention and mitigation can help. Reconstructing housing with disaster-resistant techniques and accommodating the needs of occupants can reduce vulnerability to disasters and minimize the costs, especially in poorer countries where construction quality, land registration, and other regulatory mechanisms are weak. Enforcing building standards and improving information management, such as geographic data on hazards and vulnerability, could also lower the costs.

While both rich and poor countries suffer from natural disasters, poor countries typically suffer the greatest loss of life. Most of the 3.3 million disaster-related deaths over the last forty years were in poor countries, almost 1 million from Africa's droughts alone. Economic losses due to natural disasters are twenty times greater as a percentage of GDP in developing countries than in developed countries.[25] According to the United Nations Disaster Relief Organization, damage in small island states can amount to as much as 1,200 percent of a small island's GDP.[26] The reliance on poor-quality construction often leaves poor countries less prepared. When an earthquake hit rural Maharashtra State in India in 1993, house walls—four to five feet thick and made of heavy stones loosely bonded with dirt—crushed thousands of sleeping inhabitants.

The effects of a disaster are conditioned by a community's vulnerability to a given hazard (or conversely, its ability to cope with it). Prosperity improves preparedness. In Chile, in comparison with Haiti's experience a bit earlier, seismic-resistant construction helped prevent massive casualties and economic paralysis after the 2010 earthquake. High-level risk awareness and planning helped produce the no-casualty miracle in Australia after a category 5 cyclone in 2011.[27]

Preparedness can reduce vulnerability. Bangladesh achieved this by increasing its capacity to provide flood forecasting, early warning, shelters, and systems to evacuate areas most at risk. While a cyclone and floods in November 1970 killed 300,000 people, a comparable storm in May 1997 claimed only 188 lives, a dramatic improvement.[28]

The World Bank-funded East Pakistan (now Bangladesh) Cyclone Protection and Coastal Area Rehabilitation Project ($25 million) built

260 multipurpose buildings to serve as schools and cyclone shelters, elevated coastal feeder roads, and improved the meteorological cyclone warning system. As part of the effort, the Bangladesh Red Crescent Society identified 20,000 volunteers to convey early warnings from the national weather service (supplemented by satellite data) using easy-to-understand language. Warnings went out through mass media, mosque megaphones, extension agents on motorcycles, and signals to fishing villages.[29] Even so, if additional measures are not put in place, the damage from a single, severe cyclone is expected to rise nearly fivefold to more than $9 billion by 2050, affecting the poorest households most.[30]

Samoa, an island country in the Pacific, also shows success in prevention. In 1991, cyclone Val hit with maximum wind speeds of 140 knots, causing massive damage—equivalent to 230 percent of the country's real GDP. By contrast, cyclone Heta in 2004, with winds of up to 170 knots, took just 9 percent of Samoa's GDP. While the two cyclones were not directly comparable, having different tracks and durations, the effects of cyclone Heta would have been far worse if the country had not invested in risk management for natural hazards through the 1990s. The shoreline protection systems designed to cyclone standards performed well in the second disaster, sustaining minor damage, unlike adjacent areas with substandard coastal protection systems.[31]

In Vietnam, damages from cyclones were averted by reforesting coastal zones with mangrove trees to buffer against storms. Investments of $1.1 million in mangrove replanting and other measures saved some Vietnamese communities an estimated $7.3 million a year in sea dyke maintenance. During Typhoon Wukong in 2000, the project area in the northern parts of the country remained relatively unharmed while neighboring provinces suffered significant losses of life and property.[32] Relocating people from the areas vulnerable to typhoons to a buffer zone protected by dikes also improved incomes in adjoining communities and reduced destructive practices through resettlement activities, extension services, vocational training, credit, and social support.[33]

Poor construction is a major reason for the heavy loss of life from disasters in developing countries. In Turkey and Colombia, earthquake-resistant building codes, enforced construction standards, and oversight of materials procurement practices have paid off in a major way.

Turkey retrofits its houses to make them more disaster-resistant. It aims at preparing Istanbul for an earthquake through enhancing the institutional and technical capacity for disaster management and emergency responses, strengthening critical public facilities for earthquake resistance, and supporting measures for better enforcement of building codes. The World Bank's Istanbul Seismic Retrofitting Program[34] is designed to prevent buildings from collapsing and to save lives. Under the program, public schools and hospitals are not only made safe from collapse, but also renovated and improved. According to one estimate, for the cost of one new building, the program renovated and improved five or six.[35]

Colombia established a National System for Disaster Response and Prevention in 1983 to open the way for municipalities to invest in risk-reduction measures and address fiscal vulnerability in the country's national development plan. And everywhere, better land-use planning is ensuring that people are not building homes in harm's way.

By contrast, inadequate preparedness is associated with a higher disaster loss. In 1995, Pakistan's Federal Flood Commission developed a plan to support preparedness measures, but the $1 billion needed was not in place, coordination among government agencies was lacking, and monitoring of flood protection remained unsatisfactory. In 2010, a major flood caused $9.5 billion in damages, claiming 2,000 lives and leaving 20 million people homeless.

Cost-Efficient Measures Can Enhance Preparedness

The economics of implementing prevention measures are compelling. According to the Multi-hazard Mitigation Council, every $1 spent on loss prevention avoids an average of $4 in future losses.

Building a drainage system along a road may cost a little more when constructing the road, but it will preserve the road in future floods. Similarly, investing in regular road maintenance will cost the government, but it will not have to repeatedly reconstruct the same road, saving resources. Prevention measures can be difficult to design and implement. They might not be politically attractive because the benefits are only revealed when disaster strikes, while the cost is often high and always competing with other demands.

The 2010 report *Natural Hazards, Unnatural Disasters* shows benefit–cost ratios for Jakarta, St. Lucia, Istanbul, and Rohini Basin using assumed (but reasonably typical) costs: elevating a house with

mixed wall, concrete floor, and asbestos roof by one meter in Jakarta; protecting windows and doors in a wood frame house in Canaries, St. Lucia; retrofitting a five-story building to increase earthquake resilience in Istanbul; and flood-proofing a brick house by building with new brick on a raised plinth in the Rohini Basin, Uttar Pradesh, India.[36]

In all four cases, prevention would be cost-effective if the structure lasted ten years or more. For shorter time periods, cost-effectiveness depends on the discount rate; for low discount rates, the benefit–cost ratio is greater than one for some of these measures, implying that prevention is financially feasible.

Understanding why a prevention system is not in place helps address the main issues in achieving long-term results. If people do not grasp the importance or methods of prevention, disseminating information and sharing knowledge are necessary. If people choose not to follow the building codes because of an excessively high preference for the present (versus the future), pursuing a proper discount rate and building rational expectations are necessary. If people would like to take prevention measures but are financially forced to use substandard building materials, developing the financial system and providing access to resources to help overcome the liquidity constraints are necessary.

Better spending, not necessarily more spending, can achieve the desired results. Houses can be made earthquake-resistant using the same materials but in different ways. For example, rather than blindly putting stones together, incorporating bonding stones in the corners of rooms, using "through" stones and concrete layers at certain levels, anchoring timber beams, and building mud roofs with less weight can help. Doors that open outward can make the difference between complete destruction and little or no impact during an earthquake. Other lower-cost measures for prevention are available, including high wind-rated shingles, nails instead of staples, and straps that connect the roof, floor, and foundation.

Prevention can also include providing greater access to information, sharing data across borders, and reallocating public spending to road and bridge maintenance. Many countries are not taking advantage of technological improvements in weather and hazard forecasting. Even modest increases in spending—if supplemented by international data sharing—can have enormous benefits, especially in warning people of impending hazards. Several countries, some very poor, have made

large, quick gains from such spending. The success of the cyclone warning system in Bangladesh is one example.

Participating in a risk-pooling insurance facility is an important form of financial mitigation. Paying premiums may seem like an extra burden in noncrisis times, but it can pay off when disasters hit. Through the World Bank Caribbean Catastrophe Risk Insurance Facility, St. Lucia and Dominica received about $1 million within two weeks of the November 2007 earthquake, the most severe in the eastern Caribbean in thirty years. Turks and Caicos Islands received $6.3 million after Hurricane Ike hit in September 2008. Both payouts contributed to the rebuilding.

***

Work on natural disasters exemplifies a broader problem: focusing on immediate needs alone can undermine long-term goals. Evaluative evidence shows how responses to short-term objectives can drive the entire cause at the expense of the long-term ones.

Given the rising regularity and intensity of disasters, addressing this imbalance is urgent. We have seen how building preparedness can complement immediate emergency and reconstruction efforts. The World Bank has shifted some of its focus to incorporating disaster prevention into disaster-response projects around the world, as emphasized in the new strategy for natural disasters of 2010—a welcome step.

Why do the urgent often divert the important? Because there is greater political appeal in addressing short-term problems with immediate visibility, even at the expense of critical longer-term concerns for which the gains may only accrue after policymakers leave office. This is true not only for natural disasters, but also for other areas. Lasting solutions ensuring adequate attention to the future will depend on how effectively we deal with this question of political economy.

Notes

1. Dasgupta (2008).
2. IEG (2006b).
3. World Bank (2011c).
4. Asian Development Bank (2009a).
5. Jha (2010).
6. Thomas (2011).
7. UN (2009).
8. Thomas (2011).
9. Jha (2010).

10. World Bank (2010e).
11. Primavera (2005).
12. IEG (2006b); Asian Development Bank (2011).
13. Rückversicherungs-Gesellschaft (1975).
14. IEG (2006b).
15. India Water Resource Consolidation: Haryana (P009964) in IEG (2005b).
16. IEG (2010j).
17. Linnerooth-Bayer and Mechler (2005).
18. IEG (2006b).
19. Ibid.
20. World Bank (2006).
21. Min et al. (2011).
22. Pall et al. (2011).
23. Stott et al. (2004).
24. World Bank (2010e).
25. IEG (2006b).
26. Commonwealth Secretariat/World Bank Joint Task Force on Small States (2000). "The United Nations Disaster Relief Organization has estimated that the impact on the 25 most disaster-prone countries (including 13 small island states) has entailed costs between 28 percent and 1,200 percent of their annual GDP. For instance, in 1989 a series of cyclones in Vanuatu resulted in damage and economic losses amounting to twice the country's national income."
27. UN Daily News (2011).
28. IEG (2006b).
29. Bern et al.(1993); Cuny (1983); Independent Evaluation Department (2006).
30. Jha (2010).
31. IEG (2006b).
32. International Federation of Red Cross and Red Crescent Societies (2001).
33. IEG (2009k).
34. World Bank (2011d).
35. Jha (2010).
36. World Bank (2010e).

3

Connect Links that Strengthen Results

Whatever affects one directly, affects all indirectly.
—Martin Luther King Jr.

The focus of much development work is rightly placed on how well projects are designed, financed, and executed. Yet, doing projects in the right way is not enough to achieve satisfactory country and sector results. Stronger results emerge also from the nature of the policy regime and how projects are connected with each other in augmenting their effectiveness.

Positive project outcomes do not always translate into positive country outcomes, for many other factors can come into play. The country context, the interaction among projects, the scale of projects, and the policy setting are just some of the factors that can bear on country results. The aggregate sense of outcomes from a project-by-project review does not adequately reflect direct measures of achievement for the country. And on many occasions regional cooperation among countries can impinge on or enhance country results.

Similarly, sector outcomes are determined not only by projects in the sector, but also by cross-sectoral effects. In education, for example, good schools with good teachers do not guarantee effective learning if there are no roads for students to get to the schools. In health, improving access to safe drinking water has a better chance of reducing water-related diseases if there are also efforts to promote good hygiene. So, connecting efforts in related sectors can be decisive for achieving good results in a particular sector.

In short, country development is affected by many factors other than projects, such that the ratings at the two levels do not always correspond. After all, the objectives, scope, criteria, and measures at each level are different. There are also different actors and external

influences at work. Similarly, sectoral results (and thus ratings) are a function not only of project outcomes (and ratings) but also of intersectoral and other influences. This relationship highlights the interactions within and across sectors, including the links across microeconomic and macroeconomic factors, in deciding the outcomes for a country. The main finding is that some of these indirect and perhaps unintended effects can overwhelm the direct and intended.

Project Results and Country Results Differ

Achieving satisfactory project outcomes is not the same as achieving satisfactory country program outcomes. Empirically, country program ratings have differed from project ratings.

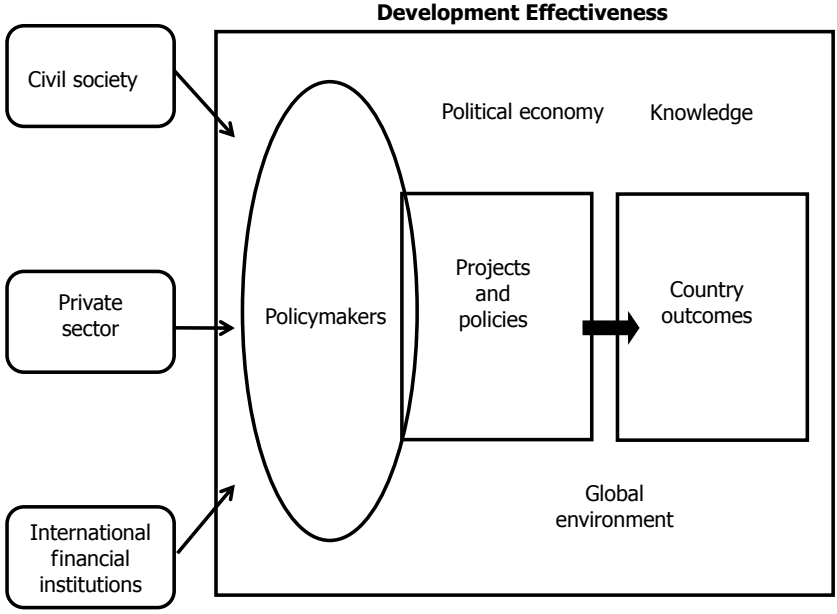
Project outcomes differ from country program outcomes for several reasons. Evaluative work is informed by an understanding of how country outcomes relate to project outcomes, but are shaped by other considerations, too. Country policymakers are central to this process (Figure 3.1). Project outcomes are frequently narrow or specific (such as access to schooling), whereas results are generally broader beyond the project (such as competitiveness). International financial institutions, the private sector, and civil society all play a role in development effectiveness directly and indirectly by influencing policymakers. Moreover, country program results are more often conditioned by interventions outside projects, some within and some beyond the control of the many players. Knowledge, the political economy, and the global environment affect both country and project outcomes.

Project and country program ratings are useful measures of outcomes at their respective levels. Project outcome ratings provided in project implementation completion reports assess the extent to which the project achieved or is expected to achieve its relevant objectives efficiently. Country program outcome ratings in reviews of country assistance strategy completion reports reflect the World Bank's contribution to country outcomes, or the results set out in the country assistance strategy. So, project ratings are not additive in producing country program ratings. Even when project ratings are high, outcomes at the country level may not be satisfactory—and vice versa.

Project and Country Program Objectives Differ

Project ratings and country program ratings measure different objectives, with no fixed relationship between the results frameworks

Figure 3.1
Political Economy, Country Program Outcomes, Projects, and Policies



for the project and the country. Outcome objectives for the country are broader and not a direct translation of those for the project.

A country evaluation must assess the size, composition, and type of lending, as well as other types of interventions. A country evaluation yields a more complete picture of the outcome of programs at the country level because it provides comprehensive coverage of the activities in a country during a given period. The country outcome may be unsatisfactory if there are critical omissions in the country assistance strategy, even if the project outcomes are satisfactory.

Country evaluations consider the relevance of the program, the achievement of objectives against standards set in the country assistance strategy, the quality of interaction with the government, and the quality and relevance of analytical work. Project evaluations assess whether the (narrowly) set objectives of the project are met. Whether the objectives in the strategy are achieved depends on the whole set of interventions that support the country's chosen objectives and program. In addition, country outcomes depend on the country's initiatives, such as policy changes and inputs from other development

partners, and developments such as natural disasters or financial crises—not solely on the success or failure of the projects.

The Chad–Cameroon oil pipeline project shows how project and country outcomes are measured against different objectives. The project was rated as satisfactory, technically well implemented, and a financial success. But the main objectives at the country level—building capacity to manage the petroleum sector and helping Chad reduce poverty and improve governance—were not met. To the contrary, the oil revenue windfall was associated with a resurgence of civil conflict and a worsening of governance. The main reason? There was a lack of government ownership of the project objectives, with repeated violations of the basic agreements.[1] No alternative program design or closer supervision would have achieved the program's objectives without a much stronger government commitment.

Doing projects right is an important part of achieving country outcomes, but factors beyond the project can have large, sometimes critical, impacts at the country level. In a review of all project and country evaluations since 1993, aggregate project outcome ratings were higher than aggregate country outcome ratings.[2] This difference holds when comparing the country program ratings with the project ratings in the same countries. As of April 2009, twenty-four of the eighty-eight reports completed had satisfactory aggregate project outcomes but unsatisfactory country program ratings (Table 3.1). The project ratings refer to those completed during the period, whereas the country ratings may reflect a broader set of projects as well as analytical and advisory activities. (The percentages of country programs and projects with satisfactory outcomes are higher after the implementation of

Table 3.1
Country and Project Outcome Ratings (Percent)

Implementation completion report reviews	Country assistance strategy completion report reviews	
	Satisfactory	Unsatisfactory
Satisfactory	64	27
Unsatisfactory	3	6

Note: Implementation completion report reviews include only investment lending and development policy lending. The outcomes of some recently completed projects are not included due to the time lags between project completion and review.

Source: World Bank database.

the results-based country assistance strategy. For the strategies that started after 2005, some 70 percent have satisfactory aggregate project outcomes and satisfactory country outcomes, 10 percentage points higher than for those started before 2005.)

There is substantial room for improvement in development effectiveness through more coherent, well-tailored country programs and through project improvements. There is no fixed relationship over time and across countries between the results framework at the project level and that at the country level. Even if the relationship between the project results frameworks and the country program results frameworks was correctly defined so that it is stable over time and across countries at the project level, factors other than project outcomes should, with the appropriate time lag, affect country program outcomes. Among the issues are the relevance of the country strategy and how the different kinds of interventions come together: policy dialogue, and complementarities with other sectors, with analytical and advisory activities, and with policy, lending, and global initiatives.[3]

Knowledge Services Affect Country Outcomes

Analytical and advisory activities, like lending, can also drive country outcomes. Such activities account for one-third of the World Bank's outlays for country services, exceeding those for lending or supervision. Economic sector work, technical assistance, and country dialogue contribute to country knowledge and performance through different channels. Project interventions are more successful when based on in-depth analytical work.[4]

One example of technical assistance having an impact is the help provided to Sri Lanka in instituting standardized small power purchase agreements to facilitate access to the power grid. Another example is the well-timed, high-quality knowledge products that helped Egypt formulate policy, reduce poverty, and develop human resources in the early 2000s, despite the World Bank's small financial contribution.

Similarly, analytical work, capacity building, and demonstrations contributed to favorable renewable-energy payment schemes, stimulating more than 20 gigawatts of installed wind capacity in China and hundreds of megawatts under construction in Mexico.

External Factors Affect Country Outcomes

External factors, besides affecting project outcomes, often play a large role in achieving country impact. Policies can have an

overwhelming effect on country programs. A recent review[5] indicates that country outcomes were correlated with country governance, measured by Country Policy and Institutional Assessment (CPIA) data, but not with country income, measured by GDP per capita. Just four of nineteen programs in countries with low CPIA governance scores (3.2 or less) had satisfactory outcomes, compared with 75 percent in those with high CPIA governance scores. When policies are off course, projects do poorly.

An evaluation showed a large difference between countries that borrowed from the World Bank for public sector reform and those that did not (Table 3.2). Overall, borrowers had a 73 percent improvement rate and nonborrowers a 48 percent improvement rate, though across regions the incidence of lending and the correlation of public sector reform lending with changes in governance scores varied.

The difference in CPIA scores between countries with and without World Bank public sector reform lending is large across all regions except Europe and Central Asia, where the improvement for countries getting public sector reform lending is the highest—90 percent—but the improvement for nonborrowers is almost as high. A common external factor explaining some of the performance improvement in

Table 3.2
Public Sector Reform Lending Can Produce Higher Governance Scores, 1999–2006

	With World Bank lending		Without World Bank lending	
Region	**Percent**	**Number**	**Percent**	**Number**
Sub-Saharan Africa	70	30	47	15
East Asia and the Pacific	70	10	56	9
Europe and Central Asia	90	20	86	7
Latin America and the Caribbean	75	20	25	8
Middle East and North Africa	57	7	0	2
South Asia	50	6	0	1
Total	73	93	48	42

Source: IEG (2008c).

this region seems to be the need to meet requirements for accession to the European Union. Almost all the countries in Europe and Central Asia that did not borrow for public sector reform in 1999–2006 were among the first from the East to join the European Union and had completed reforms before 1999.

Regional Approaches Can Improve Country Outcomes

Regional approaches can reinforce the national agendas and strengthen the aid architecture for multicountry efforts.

Regional programs offer good potential to achieve results on development issues that affect neighboring countries. Many development issues call for neighboring countries to work together—to manage water and other natural resources, facilitate trade and transport, provide reliable energy sources, and protect against the spread of disease and environmental degradation.

Some issues are best addressed from a regional perspective and some can be addressed only through regional cooperation. Focusing solely on a single country or sector cannot tackle the root issues. Regionally coordinated transportation development, for example, can help the world's landlocked countries connect to wider markets through neighboring countries. And resolving environmental problems often requires interventions across national or regional boundaries (as in the Mediterranean and Nile Basins).

The success of regional programs typically requires all relevant countries to overcome differing interests and past conflicts to achieve desired long-run benefits. But designing the programs is complex because benefits and costs need to be equitably assigned to participating countries, and regional and country activities need to be well coordinated during implementation.

An evaluation of regional programs found that in the past ten years such programs, which accounted for less than 3 percent of all international development support, performed as effectively in meeting their main objectives as single-country projects.[6] Adopting a more strategic role in supporting regional programs and integrating them into country assistance strategies could help countries realize their development potential. Africa has made the most progress with regional approaches and integrated country assistance strategies. The International Development Association (IDA), under IDA–15, made regional grants available, but the incentives and capacity for effective regional program support are still weak.

Making Vital Links Augments Results

Outcomes in one area, such as health, are often closely interlinked with those in other, related areas such as education or water and sanitation. A basic factor underlying such relations is the cross-elasticity of demand: that the demand for a service such as health responds not only to changes in its own price, but also to those of its complements (such as sanitation). Equally important is the cross-elasticity of supply. Resources can shift in response to price changes—for example, better schooling services might be provided when infrastructure around schools is improved.

Often outcomes in a sector result from outcomes in other sectors through multisectoral links. Take the Millennium Development Goals, which are highly interdependent and require effort in many sectors. Economic growth, determined by a variety of factors, is to a varying extent a driving force for the achievement of each goal. But improving one goal can improve the others in different ways. Achieving one goal often requires concerted effort from many sectors. A synergistic approach can often augment outcomes.[7]

This section first considers interdependencies, where results in several sectors can affect the outcomes in any one of them, with health as an example. Next, it looks at multisectoral interventions: in this case, objectives cut across sectors, as in postdisaster reconstruction. Then, it looks at circumstances where results depend on integrating efforts across domains, specifically the collaboration between the public and private sectors. Last, it considers circumstances where collective action is needed across government, civil society, and the general population, as in road safety.

Interdependence Affects Outcomes

The public may expect health mainly from the health sector, but the health sector needs inputs from other sectors, which may not necessarily subscribe to sharing responsibility for health improvements. The broad aspects of health and well-being are well beyond what the health sector can handle alone.

For example, public awareness through information sharing is critical for ensuring health outcomes. Only when people learn to wash their hands and to use clean water to prepare food, can improved access to water produce better health outcomes. Only when safety measures are in place, can improved road access to hospitals ensure better health outcomes. Only when public power is in place, can the

efficient maintenance of vaccine cold chains maintain the potency of vaccines to reduce child deaths and allow blood banks to save the lives of women hemorrhaging at delivery.[8]

Building health clinics is not enough. In some developing countries, making quality health care available may first necessitate ensuring that essential medicines are available. The challenge to guaranteeing a steady supply is not only related to the financial side of paying for medicines. Poor roads, limited communications, and storage problems make it difficult to keep medical facilities stocked with what they need to provide regular and life-saving care. In Zambia, a binding constraint is the medicine distribution system. Having district stores become transit points for shipments to health facilities reduced the out-of-stock rate for drugs to 1–33 percent, far below the previous 40–72 percent.[9]

Education is a necessary input to health. Better-educated women are more likely to understand and use disease-prevention measures, such as vaccines and mosquito nets. They are more likely to take a sick child to a clinic early and to follow treatment instructions. And they are more likely to understand germ theory and make clean water and sanitation a household priority.

Indeed, a mother's education affects her children's health in myriad ways.[10] An analysis of 175 countries between 1970 and 2009 found that a significant share of the reduction in child mortality over the past forty years can be attributed to the better education of women.[11] For every one-year increase in the average education of reproductive-age women, countries had their child mortality fall 9.5 percent.

Health sector outcomes are also linked to the water and sanitation sector. For example, a limited quantity of low-quality water and sanitation can impair health outcomes. Diarrheal diseases alone accounted for an estimated 1.6–2.1 million deaths annually over 1990–2000.[12] Diarrhea is one of the top five preventable causes of under-five child mortality in developing countries.[13] Poor sanitation, lack of access to clean water, and inadequate hygiene seem to account for about 90 percent of the spread of childhood diarrhea.[14]

Increasing the quantity and quality of water can reduce person-to-person water-washed transmission, fecal-to-oral waterborne transmission, and insect vector water-based disease transmission.[15] In Nigeria, water-related diseases constitute about 80 percent of the total disease burden.[16] In India, hand-washing has significantly reduced the prevalence and duration of a measure of overall diarrhea

as well as acute watery diarrhea among children under age five, but not acute dysentery.[17] In Africa and Southeast Asia, improvements in water, sanitation, and hygiene could reduce the total disease burden by 4–5 percent.[18] Improved hygiene is associated with a median reduction of 33 percent in diarrheal illness, with a range of 11–89 percent.[19]

Interlinked development programs can improve health behavior and health outcomes. In Bolivia, where Save the Children worked, children in households participating in health, credit, and literacy programs were significantly less likely than children from comparison communities participating in health-only programs to be malnourished or at risk of becoming malnourished, even after controlling for social class, the source of drinking water, and the availability of health facilities.[20]

But synergies need to be further understood and exploited. From 1997 to 2006, the World Bank invested about $5 billion in health, nutrition, and population components in 350 projects managed by other sectors, such as social protection, education, public sector management, water supply, and transport.[21] Yet, transport and water and sanitation projects with health components seldom involved collaboration with a health ministry or the World Bank's health, nutrition, and population sector. While the potential for improving health outcomes is great, it rarely serves as the primary objective driving water supply and sanitation project design and implementation. Only one in ten projects had an explicit objective to improve health. In addition, the health benefits are poorly documented—health improvements recorded are often attributed to the project, without accounting for health sector investments and other factors affecting health outcomes.

Only half the approved water supply and sanitation projects during 1997–2001 cited potential health benefits, while close to 90 percent of the projects financed infrastructure that could improve health.[22] Projects approved later (2002–2006) were even less likely to have been justified by health benefits, to have explicit health objectives, or to plan to collect health indicators. They were also less likely to target behavior change, critical for transforming infrastructure improvements into sustainable health gains. Among twenty-six completed projects, only four had documented changes in the prevalence or incidence of disease. Fewer than half the projects included behavior change objectives or activities.

Collaboration between health and transport can also be strengthened. On the positive side, there is evidence of strong collaboration between health, nutrition, and population and the transport sector on

road safety. But projects with road safety objectives or components often only documented improvements in accident and fatality statistics, with few intermediate outcomes or outputs to point to the causes of improved outcomes. In addition, projects with AIDS components rarely documented outputs or outcomes—there is almost no information on their effectiveness.

Multisectoral Connectivity Is Essential

Many interventions are multisectoral, with one sector possibly taking the lead. Consider postdisaster reconstruction. The response may span multiple sectors and themes, including urban, rural, environment, infrastructure, education, health, and social protection. An intervention need not comprise all aspects nor one agency needs to do it all, but it is important to connect the dots.

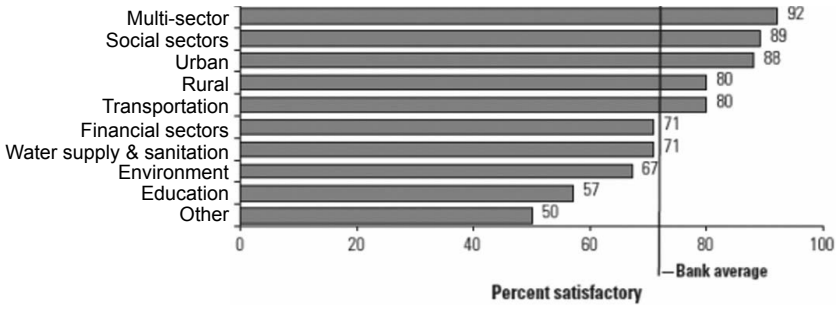
Immediately after the emergency stage, the direct and indirect effects of the event must be assessed for the social well-being and economic performance of the affected country or area. To craft a plan for reconstruction requires expertise from multiple sectors, considering the various constraints to restoring livelihoods and balancing short-term and long-term needs. This should be performed by joint teams, involving the people and institutions affected.

Getting the concept right at the outset can prevent some high or even irremediable costs in the future. For example, the layout of temporary shelter structures should consider gender-related safety. To reduce crime and violence against women, the relocation process should ensure that as many doors as possible face a common and well-lit area—avoiding the creation of passages and alleyways that are dark and poorly observed.

Between 1984 and 2005, the World Bank had 303 closed projects with disaster response activities, among which the twelve multisectoral projects performed best (Figure 3.2). The share of the latter projects rated satisfactory is the highest (92 percent), followed by social sector projects (89 percent) and urban sector projects (88 percent).

Evaluative lessons, often learned the hard way, suggest that rather than relocating communities, in situ reconstruction should generally be promoted to use existing infrastructure and community facilities and to minimize resettlement and social dislocation. It is also important to have timely social protection measures to assist survivors with jobs and cash transfers, providing not only financial means for food and clothing, but also helping survivors recover emotionally.

Figure 3.2
Project Outcomes by Sector



Source: IEG (2006b).

Rebuilding infrastructure is critical for providing services that support economic growth and reestablish normal life, generating employment, and facilitating the movement of goods and people. But reconstruction goes far beyond rebuilding physical infrastructure; it can also preserve social relationships. What counts is not only what is rebuilt—but also how and where it is done.

A recent evaluation[23] indicates that after the Haiti earthquake, the breakdown of social order, a fragile security situation, the near-complete loss of governance structures, and the failure to impose even minimum quality standards on the construction industry added to the tasks of recovery. Rebuilding homes and communities requires the safe transport and storage of building materials and, often, the formation of community groups that work to rebuild houses and infrastructure.

Distributing emergency supplies needs to be orderly, involve local leadership, and help maintain social cohesion. In Colombia, after the Armero eruption, and in Grenada and St. Lucia, after hurricane Ivan, families that did not lose their dwellings took in friends and relatives. These steps provided immediate relief to those who lost their homes and helped preserve existing social relationships to the long-term advantage of communities.

Flood response programs should focus not only on rebuilding infrastructure, but also on better adaptation and preparedness for the future in complementary investments, such as water and flood management,

cropping pattern adjustment, rural finance, enhancing the capacity of water user groups, and early warning systems.

The 2010 floods in Pakistan offer similar lessons. Restoring rural livelihoods requires restarting cropping and livestock activities and addressing land rights. (The topography in some areas may have changed, and land-rights documents may have been lost by households and the administration.)

Global and regional quick response teams, a "callable roster" of individuals who can be called on quickly in the immediate aftermath of a natural disaster, can enhance the effectiveness of disaster responses. They can also help develop a recovery strategy to facilitate collaboration among governments, nongovernmental organizations, and multilateral and bilateral donors.

Disasters typically attract many donors. During 1984–2005, about 34 percent of completed and ongoing disaster projects involved donors other than the World Bank, and 38 percent involved nongovernmental organizations. Because donors tend to focus on different sectors, effective coordination of expertise is crucial.

Public–Private Links Are Often Key

Multisectoral approaches are also relevant to the links between the public and private sectors. Across the range of country incomes, private spending accounts for more than half of all health expenditures in about 47 percent of low-income countries and about 51 percent of lower middle-income countries.[24] Public–private partnerships offer new approaches to service delivery, though evaluative evidence on institutional and financial sustainability is still limited.[25]

Given the private nature of agricultural activities and the public-good nature of agricultural services, particularly agricultural research and extension, the extent to which interventions link government and private producers makes a difference for performance. The impact of interventions by governments and international institutions will only be as good as the links with private producers.

The complex crop-production chain linking farmers to consumers requires collaboration across multiple sectors. Weakness at any point within and between the public and private sectors can hinder agricultural and agribusiness productivity. Underinvestment in research and extension, water constraints, poor rural transport infrastructure, limited access to credit, land issues, market support, and agribusiness activities all constrain agricultural productivity.[26] For

example, market-based private investments in agriculture can drive technical change, generating new sources of demand for innovation. And international institutions can capitalize on the opportunities for agriculture and agribusiness by bringing partners together to deliver practical, market-based solutions.[27]

Another example is the safeguard policies of the World Bank Group. Drawing lessons from the public and private arms can harmonize thematic coverage and guidance. Adopting strong features from each approach can improve implementation, results, and benefits. That is why IEG encouraged the IFC, the MIGA, and the World Bank to adopt and use a shared set of objective criteria to assess social and environmental risks to ensure adequacy and consistency in project categorization across the World Bank Group.[28]

For better results, the World Bank especially needs to strengthen the supervision, monitoring, and evaluation of safeguards, drawing on recent IFC experience. By the same token, the IFC and MIGA must ensure third-party verification and full and timely public disclosure—as the World Bank is poised to do—for credibility and better results in the social and environmental areas.

Collective Action May Be Necessary to Achieve Results

The scope of some issues could be so large that, beyond the key sectors, collective action by the government and society is required. Road safety, an issue of growing importance for public health and economic development, highlights the critical roles of political will and joint stakeholder efforts.

Economic growth creates demand for faster modes of passenger and freight transport. According to the International Energy Agency, the road vehicle population will grow from 170 million in developing regions in 1996 to 454 million in 2020.[29] Building more and faster roads helps move goods to the market and improve access to jobs, education, health care, and social and leisure activities. But this requires more than roads and vehicles. If safety regulations are not in place, faster travel speeds can increase traffic accidents.

The World Health Organization has declared road traffic injuries a major public health and development crisis.[30] Every year 1.2 million people are known to die in road accidents worldwide (more than 3,000 a day) and as many as 50 million more are injured, with some suffering permanent disabilities. It is predicted that road traffic injuries will become the fifth leading cause of death in 2030, accounting

for 3.6 percent of all deaths, up from ninth in 2004 and 2.2 percent.[31] The most productive age group (15–44 years) has the highest injury and death rate.[32]

The combination of unsafe vehicles and driving habits and poor road conditions translates to injuries and deaths, with a particularly heavy toll on those who can ill-afford to pay the bills. Low- and middle-income countries have higher road traffic fatality rates (21.5 and 19.5 per 100,000 people, respectively) than high-income countries (10.3 per 100,000). The economic cost of road crashes and injuries is estimated to be 1 percent of gross national product in low-income countries and 1.5 percent in middle-income countries.[33]

There is great potential to improve road safety regulation. In Australia, implementing and enforcing seatbelt laws helped reduce alcohol-related road deaths by 40 percent. And in Asia, mandatory helmet use can reduce deaths among motorcyclists by 30–40 percent.[34] Similarly, introducing speed limits, creating safer infrastructure, and enforcing blood-alcohol content limits can all sizably reduce road traffic injuries and deaths.

Effective intervention requires concerted efforts by multiple partners (governments, nongovernmental organizations, and the private sector) across many disciplines (transport, health, law enforcement, and urban planning), with a focal point or agency coordinating activities. Local and national governments can take the lead in improving public infrastructure, formulating public policy, and ensuring enforcement. Civil society can advocate and educate the public on road safety issues. The private sector can produce safe vehicles and engage communities. And people can adopt safer behaviors. The World Health Assembly resolution on road safety and health recommends that World Health Organization members facilitate multisectoral collaboration among different ministries and sectors.

Strong political will is crucial in adopting and enforcing traffic laws. Reducing drunk driving and excessive speed, while increasing the use of helmets, seatbelts, and child restraints, are among the key measures for improving road safety. But fewer than half of countries have laws to address these five risk factors, and only 15 percent have laws considered comprehensive in scope. And while many countries have improved the institutional frameworks to support road safety, challenges remain. Only one-third of countries have a government-endorsed national road safety strategy that includes specific targets and funding allocated for implementation.[35]

There is indication of recent progress in a few countries. In 2003, China suffered 220,000 road traffic fatalities, 18 percent of the total and the world's highest, with less than 5 percent of the global vehicle fleet. The government has since launched strong initiatives to improve the situation. Key measures include the creation of the Inter-Ministerial Road Safety Forum in October 2003 and the implementation of the Road Traffic Safety Law in May 2004.[36]

Road safety is gradually drawing more attention globally and collaboration is improving. Countries realize that it can be dealt with effectively only when tackled across sectors. The United Nations' Road Safety Collaboration and the World Bank's Global Road Safety Facility have brought together all the regional commissions and a wide range of road safety stakeholders to exchange good practices and provide a coordination mechanism for key multilateral and bilateral donors. A series of global events is being launched for the Decade of Action for Road Safety this year. More than ninety countries have committed to saving 5 million lives and preventing 50 million injuries in developing countries—for an estimated benefit of more than $3 trillion.

***

To achieve good results in a particular sector, putting all efforts into that sector alone may not be adequate or even optimal. The general impact of economic growth and rising incomes on a range of social indicators illustrates the broad reach of cross-sectoral links. Interventions in infrastructure and human development show that accommodating relationships among them involving one or more public or private development partners is needed to achieve the desired development outcomes.

It is common that people work in isolation from each other and that their reward structures do not adequately recognize efforts to work across boundaries. Often, operational setups in countries and multilateral or bilateral financial institutions do not provide the motivation or incentives to take advantage of links across areas of work, even when the overall results might improve from such interactions. Changing this situation through incentives, behavioral encouragement, and structural reforms should have high payoffs.

Notes

1. IEG (2009i).
2. IEG (2009a).
3. IEG (2010e).

4. IEG (2008d).
5. IEG (2011c).
6. IEG (2007d).
7. IEG (2010d).
8. Adeleye and Ofili (2010).
9. Vledder et al. (2010).
10. Gakidou et al. (2010).
11. Ibid.
12. Overbey (2008).
13. Keutsch et al. (2006).
14. WHO (2004).
15. Overbey (2008).
16. National Population Commission (2004).
17. Fan and Mahal (2011).
18. Cairncross and Valmanis (2006).
19. Huttly et al. (1997).
20. Gonzales et al. (1999).
21. IEG (2009c).
22. Overbey (2008).
23. IEG (2010j).
24. World Bank (2009).
25. IEG (2010d).
26. IEG (2010c).
27. ECG (2011).
28. IEG (2010f).
29. Bekefi (2006).
30. WHO (2004).
31. WHO (2008).
32. See note 29.
33. Freeman and Mathur (2008).
34. Krug and Toroyan (2006).
35. WHO (2009).
36. World Bank (2011a).*Source*: "Rawlings et at, [2004] should be placed at the bottom of the figure after prefaced with

Part II

Measuring Results

Part I focused on identifying the results to guide development efforts led by countries and supported by financial agencies. But doing this could be of limited value if we are not measuring those results properly. Poor monitoring and evaluation can send the wrong signals and not achieve the desired outcomes by misallocating scarce resources from higher-value to lower-value activities. But how to measure results is less obvious in practice than it is in principle.

At least three things can render measurement inappropriate. First, when there is weakness in the assumptions and methods, a composite indicator can be a poor proxy for what it aims to measure. Second, even if the right measure exists for the desired results, when the distribution is skewed, the targeted population can be left out or wind up worse off, even if the results are achieved on average. Third, even if the indicators are well defined, when the elements in the results chain are missing or links are broken, it would help to go beyond intermediate outcomes to assess the likelihood of reaching the desired results.

The motivation for ignoring these factors seems to come from the drive to carry out ever more specialized academic research and from policymakers looking for a quick fix and easy monitoring of progress. Inadequate measures may also be a matter of convenience, if not a means to promote an agenda. The antidote to these tendencies would be to continually integrate policy initiatives and build robust feedback loops so that new and contrarian findings are incorporated to reorient or halt programs.

This part addresses these issues from three perspectives. It begins by considering some of the problems in constructing, using, and interpreting composite indicators. It next examines problems associated with measuring intermediate rather than final outcomes. It then turns to questions around measures that rely on averages of the general population when the real target is a specific group.

4

Composite Indicators Can Mislead

Most businesspeople are upright citizens; but that does not change the fact that business is conducted for private gain and not for the public benefit.

—George Soros

Indicators are an essential aid to all forms of routine evaluative assessments. Discussions of country performance, for example, are hard to imagine without measures of GDP or inflation. The measure of GDP is derived from various measures covering consumption and investment income, government spending, and international trade. The U.S. Department of Labor's widely cited consumer price index is constructed from many components representing different kinds of consumer spending based on data from local samples.

Obviously, composite indicators are used extensively in many contexts and can be useful for a wide range of purposes. They are appealing because a single number can capture many dimensions of interest.

But the gap between what a composite indicator claims to measure and what it actually captures can be very wide. And the underlying hypotheses, the coverage of the indicator, and the methods used can widen it. Rigor can easily be sacrificed to data constraints. Simply lumping together several available datasets and assigning them weights without adequate justification and transparent documentation can result in "mashup indices."[1]

The general inclination to take a composite indicator at face value—interpreting it according to what it claims to measure without looking at how it measures and whether it measures validly and reliably—can lead to misleading conclusions. Worse, when the documentation of how the indicator is constructed is unclear or missing, the exercise can become a matter of guesswork.

The major concern of this chapter is whether widely used composite measures correctly capture what we need to measure to make sound development decisions. It mainly draws on evaluation research on the World Bank Group's Country Policy and Institutional Assessment and Doing Business indicators, the United Nations Development Programme's Human Development Index (HDI), and the United Nations' Multidimensional Poverty Index. The discussion covers five issues: inconclusive premises, partial coverage, arbitrary clustering and weights, the conversion of ratings to rankings, and de jure versus de facto indicators.

Inconclusive Premises Affect the Validity of Indicators

Whether an indicator can deliver what it aims to measure depends on the soundness of its premises. There needs to be a foundation for the relationship claimed between the indicator constructed and the outcomes it purports to measure. Failure to establish a link between what is measured and the expressed meaning of the composite indicator risks weakening its relevance. One area where evaluation would add value is identifying the underlying assumptions behind such a link.

The Doing Business indicators (Box 4.1) embody three ideas: less regulation is preferable, property rights and debt enforceability are important determinants of lending and investment, and lighter regulation and taxation can encourage informal firms to shift into the formal business sector. The literature is inconclusive on the causality of any of these three in business growth and job creation.

Box 4.1 Doing Business Indicators

The Doing Business indicators mainly measure the existence of laws and regulations that govern the startup, operation, and growth of businesses. They cover ten dimensions of the cost to firms of business regulations in more than 170 countries. Seven of these indicators assume that less regulation is better. The indicator is a simple average of the ten dimensions, themselves simple averages of their respective subcomponents. A single-country ranking number is created by successive stages of ordinal rankings translated from the cardinal values for each subindicator. They cover time, costs, number of procedures, and the like.

First, seven of the ten indicators presume that less regulation is better. But regulations also generate social benefits—such as safety, environmental protection, worker protection, and transparency—not only private costs. The recent financial crisis was, at least in part, a consequence of too little regulation, rather than too much. Depending on whether the country starts with a little or a lot of regulation, reducing regulation is not always better for society—for example, for employing workers, dealing with licenses, and paying taxes.

Furthermore, what is good for an individual firm is not necessarily good for the economy or society. Doing Business assumes that the fewer the steps needed to get a permit the better, and that possible benefits from safety and environmental checks can be ignored. It also assumes that the lower the overall tax rate as a share of a firm's profit the better, but this overlooks each country's fiscal requirements to raise revenue and redistribute income for equity reasons. For example, Maldives, one of the top ten countries on the subindicator "paying taxes," has no corporate income tax. But this should not be a role model for other countries—most public revenues in Maldives are raised from resort leases.

Second, Doing Business posits that poor property owners are locked out of the formal economy because they lack legal rights to their land and thus cannot use it as collateral for loans to expand their businesses or improve their properties. This derives in part from the work of Hernando de Soto. Five of the ten Doing Business indicators measure the enforceability of debt contracts and availability of collateral, including getting credit, enforcing contracts, registering property, closing a business, and dealing with licenses. But the literature does not demonstrate a causal relationship between these factors in the business environment and, ultimately, growth.

Third, the Doing Business indicators hypothesize that lighter regulation and less taxation encourage informal firms to move into the formal economy. But the literature is inconclusive about why the informal sector exists and persists and whether formalization can create more jobs and lead to higher economic growth. For example, reducing regulatory obstacles to starting a business will not necessarily encourage informal firms to formalize if they face high barriers to entry caused by their low skills and lack of access to capital. Some studies find no significant relationship between reforms as measured by changes in the Doing Business indicators and aggregate investment and unemployment rates.[2]

Doing Business can stress that regulations also have social benefits not captured by the exercise and the index is not a measure of regulatory reform. Some efforts are being made to indicate more transparently the limitations of what Doing Business measures and avoiding its identification as a measure of overall reform—steps in the right direction that need to be reinforced. And Doing Business is planning research on the connection between Doing Business indicators on firm performance and other country-level impacts.[3]

Partial Coverage Should Not Unbalance the Perspective

No composite indicator can or should include every dimension of what it tries to describe. But in some cases an indicator may provide an especially narrow focus of the underlying variables while claiming to do much more, thus creating a biased picture. If critical elements are missing or if minor ones are overemphasized, the resulting perspective would be unbalanced. The Doing Business indicators show how the partial coverage of the reform agenda can conflict with what the indicator aims to represent.

The ten dimensions that Doing Business measures are certain aspects of the investment climate—comprising mainly the laws and regulations that govern how firms do business.[4] They partly cover the factors that condition business climates and the types of firms targeted. The partial coverage itself is not a concern, as no index is to be expected to be comprehensive. The issue is the validity of what this composite indicator claims to measure.

While focusing on the importance of collateral and the enforceability of laws, Doing Business indicators omit many factors that affect firms' actual use of credit, such as the rate of interest, value of the assets, degree of intermediation, and existence of viable entrepreneurial opportunities. Nor is it clear what regulations matter in comparison with other determinants of the business environment, such as infrastructure, labor skills, and competition policies.[5]

The sample coverage is also limited: small- and medium-sized firms, formal sector firms, domestically owned firms and domestic investors, official and legal transactions and processes, firms in the capital city, and limited liability companies. But it excludes other types of firms and transactions. The same constraints may have different implications for different types of firms—for example, some regulatory constraints, such as protecting investors, are likely to be less important for informal sector entities and microenterprises.[6]

With this coverage and scope, the concern is about what the index is positioned to measure, reflect, and convey. Even using the nomenclature "ease of doing business" overstates what the indicator sets out to measure. More important, its shift from a partial interpretation of "measuring aspects of the regulatory environment for business" to claims of reflecting "regulatory reform" if not "country reform" is highly problematic. While the lively communications style has helped give the index an international profile and attracted the interest of senior policymakers, its presentation as a measure of reform could mislead its users about what the indicator really portrays and what interpretations can be drawn from it.

Arbitrary Clustering and Weights Can Have Pernicious Effects

The measurement of a single objective often includes many aspects. The Millennium Development Goals are an example. Their vision is multifaceted: a world in which developed and developing countries work together for the betterment of all, with less poverty, hunger, and disease, greater survival prospects for mothers and their infants, better-educated children, equal opportunities for women, and a healthier environment. The eight goals are measured by twenty-one targets, each with a multidimensional view, and more than fifty measurable indicators.

Consider the first goal: "eradicate extreme poverty and hunger." This includes three targets: halve the proportion of people living in poverty, achieve full and productive employment and decent work for all, and halve the proportion of people who suffer from hunger between 1990 and 2015. And there are nine indicators for monitoring progress against the three targets. For example, to measure the proportion of people who suffer from hunger, information is required on malnourishment, such as wasting and being underweight, from various age groups.

The message conveyed from the separate indicators that measure various targets could point in the same or different directions. For example, a country making good progress in employment may also be doing well in poverty reduction. Or it may not be doing as well in reducing poverty if the improvement in employment opportunities is skewed to the nonpoor, or if the increase in the quantity of jobs lacks the needed quality to lift workers and their families from poverty.

It can be tempting to construct a single composite index to simplify comparisons over time and across countries. To collapse a range of

indicators measuring various aspects of interest, different elements must be grouped and weights assigned. But adding up multiple indices of different dimensions does not always provide a more complete picture. Rigor can be compromised by arbitrary weights and clustering.

The weights, along with the ratings of each element, determine the value of the indicator. But their impact on the value of the indicator is less straightforward. For example, the World Bank's CPIA index and IDA's performance-based allocation have the same cluster elements but with different weights. The difference in country values between the CPIA and performance-based allocation demonstrates the role of weights in building a composite indicator (Box 4.2).

Box 4.2 Country Policy and Institutional Assessment

The World Bank's Country Policy and Institutional Assessment (CPIA) assesses how conducive a country's policy and institutional framework is to poverty reduction, sustainable growth, and the effective use of development assistance. It enters into a performance-based allocation formula for making International Development Association (IDA) resources available to eligible countries. Both the CPIA and performance-based allocation have the same clusters: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. The CPIA applies equal weights to each of the four clusters; the IDA allocation formula gives equal weight (8 percent) to the first three clusters and to portfolio performance, but a much higher weight (68 percent) to the governance cluster.

The CPIA's sixteen criteria are grouped in four clusters—economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions—weighted equally in the overall rating. In the IDA allocation formula—8 percent on each of the first three clusters, 68 percent on the fourth (governance), with the remaining 8 percent weighted on portfolio performance.

It might seem intuitive that a larger weight on governance in the IDA performance-based allocation formula benefits countries with better governance. But a simulation revealed that the effects of the much larger weight on governance in the formula are due not just to

the governance rating but to how different the governance rating is from ratings on other clusters.[7] A larger weight on governance than the other three elements in the formula does not necessarily make a country worse off in IDA allocations, even if its score in governance is lower than its scores in other dimensions. Indeed, all core IDA countries (excluding small states) have governance ratings that are worse than their ratings on other clusters, yet some countries gain while others lose from the larger weight on governance.[8] Whether they gain or lose depends on how much worse the ratio of their governance ratings is to ratings on other clusters in comparison with other countries.

The HDI is another example (Box 4.3). Until 2010 the HDI was an equally weighted mean of uniformly scaled attainments in life expectancy, education, and income. The 2010 HDI relaxed the assumption of perfect substitutability among its three components and switched from the original additive aggregation function (the arithmetical mean of the three components) to a multiplicative function (their geometrical mean). Theory offers no justification for any particular set of weights. The change in weights results in a significant reduction of the weight on longevity in poor countries. Based on the new method of constructing the index, a poor country experiencing falling life expectancy due to the collapse of its weak health care system still could see its HDI improve with even a small rate of economic growth.[9]

Box 4.3 Human Development Index

The Human Development Index (HDI) is a composite indicator measuring development as a single statistic with three dimensions: life expectancy, educational attainment, and income. The HDI is the average of the scores for the three HDI dimension indices, which are themselves averages of their respective subindicators. For example, the education component is the mean of years of schooling for adults aged twenty-five years and expected years of schooling for children of school age. The HDI is used to rank countries as developed (high development), developing (middle development), or underdeveloped (low development). Before 2010 the HDI was based on the arithmetic means of the subindicators, but since 2010, geometric means have been used.

Ravallion (2011b) uses the Multidimensional Poverty Index to illustrate some (perhaps peculiar) underlying assumptions behind lumping various aspects together and assigning ad hoc weights to compose a single indicator (Box 4.4). The Multidimensional Poverty Index complements income poverty measures and reflects the multiple deprivations that a poor person faces in education, health, and living standards. A household is defined as poor if it is deprived across at least 30 percent of the weighted indicators.

Box 4.4 Multidimensional Poverty Index

The Multidimensional Poverty Index assesses poverty at the individual level, with the poor being those who suffer deprivation, with the extent of their poverty measured by the range of their deprivations. Following the concepts developed in the United Nations Development Programme's Human Development Report, the index focuses on the three main aspects of deprivation: education, health, and living standards. It is an equally weighted aggregate poverty measure for each aspect, together comprising ten dimensions:

Education (each subindicator is weighted equally at 1/6)

1. Years of schooling: deprived if no household member has completed five years of schooling.
2. Child enrollment: deprived if any school-age child is not attending school in years one to eight.

Health (each subindicator is weighted equally at 1/6)

3. Child mortality: deprived if any child has died in the family.
4. Nutrition: deprived if any adult or child for whom there is nutritional information is malnourished.

Living standards (each subindicator is weighted equally at 1/18)

5. Electricity: deprived if the household has no electricity.
6. Sanitation: deprived if they do not have an improved toilet or if their toilet is shared (Millennium Development Goal definition).

Box 4.4 *Continued*

7. Drinking water: deprived if the household does not have access to clean drinking water or clean water is more than a thirty-minute walk from home (Millennium Development Goal definition).
8. Floor: deprived if the household has dirt, sand, or dung floor.
9. Cooking fuel: deprived if they cook with wood, charcoal, or dung.
10. Assets: deprived if the household does not own more than one of: radio, TV, telephone, bike, or motorbike.

By applying equal weights to education, health, and living standards, the Multidimensional Poverty Index implicitly attributes the same importance to being education-poor, health-poor, and living standard-poor. Ravallion argues that health is valued intrinsically, independent of command over commodities. Because education and health each have two subindicators, and living standards has six, each education and health subindicator is given a 1/6 weight, while each living standard subindicator gets only a 1/18 weight. This indicator thus implies that avoiding the death of a child is equivalent to alleviating the combined deprivations of having a dirt floor, cooking with wood, and not having a radio, TV, telephone, bike, or car. Or that attaining these material conditions is equivalent to an extra year of schooling or to not having any malnourished family members.[10] It is not at all clear that such tradeoffs validly measure the experience of poverty.

Recognizing the multidimensional measures of an objective does not imply that it is necessary to collapse all aspects into a single indicator. When creating a composite indicator, being clear about the data and methods behind the formulation of each element and providing information on each subcomponent can allow users to better understand what the data actually measure.[11]

The enthusiasm for composite indicators needs transparency from their developers and critical scrutiny from users. Although not a solution to the problem of which weight to apply, documenting each of the subindicators can help inform users. The Quality of Official

Development Assistance assessment, constructed in four dimensions of aid quality built up from thirty separate indicators, documents each indicator for users' information.

Converting Ratings to Rankings Can Obscure Findings

After collapsing the many indicators into one single composite index, it is often appealing to further convert the cardinal number into an ordinal one to answer the question of "where a country stands" among its comparators. The idea is that for cross-country comparisons it may seem more intuitive to explain a country's performance in, say, export clearance, by reporting that "it ranks in the middle in the world," rather than to explain that "it takes 20 days to clear exports," and relying on readers to do the detailed comparison with other countries.

The Doing Business indicators can show the pros and cons of converting a composite indicator from a rating to a ranking. Each ranking is derived from an indicator using cardinal values from its ten subindicators: time, cost, number of procedures, and so on. These cardinal values are ranked according to their respective percentiles in each of the subindicator distributions. The subindicator percentiles are then averaged to come up with an indicator-level percentile; the ten indicator percentiles are then averaged to generate the overall "ease of doing business" ranking.[12] By ranking countries on selected dimensions of business regulation and spotlighting both leaders and laggards, the Doing Business report has attracted considerable attention. But like any rating exercise, it also has provoked concern.

Some useful information, such as the extent of the difference between one cardinal number and another in the distribution, is often lost when rescaling from a rating to a ranking. Relying on successive stages of ordinal rankings to derive a composite indicator can obscure the underlying cardinal values. For example, for paying taxes, there is a 5.1 percentage point difference between the top performer, Maldives, and the next, Vanuatu, while the difference between the ten countries ranked fiftieth and sixtieth is only 0.1 percentage point. Countries typically are ranked differently based on different subindicators and the cardinal difference between countries in different parts of any one ordinal distribution varies widely.

The difference in cardinal measures also varies widely for the same difference in ordinal rankings. A country's location in the distribution affects how a reform will change its ranking. Countries can make significant changes but fail to improve their rankings if they are at the

ends of the distribution for that indicator. The change in ranking for any country is driven largely by where the country is in the distribution of countries on a specific indicator. Countries at the ends of the distribution have to work harder to change their overall ranking. Rankings can be highly volatile over time, related or unrelated to real changes in the cardinal values of the indicators for the country. Large rerankings can be generated by even very small differences in the underlying measure of interest.[13]

De Jure and De Facto Indicators Can Tell Different Stories

Another important consideration that could lead to the misuse or misinterpretation of an indicator is the difference between its de jure and de facto values. In many real-world scenarios, de jure and de facto differ sharply. One extreme example is the difference between the by-law and in-practice conditions for getting a driver's license in Delhi. The policy by law to get a driver's license is to prove identity, residence, age, and driving competence. But in practice the driver's examination is waived when the applicant hires an agent. More than two-thirds of those who acquired a license through hiring a tout reported that no one had taught them how to drive. In this case, having a license in practice has no predictive power for whether one has the required driving skills by law.[14]

Numerous examples can also be found in the literature. For instance, the difference between de jure and de facto can be clearly illustrated by the difference between the legal ad valorem tariff and the ad valorem tariff actually collected.[15] Not only is the collected ad valorem tariff often much lower than the legal tariff in the same country for the same items, but the collected tariff is also widely heterogeneous and individual-specific, while the legal ad valorem tariff is a single value across the board. Tax exemptions and evasions at the firm level often contribute to the deviations between the de jure tax revenue and the tax revenue collected.[16] Given this disparity, it is not surprising that a change in legal ad valorem tariff will not be associated with an equal and corresponding change in collected ad valorem tariff.

The Doing Business indicators use de jure analysis, primarily measuring laws and regulations as they are written—in other words the formal regulatory environment for firms. But there can be a large difference between what is listed in rules and what occurs on the ground, especially where implementation is a big challenge. The payoff to a

particular regulatory reform will depend on how significant a burden the regulation poses in practice. A recent study illustrates the difference between de facto and de jure approaches using the comparison of some indicators.[17] While Doing Business assesses the country's investment climate with information from lawyers and accountants, Enterprise Surveys ask firms what they perceive as obstacles to their business and about their interactions with governments in policy compliance. The relationship between the two measures is neither one-for-one nor linear.

Policy implementation and governance more broadly influence the gap between the de jure and de facto measures. A reliance on de jure measures may not reflect the actual environment over time and across firms. Changing de jure values does not always result in changes in de facto values, and the impact of such changes on firms varies widely across the distribution.

***

Constructing indexes and scales is seemingly straightforward in principle but difficult to do meaningfully in practice. Indeed, there is a sizable literature on the pros and cons and the pitfalls of developing such indicators.[18] Even with the best efforts of their developers, all indexes and scales have limitations, some obvious, some hidden. They continue to be used, sometimes with scant attention to these limitations—either because they are catchy and useful for particular purposes in specific contexts, or because they serve certain political or other goals.

All indexes are somewhat arbitrary, and that affects their validity and reliability. Composite or disaggregated measures can be useful and various indexes can serve useful purposes. But their developers need to lay out the underlying assumptions and methods used in generating such indexes and users need to take the time to understand them to avoid errors that can have serious consequences for achieving development results.

Notes

1. Ravallion (2010a).
2. Eifert (2007).
3. IEG (2011c).
4. IEG (2008a); European Bank for Reconstruction and Development (2010).
5. Dollar et al. (2006).

6. In Africa and Latin America the informal economy accounts for more than 40 percent of GDP.
7. IEG (2009g).
8. If country X has a governance rating much worse than its ratings on other clusters, and country Y has a governance rating only slightly worse than its ratings on the other clusters, country X will lose and country Y will gain because of the larger weight on governance and the total IDA resource remain the same.
9. Ravallion (2010b).
10. Ibid.
11. Birdsall and Kharas (2010).
12. IEG (2008a).
13. Høyland et al. (2010).
14. Mullainathan et al. (2006).
15. Pritchett and Sethi (1994).
16. Gauthier and Gersovitz (1997).
17. Hallward-Driemeier and Pritchett (2011).
18. DeVellis (2003).

5

Go from Averages to Targeted Segments

Remember: the average is as close to the bottom as it is to the top.
—Unknown

Even if the issues of composite measures are overcome, we still face the problem that our measures are built around averages. They convey important information, but in several contexts averages may not reveal the results we are aiming for. An average can provide benchmark information, but it can also mask the true results if the interest is the distribution. Take a simple example: if two locations share the same average temperature on a yearly basis, their weather could still be drastically different. Imagine that the temperature at one location is stable throughout the year at around 70°F, while the temperature at another location fluctuates from 30°F in the winter to 110°F in the summer, with an average of 70°F.

All too often we come across situations where what we measure on average for a country or state or municipality misses important constituencies and hides disparities among different population groups. The Millennium Development Goals measure progress toward eradicating extreme poverty, but they are average indicators and do not focus on the poor. Achieving them does not mean that the situation of the very poor improves. Higher average income can coexist with reduced wealth or income for the poorest segments of the population in relative (and even absolute) terms, especially if economic growth accelerates with little or no improvement in already existing income disparities—or worse, with a greater concentration of wealth.

The distribution of outcomes matters, not just the average outcome. If only averages are tracked, the impact of an intervention on the targeted segment may remain unknown. Targeting measures often need to be in place to facilitate the intended distribution of benefits.

The development community increasingly recognizes the importance of targeting.[1] To help the poor and vulnerable, many countries have some form of targeted social safety nets. Over the past decade, the World Bank began to move from a project-focused approach that emphasized delivering social assistance, to helping countries build social safety nets and institutions to respond better to poverty, risk, and vulnerability, targeting support to particular groups. Positive experiences in Colombia, Ethiopia, Georgia, and Mexico, for example, can be built on to strengthen initiatives featuring targeting.

But even if projects aim to target the poor and vulnerable, the benefits may fail to reach them if critical elements are overlooked in the link translating the actions into results. For the benefits to reach the poor and vulnerable requires more than focusing on the subgroup with less than average income, addressing their demands, releasing their financial constraints, and changing the structure of empowerment. Understanding the nature and interaction of structural obstacles for a particular category of countries, such as least-developed countries, is crucial to enhancing policy impact.[2]

In the sections below, we use examples in rural electrification, microfinance, and social funds to show how important it is to go beyond averages by keeping in mind the impact on targeted groups and how difficult it is to reach them.

Targeting Poor Areas

A well-known fallacy of poverty targeting is that everyone in rural areas is poor and that all poverty is rural, so focusing on rural areas automatically means reaching the poor. Rural areas have higher poverty rates, but because rural populations are smaller than urban, the number of poor people in urban areas can be higher. Not only would focusing exclusively on rural areas miss the urban poor, but such efforts could also miss the rural poor while disproportionately benefiting the nonpoor.

The global geography of poverty has changed. In the 1990s more than 90 percent of poor people lived in low-income countries; now more than 70 percent of the world's poor live in middle-income countries. Policy needs to be crafted in each specific country and region, taking into account the detailed nature of poverty.[3]

World Bank-supported projects have improved rural electrification, but the distribution of benefits has been neutral or even regressive.[4] Evidence from country case studies shows that electricity subsidies

are invariably less well distributed than a random allocation of funds because electricity consumption favors the better off.

Two factors underpin this pattern on the supply side: which communities get connected and which households can afford the connection once the grid is available. Both work against the poor. First, across villages, to be connected to the grid communities are often identified on a least-cost basis, which favors larger communities nearer the existing grid, roads, and towns. The poor, who are less likely to have access to grids, must pay a higher price for electricity even if they have access to off-grid electricity sources because the unit cost off-grid is higher. Second, high connection charges are a frequent barrier to connecting the poor. Although off-grid connections can serve remote communities that may not be connected to the grid for some years, they do not necessarily reach the poor better than does grid extension.

As a result of these double disadvantages, the access to electricity is still limited for poor households, though there has been substantial improvement. In Bangladesh, in 2004, the poorest 40 percent of rural households accounted for 17 percent of electrified rural households.[5] With wide variations across countries and regions, the expenditure by the poor on electricity is even lower, typically half to two-thirds that of the nonpoor.[6] In the Philippines the bottom 40 percent accounted for around 25 percent of grid connections but only 15 percent of grid electricity consumption.

In most countries increases in coverage within villages come from extensive growth (extending the grid to new communities) rather than intensive growth (connecting the unconnected in already electrified villages). In the Lao People's Democratic Republic, even in villages connected for 15–20 years, it is not uncommon for 20–25 percent of households to remain unconnected. An estimated 30 percent of its population cannot afford the $100 connection charge.[7]

It is worth noting that a price scheme featuring a lower unit price with a higher connection fee (versus a higher unit price with a lower connection fee) further burdens the poor. The unintended result is that poor people bear the brunt of high electricity prices and cannot reap the full benefits of electrification. The consequences are even worse when consumer knowledge is limited.

In Tambo, South Africa, consumers can choose between two price schemes: paying a connection fee of 200 rand and a metered charge per kilowatt hour, or a lower connection fee of 10 rand and a fixed monthly

charge of 15 rand. Given their actual consumption, most households would be better off taking the first option. But many households, particularly the poor, opt for the second because they cannot afford the high connection charge and are unsure how much they would use. To make things worse, many low-income consumers cannot always afford the high monthly 15 rand charge and so are disconnected and have to pay the 10 rand repeatedly to be reconnected.[8]

Another price scheme—the lifeline tariff—aims to help the poor by keeping the fixed charge low for the initial limited amount of use and imposing a much higher rate after a threshold is passed. In practice, this can also become an obstacle for the poor. The objective is to encourage the poor to get connected to electricity. But the unintended result is to restrict their consumption.

Although electricity can be used for production, the benefits of which are higher than the true willingness to pay, demand for consumption is unintentionally reduced by the illusory effect of a sharp jump in the unit price after the low threshold is reached. This results in false savings that not only limit the real benefits electricity can bring to the poor, but also unintentionally lower the financial viability of providing the electricity—because of low load factors resulting from consumption heavily concentrated in evening peak hours.

Helping the poor overcome the financial barrier is important if they are to reap the benefits of access to electricity. One way to ease the burden of a large, one-time, upfront payment is to allow households to spread payments, either by adjusting the tariff to an installment basis or by providing credit for this purpose. In Morocco, during the Second Rural Electrification Project, rural consumers were allowed to pay the connection charge in monthly installments over seven years. In Thailand, a credit program with loans made available to households in villages with grid connections was piloted in the Electricity Access Rural Expansion. The connection charge can be repaid over a period of up to twenty years.

Alternatively, profit-maximizing price discrimination schemes can apply differentiated prices for different types of consumers—charging a higher fee for those who can afford it (the more affluent) and a lower fee for those with less means and a higher elasticity of demand (the poor). Given that electrification rates tend to increase sharply in the initial years and then increase more slowly after reaching a certain threshold, timing provides a good signal of the differentiated willingness to pay of

the affluent and the poor. Charging a differentiated connection tariff across time (lowering the connection fee a certain number of years after a village is connected) would allow the poor to benefit from a more affordable connection fee and suppliers to maximize their profits and increase their financial viability.

Several World Bank projects have started taking these issues on board. The Cambodia Rural Electrification and Transmission Project created a Rural Electrification Fund, an explicit cross-subsidy to finance rural electrification. The Thailand Second Rural Electrification and the Ethiopia Accelerated Access Rural Expansion provided credit for rural consumers to cover the connection fee.

Targeting Issues Relevant to the Poor

Another well-known problem of targeting is that it assumes that focusing on an issue more common among the poor will benefit the poor disproportionately. Not so. For example, the health benefits of water supply and sanitation interventions do not disproportionately benefit the poor—some may not even reach them.

In a case study of India, expanding piped water had no impact on the prevalence and duration of diarrhea in the lowest two income quintiles.[9] There were health gains among the lowest quintiles only if a woman in the household had more than a primary education. In Lesotho, a 24-percent average reduction in diarrhea was associated with owning a ventilated improved pit latrine, but the largest declines were among households with better hygiene and water use behavior and higher socioeconomic status, not the poor households.[10] In Malaysia, improvements in water quality were associated with lower infant mortality only among the literate.[11]

Malnutrition is another case—it is more common among the poor, but this does not mean that children living in poverty will benefit disproportionately from an intervention. In Ethiopia, free distribution of food raised the weight-for-height z-score of children under-five in high-asset households but not in low-asset households.[12] In Madagascar, a large-scale community-based nutrition program tended to benefit the nutritional status of children in better-off communities, even though it was targeted to the poorest areas.[13]

Overall, there is scant evidence of who is benefiting and who is not and of the cost-effectiveness of interventions.[14] Fewer than half of forty-six evaluations of nutrition's impact measured the distribution of impacts by gender, mother's education, poverty status, or the

availability of complementary health services. And only nine assessed the impact on nutritional outcomes for the poor compared with the nonpoor. Among the evaluations that did examine variations in results, several found that the children of better-educated mothers or children in better-off communities were benefiting the most. Many other factors—such as proximity to a road, a hospital, or electricity, and access to safe water—can have a large impact on the distribution of the benefits across and within communities.

The average results for a population can also be an obscured measure of the effectiveness of an intervention, because an intervention can be highly effective for some segments of the population but ineffective for others.

An intervention's impact on individuals could vary by their specific characteristics. For example, with nutrition interventions, the age of children can be partly responsible for the variability of nutrition provision. Younger children are more susceptible to nutrition shocks than older children. Some interventions that may improve the nutrition of infants may not have the same effect on older children. When the results are measured for children in a fairly large age range, say under five years old, without information on the subgroups, the real effects of the intervention could be misestimated. Chances are good that an intervention effective in addressing a nutrition issue of a specific age group will go unnoticed because the average of the wrong population was measured.

During and after the East Asian financial crisis, a supplementary feeding program ran from 1998 through 2001 in Indonesia, giving young children special high-nutrition snacks. The program had a positive effect on children aged 12–24 months, reducing stunting by 15 percentage points, but it did not have a similar impact on infants aged 6–12 months or on children aged 24–60 months.[15] The difference in results of nutrition programs across children in different age groups is also observed in Uganda[16] and Zimbabwe.[17]

Targeting Constraints Faced by the Poor

A less well-known targeting problem is to assume that addressing the more severe constraints for the targeted group will benefit them disproportionately. In reality, although the targeted population is drastically in need of the specific resources, these may still leak to the nontargeted group, which may be better able to tap into them or make better use of them.

Microfinance is designed to provide a range of financial services to poor people to help them increase their incomes, build their assets, and reduce their vulnerability. By using innovative solutions to reduce transaction costs and substitute for conventional collateral, it can provide poor households loans in low dollar amounts without charging high interest rates. By improving access to finance, it aims to address people's financial constraints and facilitate the management of money, for both production and investment.

Microfinance often targets poor females, but it may be used by male household members and the less poor. Several studies and impact evaluations show that poor women did not benefit as much as programs had intended. While microfinance raised the incomes of the general population, it might not have empowered women (thus benefiting children) or reduced poverty.[18]

Two assumptions underpin the view that the targeted groups will be the actual beneficiaries. First, the financial return of increased resources through the supply of microcredit to poor female entrepreneurs equals that for the general population. Second, an increase in supply of microcredit to poor female entrepreneurs will be met by an equal increase in their demand for it. These two assumptions may not be valid in some circumstances.

In Sri Lanka, the returns to capital in female-run microenterprises are lower than those run by males.[19] Women invested grants differently from men and a smaller share of the grants remained in female-owned enterprises. Men were more likely to spend the grant on working capital and women on equipment. In this case, although microfinance could improve access to credit, due to the differences in preference for investment, female entrepreneurs may not benefit as much as their male counterparts given their lower rates of return to capital.

The returns to capital typically vary across individuals for multiple reasons. Due to gender differences in education or business networks, women might be less informed about investment opportunities. In a Peruvian group lending program for female microentrepreneurs, the limited demand for credit—rather than the supply of it—was the main constraint to female entrepreneurial success.[20] The lack of high-return means to expand their businesses implies that improving access to credit through microfinance might not disproportionately benefit females.

One reason could be the design for minimizing risk-taking through microcredit. Compared with traditional loans, microcredit has three

distinct features: repayment is required to start immediately, there is little flexibility about time of repayment, and there is zero tolerance of default. Useful as they are in minimizing default risk and keeping microcredit profitable and manageable for lenders, especially given that the loans are not backed by collateral, the narrow time frames and limited flexibility on repayment may constrain borrowers from investing in longer-term projects.[21]

The availability of money is not the whole story. The types of loans also matter for impact. In particular, poor women may naturally tend to be more risk-averse and more likely to be affected by the loan design than men. So, although microcredit targets women, the reduced returns to capital due to its rigid structure (limited or no grace period, weekly payments) could push female demand for more microcredit below that of men.

Changing the structure of microcredit might help. A microfinance program in Kolkata, India, instead of requiring clients to make weekly repayments immediately after loan disbursement, afforded them a two-month grace period and made the repayment schedule monthly.[22] Women used the loans for longer-term projects (such as selling saris) where revenue is less time-constrained. While the shift to a grace period increased their rate of default, it also increased clients' business investments in the short run and profits and income in the long run. In Uganda, the repayment behavior of a borrower may be partly driven by simple product details, such as the ease with which the borrower can pay the loan.[23] In this case, changing loan programs to ones that facilitate easy repayment or frequent reminders may improve loan repayment behavior and reduce the cost of lending.

Microfinance increases access to money, but that does not guarantee that people will invest rather than consume. Whereas poor people will benefit from a smoothing of consumption to buffer external shocks, in general, borrowing for consumption will not increase income in the long run.

Opening a microfinance institution in Hyderabad, India, had no impact on measures of health, education, or women's decision making.[24] In the short term, households with an existing business at the time of the program invested more in durable goods, while their nondurable consumption did not change. Households with high propensity to become new business owners increased their durable goods spending and saw a decrease in nondurable consumption to pay the fixed costs to enter entrepreneurship. But households with a low

propensity to become business owners increased their nondurable spending.

The channels for microcredit to work could also differ from what is expected. Microcredit, rather than working directly through the targeted businesses, might work more broadly through risk management.[25] There is no evidence that increased access to credit in Manila, Philippines, improved subjective well-being. Male and higher-income entrepreneurs benefited more from microcredit than women operating small-scale businesses, who were targeted. Business investments did not increase. Instead, the size and scope of treated businesses shrank by shedding unproductive workers, similar to how increased access to credit reduces the need for trading favors within family or community networks.

Changing Aid's Structure to Empower the Poor

Another less well-known but common fallacy of targeting is that in reaching the poorest communities with interventions focused on the issues they face, only poor households will benefit. When looking at inequality at highly disaggregated units, within-community inequality is often higher than between-community inequality. In Ecuador, Madagascar, and Mozambique, the median inequality in the poorest quintile of communities is no lower than in any of the richer quintiles.[26] So, targeting the poorest communities does not always lead to the expected results of disproportionately benefiting the poor.

The objectives of social funds and community-based/community-driven development typically include reaching the poor, vulnerable, and underserved. But an evaluation[27] of social funds shows that having the good intention of targeting the poor does not always translate into pro-poor results. There could be a difference between the intended and actual beneficiaries related to the design and implementation of targeting. Social funds establish menus, procedures, and targeting criteria to support investments benefiting the poor.[28] In most cases poor people are the explicit target group—80 percent of social funds aimed at targeting the poor, 46 percent of the poorest, 44 percent the vulnerable, and 10 percent the low income.

Social funds and community development projects often focus on increasing access to service delivery infrastructure, such as schools and health centers, for remote communities. But as with other projects, greater access to infrastructure does not always translate into effective service delivery. Nor do the poorest always benefit. The

better-off households may have better access to these services and may be better able to afford them. There is little hard evidence on the poverty-reducing and community capacity-enhancing impact of the projects.

Timing, composition, and geographic distribution of expenditures in democratic regimes are often related to electoral interests.[29] The decisions about social fund financing were made on the basis of both political and technical criteria. In a case study in Peru, marginal voters and core supporters both received a disproportionate share of social fund expenditures.[30]

Some studies find that social funds generally reach the poor,[31] while others find that better-off regions and provinces often receive more financing per capita.[32] A worldwide study on poverty-targeted social programs found that while the median program transferred 25 percent more to poor individuals than universal allocations would, a quarter of these targeted programs were regressive, with benefits tilted toward the nonpoor.[33]

One distinguishing characteristic of social funds and community development projects is that rather than implementing investment decisions predetermined at project appraisal, they allow local stakeholders to determine these decisions through subproject proposals they submit during project implementation. So, for the poor to benefit more, they must participate and get their voices heard.

Having a participatory approach built in does not always lead to the participatory outcomes expected. Even if the poor are formally included in a participatory space due to project requirements, their views and priorities can still be excluded from collective decision-making processes.[34] In the Matrouh Project in Egypt, where there had been a substantial focus on women, the percentage of women who believed that they had benefited from the project was highly variable. There were no elected women leaders or women's associations, and there were substantial concerns about marketing products women produced. The difference in the impact of social funds in Malawi and Zambia is a good example illustrating the importance of true participation.[35] Looking at how social funds have operated at the village level casts light on why social fund projects have limited impact in building social capital despite their participatory model—because a degree of social organization is required to apply for social fund resources.

Holding public meetings is not sufficient for the community to participate actively in decision making. In Zambia, headmen often

make the decisions; there is little room for dissent at public meetings as the meeting takes place only after considerable work has been done and the headmen have sought the backing of the chief. As a result, the majority of the community participates actively in making bricks but more passively in making decisions. In Malawi, traditional leaders often mobilize the parent–teacher association, giving the community more influence over decision making.[36]

As a targeting tool, social funds have defined menus, which feature basic services the poor are more likely to demand. Whether a community development project benefits the poor more than the nonpoor is closely related to the extent that the choice of project is consistent with the preference of the poor.

Within a community, there can be a wide difference in preferences between people of different income levels and between men and women. In Indonesia's Kecamatan Development Program, the poorest households favored irrigation and health more than the better off, who favored roads much more highly. Women preferred health, education, and drinking water projects more than men did.[37]

Although the types of projects chosen should be conditioned by other factors, such as technical feasibility and financial constraints, whether there is elite capture and whether the process is participatory can make a critical difference. The better-off, better networked individuals often dominate the participation process and are more likely to have their priority needs satisfied.[38]

In Jamaica, Malawi, Nicaragua, and Zambia, the subproject selection process could not be counted on to meet the highest priority problems of the majority of community members.[39] Even where virtually the whole community participated in some aspects of the subproject, the community as a whole did not necessarily drive subproject choice. As "prime movers," community leaders were critical to mobilizing support for and preparing a successful subproject proposal and their interests were determined by their position. For example, if the prime mover is a headmaster or health worker, there is often a bias toward subproject investments in schools or in health facilities. Although it is natural and appropriate that prime movers bring project ideas to the community, it is important to ensure that the ideas of the leaders align with the preferences of the community, including those of poor households.

But elites often disproportionately influence community decision making. In Ecuador's Social Investment Funds, communities with

higher levels of estimated consumption inequality were less likely to have projects that provided excludable goods, such as latrines, for the poor.[40] The community decisions on whether to apply for a public-good or a private-good project hinged on three factors: the way individual preferences vary with incomes, the distribution of income, and the relationship between political influence and income. In this case, the income share of the elite was the strongest determinant of project choice.

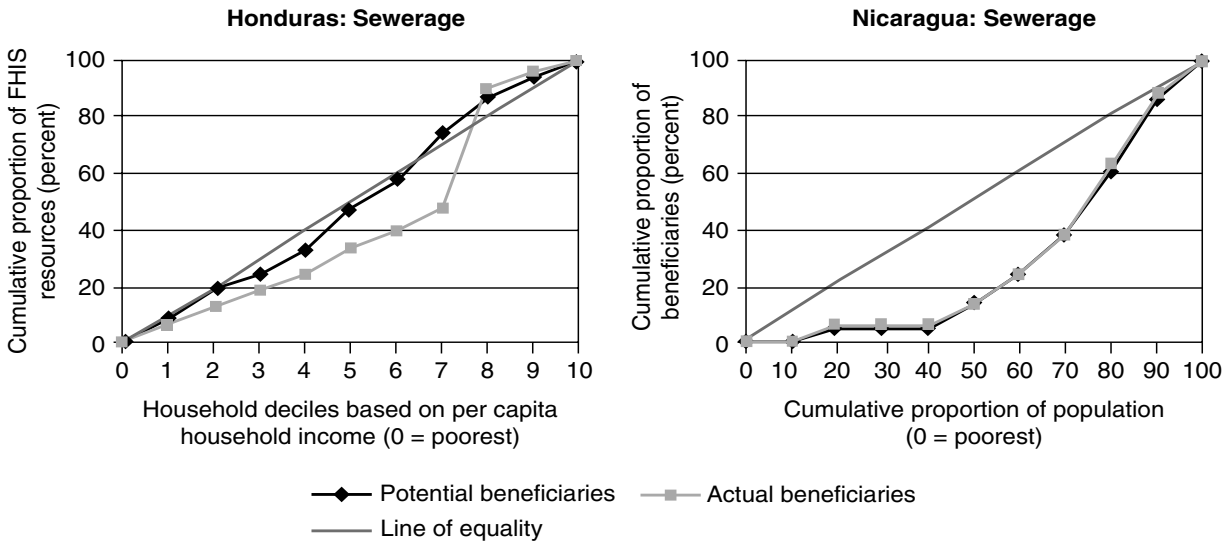
Potential targeted beneficiaries can differ from the actual ones, especially if the poor are left out of decision-making processes. For example, sewerage projects, where the incidence of expenditure is regressive, often benefit wealthier households than poorer ones (Figure 5.1). In Honduras, the high costs of connecting to the system were prohibitive for many poor households. Because they are not connected to the sewerage line, they are not actual beneficiaries. In Nicaragua, the sewerage systems tended to serve the better-off neighborhoods, with the poor households in the less-affluent areas less likely to have access.

There are even cases where the position of the poor actually worsened.[41] In Benin's Borgou pilot project, the community contribution typically required in World Bank interventions created hardships for the poor. It is very difficult for the poorest to make a cash contribution, so they usually have to contribute time and labor, which takes them away from income-earning activities. And where the rich contribute on behalf of the community, the position of the elite is strengthened relative to that of the poor. Moreover, reaching the poorest requires fundamental social and cultural changes, which take considerable time and sustained effort, unusual in a Bank-supported project.

When looking at the outcome of a program, it is often assumed that the difference between the treatment group and the control group can be attributed to participation in the programs. But a selection bias can question this assumption. For example, it is often concluded that a social fund program has achieved its goal simply by observing a difference in social capital between treatment villages and control groups. But communities with high levels of social capital are more likely to apply for social funds for community development programs, so the difference ex post could reflect the difference ex ante rather than indicate the impact of the program.

For community-driven development projects, improving transparency in decision making and lowering barriers to participation

Figure 5.1
Sewerage Projects in Honduras and Nicaragua Benefit the Better Off Disproportionately



Source: Rawlings et al. (2004)

can help. But there is still scant evidence of what can improve the pro-poor benefits of such projects. Helpful steps include requiring a minimum attendance at critical decision meetings, regulating procedures for forming community project committees, and improving transparency. For poor communities that still have difficulty accessing resources, additional measures may be needed to improve equality in results, perhaps by providing early assistance for community organization, waiving requirements for counterpart funding, or extending technical assistance.

***

The impact of an intervention is context-specific: what works in one setting may not work as well in another. The effect of an intervention can vary widely across segments of the population, and different population groups may react differently to a given intervention. Generalizing the findings and assuming a uniform distribution of benefits could be misleading. Assuming the results of a project scaled up to a larger population to be a multiple of those in a pilot applied to a specific group, even though the interventions are a replication, could be problematic.

For targeting to be effective, a clear understanding of how the underlying conditions bridge actions to results and how the variations in such conditions may affect the results of targeted group is key. When the critical conditions for an intervention to work are missing for the targeted group, they may remain left out even if averages improve.

A recent evaluation of social safety nets indicates that the World Bank's results frameworks improved throughout the 2000s, but further improvements are necessary. (Similar observations can be made of other multilateral development banks.) Results frameworks have not focused enough on the poor and vulnerable.[42] Objectives and performance indicators were often not specific enough to ensure effective monitoring of the effects of social safety nets on the poor or vulnerable. Only 59 percent of operations supporting social safety nets had objectives that specifically targeted the poor and vulnerable, and 47 percent of operations supporting them did not have even one indicator to monitor progress on reaching the poor. When the poverty focus was mentioned, it was often in general terms of poverty reduction rather than part of a time-bound objective.

Notes

1. IEG (2011b); Asian Development Bank (2011).
2. Guillaumont (2009).
3. Kanbur and Sumner (2011).
4. IEG (2008d).
5. Ibid.
6. Ibid.
7. Ibid.
8. Ibid.
9. Jalan and Ravallion (2003).
10. Daniels et al. (1990).
11. Esrey and Habicht (1988).
12. Quisumbing (2003).
13. Galasso and Umapathi (2009).
14. IEG (2010k).
15. Giles and Satriawan (2010).
16. Alderman (2007).
17. Hoddinott and Kinsey (2001).
18. De Mel et al. (2008); Karlan and Valdivia (2008).
19. De Mel (2008).
20. Karlan and Valdivia (2008).
21. Banerjee (2010).
22. Field et al. (2010).
23. Cadena and Schoar (2011).
24. Banerjee et al. (2010).
25. Karlan and Zinman (2009).
26. Elbers et al. (2004).
27. IEG (2002).
28. Jorgensen and Van Domelen (1999).
29. Rogoff (1990); Schuknecht (1996).
30. Schady (2002).
31. Rawlings et al. (2004).
32. Jorgensen et al. (1991); World Bank (1997).
33. Coady (2004).
34. Kumar and Corbridge (2002); Turton and Farrington (1998).
35. IEG (2008b).
36. In Zambia, only one-third of people knew of the meeting held for the subproject selection, one-fourth attended the meeting, and one-seventh spoke at the meeting; in Malawi, almost four-fifths knew about the meeting, three-fifths attended, and one-sixth spoke.
37. Olken (2006).
38. Rao and Ibanez (2005).
39. IEG (2002).
40. Araujo et al. (2005).
41. IEG (2005).
42. IEG (2011b).

6

Align Intermediate and Final Goals

Government efforts are still focused on showing increases in enrollment rather than improvements in learning.
—Abdul Latief Jameel (Poverty Action Lab)

The frequent discontinuity between intermediate outcomes and final outcomes has implications for whether development efforts succeed. Aligning measures of intermediate and final goals is crucial to get a picture of development results. Unlike the earlier discussion of the urgent versus the important, we consider here how efforts to meet intermediate objectives may not lead to achieving final goals. Many promising intermediate outcomes are only necessary but not sufficient conditions for the achievement of goals.

Money can buy inputs and outputs, but to achieve development outcomes, policy must induce behavioral change. Controlling and measuring the inputs and immediate outputs of a program—for instance, how much money is spent and how many textbooks are distributed in schools—is important but not enough to achieve the desired outcome. Sometimes an exclusive focus on intermediate steps, such as a rapid increase in enrollment rates, can come at the expense of the desired results, such as improving learning outcomes, if the former crowds out the needed resources for other steps.

Projects often need to target intermediate outcomes to make things manageable, and they often (rightly) do so, keeping in mind frameworks of assumed links among inputs, outputs, outcomes, and impacts. Often the intermediate outcomes, even if imperfect, are one set of useful proxies for where things are headed. They are also easier to monitor, reflecting changes in a timely manner. But promising intermediate outcomes may not lead to the desired final outcomes if critical links in the causal chain are missing. And without the desired

results spelled out and monitored, program achievements will remain unknown.

These lessons are demonstrated here for education and health, though similar lessons could also be drawn from other areas. The examples tell a similar story in both sectors: a focus on intermediate objectives does not necessarily ensure the achievement of ultimate goals; but the way we measure performance often overlooks this disconnect, undermining success in delivering results.

All this does not mean that each step needs to focus on the subsequent links, but it is important to see the connections. One possible solution is to develop mechanisms that can overcome the disconnects and link inputs and ultimate goals. The chapter ends by discussing output-based and outcome-based aid, their potential, constraints, and pitfalls.

Education: From Access to Learning

There has been promising progress in education.[1] Worldwide, about 85 percent of primary school-age children are enrolled in school.[2] The number of primary school-age children out of school fell from 106 million in 1999 to 68 million in 2008. Even in the poorest countries, average primary enrollment rates surged above 80 percent, and completion rates, above 60 percent. And between 1991 and 2007, the ratio of girls to boys in primary and secondary education in the developing world improved from 84 to 96 percent, with even larger gains in the Middle East and North Africa and in South Asia.

Gains in access to school have turned attention to improving the quality of education and accelerating learning. The overarching goal is not just schooling, but learning, better labor force outcomes, and earnings. Growth, development, and poverty reduction depend on the knowledge and skills that people acquire, not the number of years that students sit in classrooms.

The goal of education is building human capital, enabling people to realize their potential and contribute to development. The driver of development will ultimately be what individuals learn, both in and out of school. Recent research shows that the skills in a workforce predict economic growth rates far better than do average schooling levels. It is estimated that an increase of one standard deviation in student reading and math scores (roughly equivalent to improving a county's performance ranking from the median to the top 15 percent)

is associated with a (large) increase of 2 percentage points in annual GDP per capita growth.[3]

In the past ten years the World Bank's support to postprimary education has expanded, but the share of projects with labor market objectives has not.[4] Education projects most often aimed to improve the quality of education inputs, increase access to schooling, and improve the equity and efficiency of education. Bank support has succeeded most at increasing access to education and improving its equity, whereas fewer than half of projects have achieved education quality or labor force, management, learning, or efficiency objectives.

Various factors enhance learning.[5] School hardware (school buildings, classrooms, and sanitation facilities) and school software (curriculums, pedagogy, textbooks, and writing materials) need to attain a minimum threshold to facilitate learning. The availability and quality of teachers have a big role. School management and institutional structures, such as school autonomy and examination systems, are associated with learning outcomes. Contextual factors—including individual student health and nutrition status, family background, and community characteristics—also influence students' learning.

Many primary completers in low-income countries still have difficulty with reading and basic calculations. If school completion rates are raised by automatically promoting children to the next grade and not by attending to student learning outcomes, higher completion rates will not reflect improved knowledge and skills. Although getting children in school has its own benefit, mechanically pursuing an increase in the number of students adds little to achieving the goal of learning.

Tanzania rightly won a United Nations award in September 2010 for its progress toward attaining universal primary education five years ahead of the 2015 target set under the Millennium Development Goals.[6] But the results of a recent nongovernmental organization-led review are sobering—the learning outcomes of the children remained low. The survey results suggested that about 20 percent of the children who completed seven years of primary school could not read their own language, Kiswahili, at the grade 2 level; half could not read English, the medium of instruction in secondary education; and about 30 percent could not solve grade 2 multiplication problems.[7]

The World Bank supported the Tanzania Primary Education Development Program (2001–2004) and the Secondary Education Development Program (2004–2008) to improve access to education and

learning outcomes. School fees were eliminated for primary education, tuition fees were reduced by half in government schools, and a capitation grant was provided for books and learning materials. As a result, intermediate outcomes were promising—primary enrollment rates increased dramatically, the number of government-supported secondary schools rose almost fourfold, and secondary enrollments nearly tripled.

But the promising expansion of enrollment did not translate into the desired final outcomes of improving learning. Rapid expansion was constrained by a shortage of qualified teachers. Learning outcomes in secondary education suffered.[8] Although the number of secondary graduates increased, the pass rate to continue study declined substantially, and a large gender gap in outcomes persisted.

Even substantial increases in domestic budget and development partner support were not enough to finance all the goals of primary and secondary education development. The capitation grants for nonsalary items, such as books and learning materials, were cut by a third and the pupil–teacher ratios doubled, with negative effects on learning.

Tanzania is not an exception; many other countries share similar lessons. Despite the large investments in and high coverage of minimum operational standards for schools, learning achievements fluctuated. Empirical results in Brazil's Bahia State and St. Lucia indicate that, on average, improving the quality of inputs was not associated with better learning outcomes.[9]

Teacher absenteeism is often associated with poor student learning. But higher teacher attendance cannot be presumed to lead inevitably to more learning, when other constraints in the supply side and demand side are binding. In Kenya, ICS, a Dutch nongovernmental organization, provided schools with funding to hire a local contract teacher to address classroom overcrowding. While hiring an extra teacher on a short-term contract had a generally positive effect on test scores, the impact depended heavily on how the program was implemented. Training school committees to monitor teachers in conjunction with hiring contract teachers were found to increase program effectiveness.[10]

Seva Mandir, an Indian nongovernmental organization, placed a second teacher in the nonformal education centers it runs in Indian villages. In half the forty-two centers randomly selected to receive a second teacher, teacher attendance increased, as hoped, but test scores remained the same.[11] Further enquiry of the program indicated

that poor schooling quality and economic deprivation could be major factors that forced many children to undertake work and lose out on studying.[12]

Access to technology can help students learn. But having computers in schools cannot be presumed to lead inevitably to better learning. The project "Computers for Education in Colombia" places computers in public schools but there is no evidence that doing so produced measurable increase in student test scores.[13] Part of the reason may be that, although computers were available, teachers did not use them very often, and when teachers did use the computer, it generally was not as a teaching aid or for other classroom activities. The key contribution to learning—incorporating the technology into teaching—was missing.

The general lesson: focusing only on intermediate outcomes through improving inputs risks not achieving the desired results. These lessons can be discerned from the experiences of multilateral development banks. Overall, two-thirds of World Bank primary education projects during 1990–2005 focused on increasing enrollment and reducing dropout rates.[14] A fifth of them had objectives explicitly covering the expected results, such as improving reading, writing, and math skills, and other learning outcomes. Further, there was no clear increase in explicit objectives for learning outcomes, though projects have increasingly financed development or implementation of learning assessments over the last decade, as indicated in a recent portfolio note.[15] In all education sector projects approved in fiscal 2001–2009, only one in five had an objective related to the labor market, such as generating employment, creating human capital, or increasing the market relevance of education. Fewer than half achieved their objectives for education quality, labor force, management, or efficiency.

The cost of such inaction is high. The results for about a third of the projects aimed at learning outcomes are unknown because of inadequate or nonexistent learning outcome indicators. Country case studies in Bulgaria, Georgia, Mali, and Serbia show that even if results for learning outcomes are an explicit objective of the projects, no information is available for projects.[16] Most of the projects that supported new learning assessments could not apply them more than once during the project, making it difficult to track trends.

Improved student enrollment and retention, along with enhanced teacher attendance and improved technology, can help in achieving better learning outcomes. But measures are often missing to bridge

the targeted output of enrollment and completion with the desired outcomes of improved learning. The absence of planning for improved learning outcomes and the lack of political commitment for reaching those outcomes can result in unnecessary tradeoffs between improved access and student learning gains, especially among the poor.

Tracking intermediate as well as final outcomes are important for understanding what works in bridging school access to learning and what conditions influence outcomes. Too often, inadequate results frameworks lead to the adoption of inadequate indicators. For example, employment rates of graduates do not reveal the type of occupation, sector of employment, earnings, opportunities, or duration and stability of employment, which could still be partial to measure education outcomes. Without appropriate monitoring and baseline information, the evidence is inconclusive on whether investments in vocational and higher training contribute to meeting labor market demands.

The World Bank's new education strategy aims at supporting countries rightly calls for stronger systems to improve the quality and reach of education. It calls for prioritizing and financing reforms of countries' education systems to improve the quality of student learning, matching new education financing with results, and building an evidence base of what works in education reform and what does not. This is an important development.

Health: Reaching Mothers and Children Is Not Straightforward

The value of strengthening links between improving inputs and achieving outcomes is also evident in the health sector. Building health clinics and increasing the medical staff and hospital bed to population ratios are only steps to strengthen health service delivery. Besides the lack of access to good health care services, poor nutrition practices can be a major contributor to poor health outcomes among other actions from the household side, such as inadequate care for illnesses and careless handling of water and waste. The same level of health inputs can translate to different health outcomes, depending on elements from the supply and demand sides.

Take children's nutrition. High levels of child malnutrition in developing countries contribute to mortality and have long-term consequences for children's cognitive development and earnings in adulthood. Educating caregivers is often among the main interventions to improve the nutrition status of children. But they are not sufficient

for improving nutrition outcomes.[17] They are at best necessary steps toward the desired goal.

The results chain from interventions to nutrition outcomes is long and complex. Many factors need to be in place to translate to these actions into children's better nutrition. Public policy can have an impact through government finance and regulation of many types of services, but its impact on nutrition outcomes depends on the local context and on the behavior of service providers and households. Having the nutrients available and the knowledge or willingness to change behavior is no guarantee that real changes will take place. Only when behavioral changes occur—along with other necessary conditions, such as access to food, hygiene, and preventive care—can nutrition outcomes be achieved.

In Ethiopia, while the Health Service Extension Programme seems to result in a larger percentage of women making their first contact with a skilled health service provider significantly earlier during pregnancy, very little effect is detected on other prenatal and postnatal care services. Moreover, the program has not reduced the incidence and duration of diarrhea and cough diseases among children under age five.[18]

Community nutrition programs reduced stunting in Madagascar and Haiti but not in Bangladesh. The Bangladesh Integrated Nutrition Project aimed to educate pregnant women and new mothers in nutrition to improve maternal and child health. The project was implemented correctly as planned with mothers attending the programs and gaining knowledge of children nutrition. But changes in mothers' knowledge did not reduce child malnutrition.[19]

One reason could be that in Bangladesh, husbands do the shopping and mothers-in-law have a stronger say in children's food consumption. The assumed causal pathway linking mother's improved knowledge of nutrition to behavioral changes to feed babies with more nutritious food so to achieve the final outcomes of improving children's nutrition status was mis-specified. The intervention resulted in very little reduction in malnutrition, even though 90 percent of mothers attended the program. They could not translate the knowledge gained to changes in practice that would improve their children's nutrition outcomes because they are not the main decision maker for their children's nutritional intake.

Assuming that demand for the service provided by the intervention always exists and that the targeted population shares the same

objectives of the project designers is not always valid. Addressing such constraints to align the objective of project intervention with the target population can be required.

The "lentils for vaccine" program is a good illustration.[20] In some places in India, where vaccinations are free in public health facilities, immunization rates have seen some increases but remain low. The lack of understanding of the benefits of vaccination or even a suspicion of government health services could be contributing factors in some contexts, in addition to the high absenteeism of health care providers and unreliable supplies of vaccines. The combination of a well-staffed mobile immunization camp and small nonmonetary incentives (such as a one-kilogram bag of lentils per immunization, plus a set of plates for completing an entire immunization schedule) leads to more children receiving vaccines and a greater likelihood of full immunization.[21]

A given intervention should not be expected to yield the same results, even if it could be implemented in exactly the same way in different places. In the long pathways connecting public policy to outcomes, service providers and households play a role in determining health outcomes. It would be problematic to take for granted that the same results as those in randomized trials in a controlled setting would come from interventions in various local contexts. The relevance and impact of health impact evaluations could be enhanced by collecting data on service delivery, demand-side behavioral outcomes, and implementation processes to better understand the causal chain and its weak part.

Output-Based Aid and Outcome-Based Aid Imply Different Incentives

The model of aid delivery and the appropriate stage of result targeting are current topics of debate. Output-based aid pays for specific outputs. Outcome-based aid provides funds only on condition that the recipient achieves certain outcomes. By focusing on different stages of the results chain, output-based aid and outcome-based aid have different incentives built in for their reward structure. Importance is to strike the rebalance of focus given the difference in causality from output to intermediate outcome and final desired outcome in a specific context. A good base of aid delivery needs to be measurable (so it can change midcourse if needed) and have a causal relationship with the desired results.

Both output-based aid and outcome-based aid have their advantages and disadvantages. Output-based (or intermediate outcome-based) aid is easier to measure and track progress. It is often desirable when there is strong causality between the output (or intermediate outcome) and the ultimate result. But the output (or at best, the intermediate outcome) indicators may not provide accurate indications of long-term beneficial changes when the causality is weak.

Outcome-based aid can provide incentives to pursue final results, fostering accountability between funders and their citizens, between recipient governments and their citizens, and between funders and recipients. But it is often challenging when the desired outcomes are intangible or difficult to measure and when attribution is an issue. If it takes too long for the desired outcomes to emerge, not having appropriate tracking of intermediate outcomes can risk not knowing the progress of the interventions, so adjustments or scale-ups cannot be made at critical moments.

Evidence from impact evaluation on the long-term impact is scant, but some offers useful lessons. When the links between intermediate outcome and final desired outcomes are strong, targeting intermediate outcome using conditional cash transfer, a traditional output-based aid, is effective. For example, in Pakistan the Punjab Female School Stipend Program, a conditional cash transfer program targeting girls, helped narrow gender gaps in education. Evaluative findings suggest that four years into program implementation, adolescent girls in districts with stipends were more likely to progress through and complete middle school than those who do not receive stipends. Girls exposed to the program later on—and who are eligible for the benefits given in high school—also increased their rates of matriculation into and completion of high school, contributing to human capital development.[22] Evidence from impact evaluation suggests that financial incentives help to increase the use of key health services on which the cash transfer is conditioned, if the beneficiaries know about this condition. However, if the value of the services is low or unknown, results are mixed with respect to nutrition and health outcomes.[23]

But the links in service delivery between outputs and intermediate outcomes and final results are not always straightforward. Targeting the intermediate outcomes too early may not ensure the achievement of final desired results due to potentially weak links if the two do not align. In response to the rising demand to promote results-based programs, some outcome-based aid models emerged.

The program-for-results assistance developed by the World Bank Group and the cash-on-delivery aid proposed by the Center for Global Development are two new initiatives whose effectiveness are yet to be tested. Both program-for-results and cash-on-delivery aid focus on outcome targets, but at different levels.

Program-for-results focuses support directly on improving sectoral or other development programs and finances specific expenditure programs. Disbursements are made based on achievement of results and performance indicators, and hence determined by progress on monitorable performance indicators rather than simply by whether expenditures were incurred.

For program-for-results, the disbursement-linked indicators can be at different levels. For example, they can include desired outputs or outcomes (such as the number of service connections of a particular specification or the number of children vaccinated or confirmed as going to school). They can also include intermediate outputs or process indicators (such as confirmation that specified communities participated substantively in decision making). And they can include financing indicators (such as the share of a specific type of project in total expenditures). The challenges are to strike the right balance between the monitored output and intermediate outcomes to adjust mid-course as needed—and to target the real desired results in the long run.

Cash-on-delivery aid proposes an even more substantial and fundamental change in targeting final outcomes. At no point does the funder specify or monitor input, and the funder makes payments for outcomes, not inputs. Rather than focus on disbursements and verifying expenditures, it links payments more directly to a single specific outcome. Its purpose is to overcome the problems in many current aid programs by strengthening the donor–recipient relationship to achieve results.[24]

Compared with traditional approaches to aid, the cash-on-delivery aid funder embraces a hands-off approach. The recipient has complete discretion and responsibility from the initial design and planning right through to implementing strategies.

For outcome-based aid, disbursements are contingent on the achievement of specific outcomes. Funding is disbursed only after the mutually agreed results are achieved. It could reduce the risk of ambiguity or broken links among outputs, intermediate outcomes, and final outcomes. Valid and reliable measures of those outcomes are critical

for the aid to succeed. But inappropriate outcome measures could lead to unintended consequences, and domestic resources might be diverted from the needed sectors, as mid-course adjustment is more challenging due to the limited progress-tracking mechanism.

In pursuing these instruments it is vital to ensure that there are adequate oversight frameworks to mitigate environmental, social, and financial risks. The pursuit of greater inclusion in the growth process is becoming a top priority in many countries. Equally, environmental, social as well financial sustainability, challenging as they are, represent crucial objectives for countries. These financial instruments, therefore, ought not to turn out to be mechanisms that serve to avoid or sidestep the vital concerns over social, environmental, and financial sustainability.

The World Bank portfolio, for example, has evolved toward programmatic lending, with up to one-third of current lending projects becoming ill suited to the existing safeguards framework that applies to investment projects. As programmatic financing replaces project financing, there will be a greater role for alternative and adequate oversight mechanisms. In the case of environmental actions, there will need to be a special treatment of high risks involved, third-party verification of compliance with safeguards, and full disclosure of the findings, as well as building in grievance mechanisms.

***

The alignment of intermediate outcomes and final desired results is conditioned by the causal relationship between the two. An understanding of the logical links among inputs, outputs, and intermediate and final outcomes, grounded in both theory and observation, is crucial to bridging interventions to final results. And a willingness to refine or even change intermediate targets as information and experience accumulate can mitigate the concerns for development effectiveness.

Ultimate outcomes are not only more difficult to measure than intermediate outcomes but also more challenging to achieve. For example, improving learning is much harder than increasing enrollment rates; improving the nutrition of children is much harder than launching health service campaigns. There can be interest in favoring the more tangible and visible things in the short horizon. But leaving the desired outcomes untracked could incur high costs because achieving the intermediate outcomes does not ensure that outcomes' realization.

In the case of education, the World Bank's new education strategy is an important step in recognizing and linking intermediate and ultimate goals in education. It acknowledged the findings of evaluations of education projects and focuses on increasing accountability and results as a complement to providing inputs. It seeks to support institutional changes to make structural and behavioral shifts. Reforms require buy-in from a large group of stakeholders. Navigating a nation's political economy is an important challenge besides getting the technical details right and building its implementation capacity. The targeting of learning outcomes in this strategy has implications for setting the direction not only in education but in other areas such as health.

Experiments with new mechanisms to deliver aid in ways that focus attention on ultimate development goals bear watching closely. Early evaluations could help ascertain if and how they promote greater development effectiveness, including if they compromise social, environmental, and financial sustainability, and suggest lessons for whether they should be scaled up.

Notes

1. World Bank (2011b).
2. UNICEF (n.d.).
3. See note 1.
4. IEG (2011a).
5. IEG (2006a).
6. *The Citizen* 2010.
7. Devarajan (2010).
8. IEG (2010g).
9. See note 4.
10. Duflo et al. (2010).
11. Banerjee et al. (2001).
12. Jameel Poverty Action Lab (2009).
13. Barrera-Osorio and Linden (2009).
14. See note 5.
15. See note 4.
16. Country case studies in Mali, Pakistan, Peru, and Romania for IEG (2006a).
17. IEG (2010k).
18. Admassie et al. (2009).
19. White and Masset (2007).
20. Banerjee and Duflo (2011).
21. See also Karlan and Appel (2011).
22. IEG (2010b).
23. Gaarder et al. (2010).
24. Birdsall and Savedoff (2010).

Part III

Applying Lessons to Boost Results

Focusing on results and measuring them can be of far greater value if we can use the findings. For evaluative findings to have an impact, it pays to see their use in context and to present them in time to the audience that can use them.

We can think of at least three circumstances when even good evaluative findings can be underused or misused. First, in a rapidly changing context, simply replicating what has worked in the past might not help future work. Sometimes, it is not enough to repeat successful projects, for some challenges may have been set aside because of risk aversion or inertia-ridden bureaucracies. Focusing on the underlying conditions and adapting to a dynamic situation can be crucial.

Second, focusing only on what has worked and what has not could overlook potential areas where evaluation can bring value: capturing missed opportunities. Evaluating how things worked in relation to stated objectives may not be enough. There are numerous instances when the stated goals missed out on crucial priorities, again because of risk aversion or inertia.

Third, doing the right thing in the right way is most effective when lessons are learned and applied at the right time. Seizing the crucial moment to draw lessons and to influence policy is of the highest importance for evaluation to contribute to greater development effectiveness. The timely dissemination of evaluation results should improve the design of policies, instruments, and institutions.

7

New Challenges Call for Shifts in Direction

There is nothing noble about being superior to some other man. The true nobility is in being superior to your previous self.
—Indian Proverb

Under rapidly changing conditions, evaluative findings from past experience might need to be qualified when applying them to future efforts. When contexts change, the essential nature of the problems might change too. In response, ways of addressing them may also need to change.[1] Replicating past projects, even with highly satisfactory ratings, cannot be assumed to ensure continuing success.[2] And when the emerging needs and challenges have changed, high performance ratings for a past portfolio based on one set of projects do not necessarily indicate that repeating that set would produce high development impacts in the future.

For example, the water and transport portfolios of some of the multilateral banks have had above-average success rates when measured against stated objectives. But the approaches taken thus far in these areas have underemphasized some of the most difficult challenges, which also are crucial priorities for development. While this has allowed the projects actually chosen to achieve higher success rates relative to their objectives, it has left some critical needs unmet.

For example, involvement in water faces greater challenges from climate change, migration of people to coastal zones, and the declining quality of the water available to most major cities and businesses. This is a unique moment for the countries and financial agencies to confront growing water scarcity head on. It calls for a shift in emphasis from some of the easier avenues in building infrastructure to the tougher challenges of managing water resources and the environment.

Similarly, much has been achieved in building transport infrastructure in developing countries, but with the growing population density and rising environmental vulnerability, the solutions of the past—with a heavy focus on roads—will no longer be adequate. Countries and the development community will need to pay far greater attention to efficiency, safety, health, and the environment using cross-cutting approaches.

Although this chapter draws on experience from water and transport, the lessons pertain to many other sectors. It is crucial to learn lessons from the past and to understand links between actions and effects in a given context, beyond replicating the past interventions.

Water Needs Are Changing: More than Access

Water has long been a major focal area for World Bank lending to developing countries and projects in this area have done well. The growing concern about water stress, however, renders the past focus of increasing water accessibility inadequate in addressing the new challenges as the management and sustainability of water use are more prominent. This recognition is a pillar of the institution's new water strategy.

So, sustainable management of water resources has acquired a new urgency. Maintaining a sustainable relationship between water and development requires that current needs be balanced against the needs of future generations.

Growing Concern about Water Scarcity and Water Quality

For almost a century, water use has been growing almost twice as fast as population. To meet the demand for water, numerous parts of the world have exceeded sustainable limits of water withdrawal from rivers and groundwater aquifer. Underground aquifers below New Delhi, Beijing, and many other booming cities are falling rapidly. Major rivers such as the Ganges and Yangtze, or the Nile and the Jordan are overtaxed and regularly shrink for long periods during the year.

Water shortages already loom in many parts of the world. One-third of the world population, concentrated in developing countries, lives in basins where the water deficit is larger than 50 percent.[3] About 700 million people in forty-three countries face water stress, unable to obtain the minimum need of 1,700 cubic meters of water per person per year.[4] And climate change aggravates the erratic rainfall patterns, compounding the challenges.

The global water footprint reached 9,000 billion cubic meters a year in 1996–2005.[5] Irrigated agriculture accounts for more than 80 percent of water use in developing countries. Yet, feeding more people and coping with the changing dietary demands from a richer population will require more efficient water use. Without sufficient water, future economic progress could be severely constrained.[6]

But water stress is about more than availability. Rapid economic growth increases not only water use but also pollution. It has changed natural water reservoirs—directly, by draining aquifers, and indirectly, by melting glaciers and the polar ice caps. And overexploitation of groundwater results in salinization, while industrial and agricultural waste pollutes water sources.

The economic benefits of better managing water resources are big, as are the economic costs of inaction. Country examples indicate that proper water management could increase gross domestic product by 5–14 percent. In the Middle East and North Africa, where water shortages are most acute, the cost of environmental degradation from water pollution and excessive withdrawals is estimated at between 2.0 and 7.4 percent of GDP.

The Focus on Water Access Needs to Be Revisited

The context has changed. Much of the low hanging fruit has been collected. And at the project level, replicating what was rated successful in the past, therefore, may not yield the desired results in the future. Water availability and water quality both have become tougher challenges.

Connecting households to water supply systems used to be the main way to provide access to safe water. But this can no longer be replicated because water quality has been degraded. In many countries water needs to be pretreated to improve quality. Restoring water resources at high cost has become a prerequisite for supplying water to communities.

Vietnam and Brazil provide examples. In Hanoi, ammonia has been found in groundwater since the early 1990s from a series of wells near the Red River, which the water system tapped.[7] To remove this contaminant, as part of the 1997 Vietnam Water Supply Project, equipment costing $2.4 million was installed in the water plant before delivery to households.[8] In the area of Vitória, Brazil, cutting trees to create pasture areas for cattle farms degraded the environment. With cattle farms on steep slopes adjacent to the riverbanks upstream, the

runoff into the river caused high sedimentation, shutting down a dam that provided drinking water and electricity to the city. Clearing the sediment from the reservoir then cost more than what the cattle ranchers earned from farming.[9]

Further, project outcome ratings are measured against stated objectives. But when the negative externality is not accounted for, repeating the same interventions, even rated as satisfactory in the past, can do even more harm than good in the future. High cost is often required to mitigate the loss. Shrimp farming in coastal areas and large-scale irrigation projects are two examples.

Shrimp farming. Investment in shrimp farming in coastal areas was considered a viable approach for job creation and poverty alleviation in the 1980s and 1990s. But since the 1980s, as farmers have converted mangrove swamps into aquaculture ponds, the world has lost 5 million hectares (or one-fourth) of mangrove forest. Shrimp farming also breeds fish and shrimp diseases and causes the salinization of soil and fresh water, further polluting water and the environment.

South and Southeast Asia have experienced particularly dramatic losses. As a result of depleted mangrove forests and drained wetlands, increased sediment runoff reduces the quality of water.

Externally funded shrimp and fish culture projects in three Indian states (Andhra Pradesh, Orissa, and West Bengal) were conducted in the early 1990s to alleviate poverty through employment creation. But shrimp farming increased the vulnerability of local communities to storms and floods and degraded freshwater resources.[10] Two cyclones, one in Andhra Pradesh in 1997 and one in Orissa in 1999, completely destroyed the shrimp farming sites. New instruments, such as replanting mangrove forests and improving water quality, are now required to restore and protect the environment and improve people's livelihoods.

In coastal Vietnam, inadequate attention to mangrove preservation contributed to widespread eradication, with catastrophic environmental impacts. In 2000, the Vietnam Coastal Wetlands Protection Project[11] replanted 370 million trees along 460 kilometers of the coast to restore ecosystems and reduce coastal erosion.[12]

Large-scale irrigation infrastructure. This greatly improved agricultural productivity in the 1970s. Pumping groundwater at a faster rate allowed more hectares to be irrigated and more crops produced. Building dams and reservoirs increased water supplies, particularly

where water was scarce and unevenly distributed.[13] But ignoring the environmental harm of irrigation infrastructure resulted in large and irreversible costs in the long run.

For example, Morocco has managed water security with dams and reservoirs over the past fifty years. It is expected to achieve the Millennium Development Goal for access to safe drinking water thanks to large capacity increases in water storage from heavy investments in dams and reservoirs—the number of dams increased from 12 in 1960 to 114 in 2006.

But there is a limit to this approach. Morocco's dams and reservoirs have indeed sustained a large irrigation system. In 2006, the country had a total irrigated area of 1.4 million hectares, up from 1 million hectares in 1998 and 8,350 hectares in 1967.[14] Yet, further investments in dams will not yield as much benefit as in the past; of the possible dam sites, 80 percent have been filled, leaving only the most difficult sites. Marginal returns have declined. And due to the decline of water quality, not only does building new dams become more complicated and expensive, but the life of the existing dams also becomes much shorter than predicted. Farming and grazing on steep hillsides resulted in environmental degradation and siltation.

Three large pumping stations in China use water from the Yellow River to irrigate farmland. But the shifting of river channels and high sediment have seriously disrupted water supply, at times leaving the stations without water. And concerns are increasing that the Yellow River will no longer reach the sea. In 1997, the Yellow River did not reach the Bohai Sea for 226 days. The Shanxi Poverty Alleviation Project aimed to irrigate 123,000 hectares with river water, but by the time the project was completed in 2004, only 24,400 hectares (or 20 percent) were supplied with irrigation water, leaving 80 percent of the target areas continuing to rely on tube wells, overexploiting groundwater.[15]

So, simply replicating dam construction is not feasible in financial or technical terms. With groundwater threatened by overexploitation and contamination, the sustainability of such large-scale irrigation infrastructure becomes questionable. Today, with growing water scarcity, irrigation schemes often rely heavily on extracting groundwater because of limited and unreliable access to river water. Continuous overexploitation of groundwater can deplete resources. Due to the growing constraints on groundwater and increasing water demand, stringent water use restrictions are required.

Managing Sustainability

Meeting today's water needs while putting in place innovative strategies to address future requirements is a global challenge. Many projects have largely met their stated goals in recent years. The continuing challenge is to reinforce and scale up innovative ways of confronting the mounting problems, including some of the most difficult ones, such as coastal zone management, pollution reduction, and groundwater conservation.

New challenges emerge and strategy needs to change accordingly. For example, coastal zones are home to an ever-growing concentration of people and economic activity, yet they are also subject to climate risks, including rising sea-levels and the greater intensity of tropical storms and cyclones. East Asia and the Pacific and Latin America and the Caribbean account for about two-thirds of the large price tag of total adaptation costs of $75 billion to $100 billion a year.[16] But between 1997 and 2007, the World Bank financed 1,059 projects for water supply and sanitation and only 602 for coastal zones, rivers and lakes, watershed management, and groundwater conservation.[17] New ways need to be found to help countries make water's sustainability a cornerstone of development plans.

Investing in new technologies (such as sprinkler, drip, and pipe irrigation) and implementing regulations (such as water quota systems to raise the efficiency of water use) show promising results.

For example, to increase the efficiency and productivity of agricultural water use, the China Second Tarim Project in 1998 introduced a water quota system restricting agricultural water use.[18] Water-conveyance efficiency was increased from 60 percent to 95 percent and an estimated 600–800 million cubic meters of water were saved each year by lining canals with concrete over geomembranes to prevent leaks. The saved water was reallocated to environmental, municipal, and industrial uses, enabling the reclamation of land and the expansion of irrigation to more than 41,000 hectares of new farmland. Within the project area, nearly 70 percent of people were lifted out of poverty.

To prevent sedimentation, an innovative Global Environment Fund project in Brazil, the Espirito Santo Biodiversity and Watershed Conservation and Restoration, provided a startup fund for transfers between water users in different economic sectors and allowed the power company to provide cattle ranchers with payments for not farming.[19]

Other country examples include paying more attention to wetlands (Vietnam), caution in expanding irrigation that relies on falling underground water tables (the Republic of Yemen), and confronting agricultural water pollution (Morocco).

Evaluations of water projects have fed into the World Bank's water strategy. There is some realignment of the Bank's portfolio and investment priorities, including multipurpose hydraulic development, water supply and sanitation, and watershed management. The sustainability of the resource base is being incorporated into more Bank operations. And the Bank is embarking on the formulation of a strategic approach to ensure that water management and the provision of services are clearly aligned—and that water service operations include proper management strategies.

A water mid-cycle progress report places water at the forefront of the World Bank Group's mandate for sustainable development.[20] Highlighting many critical water issues, it recognizes that changes are needed to respond to the complexities of the rapidly changing environment. It accepts the need to do more about the vulnerability of settlements to flooding and the need for coastal zone management. Addressing water issues at the river-basin level and sequencing approaches have acquired growing importance. More recently, the World Bank Group has been assessing options to reduce the gap between water availability and use by involving all stakeholders in efforts to support government transformation programs. More projects are systematically linking the availability of water with its use, with a more holistic look at the impacts on health, poverty, and the society. To improve knowledge of water resource management, more evidence needs to be provided on the actions taken to ensure that projects pay adequate attention to conserving groundwater and that the quantity extracted is sustainable. And more evidence is required on effective ways to help countries address coastal management.

Transport Needs Are Changing: More than Roads

Sharply rising populations in coming decades, especially in the developing world's urban areas, coupled with continuing globalization and trade liberalization is expected to significantly accelerate demand for transporting people and goods. Rapid urbanization, congestion, pollution, and resource overuse exacerbate the negative impacts of transport investments.

Meeting the demand for transportation requires more than building highways. Urban transportation issues have to be dealt with to support passenger and freight mobility in large urban agglomerations.

Rapid Urbanization Poses New Challenges

The number of cities exceeding 1 million inhabitants is on track to surge from 268 in 2000 to 358 in 2015.[21] In the one of the most dynamic regions, East Asia and the Pacific, cities account for 70 percent of GDP growth, a trend likely to continue.

Current global trends indicate growth of about 50 million urbanites each year, roughly a million a week. More than 90 percent of it occurs in developing countries, placing intense pressure on urban infrastructure, particularly transport. Since 1950, the world's urban population has more than doubled as a result of natural increase and migration, to reach nearly 3.5 billion in 2010, about 50.6 percent of the global population.[22] By 2050, 6.4 billion people, about two-thirds of humanity, are likely to be urban residents. Urban mobility problems increase proportionally with urbanization, and in some cases exponentially, since it concentrates mobility demands in specific areas.

Surging numbers of people and vehicles threatens the environment, the climate, and the quality of life. With the global pool of motor vehicles growing 3 percent a year, the pressure for additional road space is relentless. Vehicle miles traveled in the United States has outpaced population growth since 1970.[23] More recently, car ownership has surged in the developing world, including the most populous countries, China and India. Air pollution causes the premature deaths of 650,000 people a year in developing countries.[24] The consequences of pollution and environmental degradation fall disproportionately on the poor, and road transport already accounts for nearly a quarter of manmade gases contributing to climate change. If the past trend of transport development were to continue, transport greenhouse gas emission would increase even more dramatically. Policies to guide demand to low-emission modes and technologies become more and more important.

Transport can help resolve the nexus of issues associated with energy, land use, urbanization, and climate change. But economic activities concentrate on cities, and transportation in urban areas is highly complex because of the multitude of modes, origins and destinations, and the amount and variety of traffic.

The Focus on Highways Needs to Be Reassessed

The World Bank's past support for transport has been well managed and effective, with above-average project ratings. But to meet the new challenges, shifting away from the focus on highways is required. More attention to efficiency, safety, health, and the environment and a shift to urban-focused and multimodal transport are required to address the tougher issues, such as growing population density and rising environmental vulnerability.[25]

Between 1995 and 2005, the World Bank's experience in transport was rated successful overall.[26] Of ninety-seven Transport Sector Board road projects with specific physical upgrading and rehabilitation objectives completed and evaluated, 79 percent were rated moderately successful or better on outcomes. The average economic rate of return was 29 percent, above the median for all Bank projects.

Sound design and construction contributed to the positive outcomes of these projects.[27] Armenia rehabilitated more than 60 kilometers of railway tracks between Yerevan and the Georgian border and undertook pricing and operational policy reforms to improve railway finances. The China Qinba Mountains Reduction Project surpassed all rural infrastructure targets for road construction. In Peru, rural accessibility improved, and transport tariffs for freight declined 9–15 percent. In Nicaragua, the project's success was in part due to successful donor coordination.

But there has been a predominance of road projects (Table 7.1). For fiscal year 2001–2006, 73 percent of total commitments for transport went to roads, but only 8 percent to railways, 3 percent to ports, and 3 percent to aviation. Urban transport, mostly classified under "general transport," stood for only 13 percent of the total commitments. If urban roads and streets are separated from "general transport" and added to the roads category, the percentage share of road-related transport commitments rises to almost 80 percent of the portfolio.

A focus on roads is not surprising, given that they carry more than 80 percent of passenger kilometers and a significant percentage of freight ton-kilometers in all countries.[28] But continually constructing roads is not likely to produce satisfactory outcomes as new demand emerges alongside rapid urbanization.

A transport project rated successful against its stated objective could produce unaccounted negative externalities. Expanded land

Table 7.1
Predominance of Road Projects

Transport mode	Fiscal year 1996–2000		Fiscal year 2001–2006	
	World Bank Commitments ($ billions)	Percent	World Bank Commitments ($ billions)	Percent
Roads	13.0	73	11.9	73
Railways	1.5	9	1.3	8
Ports	1.2	6	0.5	3
Aviation	0.1	0	0.5	3
General transport	2.2	12	2.2	13
Total	17.9	100	16.3	100

Note: Multimodal projects have been redistributed to the appropriate modes. Totals may not add up exactly as a result of rounding.

Source: World Bank data.

use for transport facilities, deforestation in rural areas, and increasing motorization in urban areas can increase air pollution, noise, and traffic accidents. As the environment has become more important, simply replicating transport projects with high outcome ratings in the past would not always be good for the future.

Roads, if not managed appropriately, can be a source of water pollution as well. Unpaved roads can contribute to basin-wide runoff and stream sediment. Because water begins flowing over land on compacted earthen surfaces after only a little rainfall, the flows go directly to nearby streams. Surface preparation processes performed when these roads are maintained renew the supply of easily transportable surface sediment. According to some estimates, unpaved roads appear to be on the same order of importance as agricultural lands in contributing sediment to stream networks, despite occupying a fraction of the total surface area in basins.[29]

Large-scale paving can reduce infiltration. Natural storage of groundwater can be reduced by improving drainage (which, by definition, removes water to another area). Streams are often constricted by roads. Bridges may constrict the flow of water, especially

when designers are unaware of peak rainy season flows, and they can act as unintended dams if debris jams their openings. In some geographic areas the construction of roads can even exacerbate the conditions that cause floods.[30]

For example, burning forests to clear land for agriculture has spread rapidly in Brazil.[31] But building roads in forested areas can also result in deforestation and erosion. The Brazil Northwest Region Integrated Development Program, or Polonoroeste program, aimed to provide sustainable settlements by expanding infrastructure and supporting agriculture and social services in the agricultural frontier areas of Rondonia and Western Mato Grosso. The program supported paving the federal highway between Cuiaba and Porto Velho, extending feeder road networks in both states, consolidating existing settlement schemes, and establishing new ones. It also aimed to improve rural social services, including management of the natural environment and protection of indigenous people. In the absence of effective environmental concerns and controls, however, it led to land use changes, uncontrolled deforestation, and environmentally unsustainable development in the 1980s and beyond.[32] As a result, the increase in runoff made its way into the rainforest through the road system, with close to 75 percent of the deforestation taking place within 50 kilometers of a paved road. The sad lessons of the environmental devastation caused by such infrastructure projects may still not have been learned—at a high cost to society.

The Future Direction: Multimodal and Environmentally Friendly Approaches

The challenge for transport is to support economic growth, but also to ameliorate the negative social and environmental impacts. Overlaying the general scenario of economic growth is a second phenomenon of rapid urbanization, congestion, pollution, and resource overuse. The transport sector must look for more sustainable solutions than in the past. The way forward is to ensure that transport operations go beyond intercity highways and give more attention to environmental damage, energy efficiency, climate change, traffic congestion, and safety.

Evidence shows a strong rebalancing between the various transport modes in the World Bank's portfolio in the recent years. As a share of the Bank's lending in fiscal year 2010, roads and highways came down to roughly 50 percent, and urban transport rose to 30 percent, railways

to 15 percent, and ports and air transport to 5 percent. More attention has been given to energy efficiency and climate change, particularly in urban transport. Greater innovation is required to development infrastructure that is environmentally and socially sustainable.

The next generation of projects needs have a much more urban focus. Much of the growth in the world's population for the foreseeable future will take place in the cities and towns of the developing world. It will be essential to see transport opportunities with a multimodal setting of integrated urban and rural concerns.

The Jamuna River Bridge in Bangladesh is an example how a multifunctional structure can stimulate development and reduce poverty.[33] Before the bridge was constructed the only connection was a slow-moving ferry; traffic jams at the ferry terminals often lasted for days. The bridge, nearly 5 kilometers long, carries a four-lane highway, rail line, utility connections, and fiber optic cables. It aimed at connecting Bangladesh's less developed northwestern region with its more developed eastern region to accelerate economic growth in the northwest and to integrate the area more fully into the economy of the nation. The completed project reduced journey times and transport operating costs and stimulated interregional trade. And even though the road was subject to tolls, the level of traffic has been 41 percent higher than expected.[34] Computer simulations predict major shifts of persons out of abject poverty locally, with wider distributional improvements to both local and national welfare.

To meet the emerging challenges, urban planning programs are also required. To improve the integration between services and increase the access of urban poor to employment centers, health centers, and educational facilities, integrative public transport, land use, and air quality strategies are important. For example, extending and modernizing commuter rail systems and bus corridors is an effective way to alleviate heavy urban traffic congestion in Brazil. And the installation of centralized traffic management systems helps to regulate traffic flows in Bangladesh and Vietnam.

Given the increasing links with energy, land use, urbanization, the environment, and climate change, transport will require innovative cross-cutting approaches. One example is new parking strategies to improve traffic flow, generate revenue, and discourage car usage. The Bangladesh–Dhaka Urban Transport Project introduced the Non-Motorized Traffic/Transport conversion zone in several cities in Bangladesh. As a result, accidents have fallen 85 percent, and fatal

accidents 33 percent, on the Non-Motorized Traffic/Transport corridor during the three months of conversion.

Although the demand for highways will drive it core business, it is anticipated that clients will increasingly seek support for more complex projects and that this will gradually lead to a significant redeployment of resources and a reexamination of priorities.[35] In China, where traditional expressway projects were among the most successful in the sector, the government has enlisted World Bank support for transport projects connecting interior regions with the coast. The project pipeline is shifting toward railway, inland waterway, and urban transport projects, so the share of such projects in technical assistance and investment lending continues to increase.

In the new dynamic context, developing environment-friendly and energy-saving transportation is under the spotlight. Projects in several countries show promising results. In Dhaka, the removal of the highly polluting three-wheeled taxis with two-stroke engines, under the World Bank-supported Air Quality Management Project, is one example of addressing the worsening quality of urban air. In Mexico, the Transport Air Quality Management Project for the Mexico City metropolitan area reduced ambient concentrations of pollutants and resulted in fewer respiratory illnesses and other acute syndromes of poor quality air.

Low-carbon technologies might seem costly at the beginning, but can pay off in the long run. The durability of transport equipment, the longevity of transport infrastructure, and the high fixed costs mean that current investments lock in the modal structure of transport for decades.[36] Once investments have been made, the expenditures are sunk. It is of vital importance to address sustainability.

Overall, however, evaluative evidence of whether and how World Bank projects worked in addressing the environmental issues is scant. Only six Bank projects (between 1995 and 2005) for the urban environment and air quality in ten years have been completed and evaluated. So, efforts to strengthen monitoring and evaluation are still required.

***

The pace of change in the world is accelerating. The positive impact of economic growth is coupled with relentless damages to the environment. New challenges are emerging. Replicating past experiences, successful in their original context, may not yield satisfac-

tory results in the future. The focus needs to shift to tackle the toughest challenges in order to achieve better development results.

Such shifts might call for moving beyond repeating easy-to-implement projects and taking on riskier activities in the interest of greater development impact. The challenge is to ensure that project performance ratings take into account that results are being achieved under more difficult circumstances and that there is continuing encouragement from evaluation for taking on riskier directions.

The urgency for adjusting approaches to the emerging needs and priorities applies to many areas of concerns, but making such shifts is sometimes not adequately motivated when evaluations rate past successes without bringing out their dynamic context. In such a context, replicating the past interventions, like mechanically ticking the boxes on the check list, would not be adequate or even useful. For evaluative lessons to improve development effectiveness, the key is to understand what worked and what did not, and under what conditions.

Notes

1. Sachs (2008); Kohli et al. (2011).
2. Lin and Monga (2010).
3. The 2030 Water Resources Group (2009).
4. UNDP (2006).
5. Water Footprint. "Introduction." www.waterfootprint.org.
6. The state of world water has been characterized as a crisis by the United Nations and other international organizations as well as in the popular press. See, for example, DfID (2008), Lall et al. (2008), and Rogers (2008).
7. IEG (2006b).
8. Ibid.
9. Espirito Santo Biodiversity and Watershed Conservation and Restoration Project (P094233).
10. Primavera (2005).
11. Vietnam Coastal Wetlands Protection Project.
12. Ibid.
13. Parker and Skytta (2000).
14. Swearingen (1987).
15. China Shanxi Poverty Alleviation Project.
16. World Bank (2010c).
17. IEG (2010i).
18. China Second Tarim Basin Project IEG (2009k); World Bank (2007a).
19. See note 9.
20. World Bank (2010f).
21. UN Habitat (2001).
22. Rodrigue et al. (2009).
23. U.S. Census Bureau (2001).
24. IEG (2007a).

25. Ibid.
26. Ibid.
27. Ibid.
28. Ibid.
29. Ziegler et al. (2004).
30. See note 24.
31. Alves (2002).
32. Redwood (1993).
33. Asian Development Bank (2005); Luppino et al. (2004); IEG (2000).
34. Ibid.
35. See note 24.
36. Kopp (forthcoming).

8

Capture Opportunities

Two roads diverged in a wood, and I, I took the one less travelled by,
And that has made all the difference.

—Robert Frost

One of the important roles for evaluation is to draw lessons from the past to inform future actions. Traditionally, development evaluation has focused on what actually happened—what went right or wrong—in projects, programs, policies, or other activities to extract such lessons. This approach generally makes sense, since it relies on the facts on the ground for its analytical base, providing observational support for the lessons.

In a rapidly evolving and increasingly complex environment, an often underappreciated role for development evaluation is to identify opportunities in operational work—a form of counterfactual analysis. It looks not at what *did* happen, but what *might have* happened. The intent is to identify opportunities missed in practice but that might have made a difference if implemented. As evaluation becomes more forward-looking, this kind of analysis becomes more important.

This chapter considers issues around environment and climate change, structural constraints, and the use of cost–benefit analysis for decision making. It focuses on three kinds of constraints. The first involves overcoming apparent conflicts between policy objectives by finding opportunities to exploit win–win situations. The example here has to do with climate change and environmental protection and how addressing them could support economic development. The second is overcoming internal institutional constraints, such as limitations placed by governing charters and policies. The examples cover the use of guarantees and the application of safeguards. And the third relates to information constraints that may make it more difficult to choose rationally among options, forgoing good opportunities. The example comes from the neglect of cost–benefit analysis.

Environmental Protection and Development Can Be Win–Win

Climate change and environmental protection have emerged as major development issues over the past couple of decades. Both developed and developing countries have been struggling with how to cope with them, leading to often-contentious disputes over the distribution of costs and benefits. One especially vexing concern is that policies to protect the environment and reduce or mitigate climate change will slow economic development. This kind of tradeoff is difficult for policymakers to embrace or to persuade their constituencies to accept. But if no actions are taken, there is danger of irreparable harm, which could have particularly adverse consequences for developing countries and their people.

While there are no easy solutions to this conundrum, there may be opportunities to use sound environmental policies to promote development while still addressing climate change. Indeed, recent work by IEG provides some examples related to energy efficiency and protected forest areas that show how identifying missed opportunities can provide practical guidance to achieve just such results. These potential win–win strategies can address urgent environmental concerns and economic needs simultaneously.

A 2°C warmer world will experience more intense rainfall and more frequent and more intense droughts, floods, heat waves, and other extreme weather events. That would have dramatic implications for how countries manage their economies, care for their people, and design their development paths. The cost of adapting to a 2°C warmer world is in the range of $75 billion to $100 billion a year between 2010 and 2050.[1] This range is of the same order of magnitude as the foreign aid that developed countries now give developing countries each year.

The cost of adaptation is high. The less done to mitigate climate change, the more severe and expensive are the consequences. Delays in action increase the costs because impacts worsen and cheap mitigation options disappear as economies become locked into high-carbon infrastructure and lifestyles—more inertia. Some negative impacts are irreversible if actions are not taken in time.

Energy Subsidies Can Be Better Targeted

A byproduct of economic development is the production of greenhouse gases that threaten climate change. A 1 percent increase in per capita income induces—on average and with exceptions—a 1 percent

increase in greenhouse gas emissions.[2] But some exceptions offer opportunities to promote strategies that both promote growth and limit emissions.

This implies that strategic tradeoffs can be addressed by development policy. In energy the World Bank has long supported reductions in subsidies, coupled with improvements in efficiency and greater availability of energy for the poor. On one level, these three policies appear inconsistent. But this need not be so.

Subsidies are a large but poorly monitored drag on developing economies. Removing them would increase economic efficiency and reduce greenhouse gas emissions. The Organisation for Economic Co-operation and Development has about €29 billion in subsidies, mostly to energy producers.[3] The International Energy Agency estimated that there was about $250 billion in annual consumption subsidies for electricity and fossil fuels outside the Organisation for Economic Co-operation and Development in 2005.[4] The largest subsidizers in absolute terms were the Russian Federation, the Islamic Republic of Iran, Saudi Arabia, India, Indonesia, Ukraine, and Egypt—all with more than $10 billion a year in subsidies. Subsidies are 2–7.5 times larger than public spending on health in Bangladesh, Ecuador, Egypt, India, Morocco, the Republic of Yemen, Pakistan, Turkmenistan, and Venezuela.

The bulk of energy subsidies, however, go to better-off consumers. Most poor people in developing countries get no direct benefit from fuel and gasoline subsidies, because they are not connected to the electric grid and do not own cars. They receive only indirect benefits through lower prices for energy-intensive goods and services such as public transit. Even when the indirect benefits are considered, the bottom 40 percent of the population in Bolivia, Ghana, Jordan, Mali, and Sri Lanka received only 15–20 percent of fuel subsidies. In principle, moving to higher energy prices could yield fiscal dividends and long-run reductions in greenhouse gas emissions. Also in principle, reallocating the savings from lower subsidies and lower energy use could benefit poor people and society.

Indeed, in European transition countries cutting energy subsidies led to increased efficiency and reduced emissions as prices to end-users rose toward their full production cost.[5] Subsidies tend to go disproportionately to the better off, providing few benefits for the poor. Increased access to energy for the poor has very little effect on emissions.

Using energy-efficient instruments can also help. Compact fluorescent lamps draw only 20–30 percent as much power and last much longer.[6] Substituting them for all the incandescent lamps in sub-Saharan Africa would reduce peak power consumption by 15 gigawatts, roughly 23 percent of the installed capacity.

Together, these findings suggest that a win–win strategy could be built around introducing efficiencies, while reducing subsidies and better targeting them to the poor. This would simultaneously reduce the strain on government budgets, free resources to allow extension of energy sources to the poor, and promote more efficient energy use.

But opportunities to pursue such a strategy need to be taken. Subsidy reductions could be linked to efforts to improve energy efficiency as a way to reduce the burden of adjustment to higher prices. And lending from multilateral banks for energy efficiency improvement has been modest—there has been some work on pricing.[7] A major reason for the imbalance is that internal Bank incentives tend to work against such projects, which often are small but place heavy demands on staff time and preparation funds and may require persistent client engagement that could take years. Moreover, investments in power generation are highly visible and easily understood, unlike investments in efficiency, which are less visible and involve human behavior (rather than electrical engineering), whose efficacy is harder to measure.

In addition, a general lack of monitoring and evaluation in energy projects reinforces the negative view of efficiency-oriented projects as demanding but not necessarily effective. There is no timely, comprehensive, and consistent international monitoring of energy subsidies or prices, and basic national data on key factors related to energy efficiency are mostly lacking. Timely and accurate data on household, commercial, municipal, and industrial consumption and expenditures on energy also are in short supply. Without such basic data the win–win opportunities for greater energy efficiency, reduced subsidies, and better targeting to improve access of the poor are difficult to identify and measure. And so they remain missed.

Protected Forest Areas Can Be Better Managed

Another way to address climate change is by protecting forest areas. The loss of tropical forest is a major contributor to climate change, accounting for some 17 percent of the human contribution to greenhouse gas emissions, primarily through the burning of forest lands to clear them for other uses.[8]

The Reduced Emissions from Deforestation and Degradation (REDD) agenda, under the United Nations Framework Convention on Climate Change, is intended to protect forest areas while safeguarding the local environment and meeting development goals. Of course, efforts to protect the forest are much older than REDD, stretching back decades and focused primarily on fostering biodiversity and maintaining sustainable timber supplies. Reducing greenhouse gas emissions has been a side benefit of such policies, but not their primary purpose. REDD puts climate change concerns explicitly on the deforestation prevention agenda. But a lack of evaluation leaves the effectiveness of all these efforts largely unknown.

An evaluation using forest fires as a proxy found that, on average, protected areas significantly reduce tropical deforestation and associated carbon release, thus reducing carbon emissions while preserving biodiversity.[9] The study examined whether areas subject to strict protection—with essentially no use allowed—fared better than those in which some activity was permitted.

The expectation was that, all things equal, strict protection would have the bigger impact on reducing the incidence of fires, considering differences in deforestation pressures. But evaluation found instead that the impact was actually greater when the protected areas allowed sustainable use by local populations than when they did not (Table 8.1). This finding is true for Africa, Asia, and Latin America and the Caribbean when comparing the mean reduction in fire incidence from strict protected areas with that from multiuse protected areas. In Latin America, where indigenous areas can be identified, the impact on fire incidence is extremely large.

The value of protection was even greater in areas where people use forests to generate income. This counterintuitive finding suggests that giving local people a stake in the preservation of the forest, through sustainable and economic activities, can simultaneously address forest protection and economic development within the REDD agenda. Some kinds of land use restrictions—rather than fencing off the areas—can contribute to biodiversity conservation and climate change mitigation. This implies a degree of compatibility between environmental goals (carbon storage and biodiversity conservation) and support for local livelihoods. And it sheds important light on the debate comparing the effectiveness of strictly protected areas with areas that allow local people some degree of sustainable use.

Table 8.1
Estimated Protected Area Impacts on Fire Incidence (Percent)

Area	Mean fire incidence	Mean reduction due to strict protected areas	Mean reduction due to multiuse protected areas	Mean reduction due to indigenous areas
Latin America and the Caribbean	7.4	2.7–4.3	4.8–6.4	16.3–16.5
		3.8–7.7	*6.2–7.5*	*12.7–12.8*
Africa	6.1	1.0–1.3	(0.1) –3.0	Not applicable
		2.4–4.5	*Not calculated*	
Asia	5.5	1.7–2.0	4.3–4.9	Not applicable
		2.9–3.1	*5.1–6.7*	

Note: The table includes estimates of two periods as a robust test. The first line (non-italics) indicates estimates for the pre 2000 protected areas. Italics indicate estimates for protected areas established between 1990 and 2000.

Source: Nelson and Chomitz (forthcoming).

Zoning for sustainable use may be more politically feasible and socially acceptable than designating strict protection in areas less remote and with higher population densities. In short, REDD can be win–win, not merely putting up fences to conserve forests but also addressing the fundamental interests of communities and promoting local environmental and development goals.

Structural Constraints Affect Development Performance

Opportunities can also be missed because of structural constraints that affect the ability of development organizations to provide optimal services to clients. Two examples of such constraints are the limitations on the MIGA's lending activities and the gaps in the World Bank's safeguard policies.

The Multilateral Investment Guarantee Agency's Convention Limits Its Ability

The MIGA's fundamental purpose is to provide political risk insurance to promote private investment. The kinds of risks it may

insure are governed by its convention, adopted in 1988. The problem is that MIGA's basic product range for risk insurance has not changed since, despite the major changes in the global financial markets in the 1990s and 2000s.

A study found that other providers of risk insurance were developing a range of new products in response to the changing market conditions.[10] MIGA was expected to take the lead and expand its product lines to meet new areas of demand. In response, it did take some steps in its 2005 strategy to address the concerns in the study. But its convention and operational regulations restricted the eligibility requirements for insurance and the types of risks covered. These limitations hindered MIGA's ability to adapt to market trends, and the agency has not been sufficiently aggressive in innovating within the flexibility allowed by the policies.

While MIGA was facing these limitations, the international political risk insurance industry saw rapid growth—tripling in size over 2004–2007—and MIGA's share in the market declined from 6 percent to 4 percent. This in part reflected growth in the political risk insurance business that MIGA could not offer. MIGA's annual business volume did increase, but only because the average size of guarantees went up, even as the number of supported projects declined.

The evaluation further concluded that the convention on eligible risks had hampered MIGA's effectiveness—including in IDA countries—and that restrictions were limiting its ability to respond to country needs during the global financial crisis.[11] The result was that MIGA was missing opportunities to enhance development because of these structural constraints.

The study came out when the global financial crisis was increasing the need for MIGA to enhance its range of products to better serve client countries. MIGA has begun addressing institutional effectiveness, amending operational regulations and some policies, conducting a business process review, and taking a first step toward implementing self-evaluation. In late 2010, its board of governors amended the convention to address the constraints.

Unbalanced Safeguard Policies Underestimate Risks

In 1997, the World Bank identified ten policies as safeguards that would direct development in ways that would do no harm. These included six policies on aspects of the environment, two on social

concerns, and two on legal issues, marking an important milestone for the Bank in addressing some of the concerns about potential (unintended) effects of development activities financed through Bank loans. Subsequently, other development agencies adopted such policies, initially based on the Bank's. The IFC and MIGA developed a set of performance standards separate from the World Bank's safeguard policy framework.

A 2010 evaluation found that the World Bank's safeguard policies needed to be more balanced, with several enhancements in the area of social risks.[12] It concluded that the safeguards framework excluded from scrutiny some project-induced risks covered by IFC/MIGA performance standards, as well as other World Bank policies and guidelines. More specifically, the coverage of social issues is too narrow. For example, the Bank does not examine labor and working conditions, community health, safety, and security, as IFC and MIGA do.

These coverage limitations lead to an underestimation of risks, which can impair development results. If a development risk is identified in advance, steps can be taken to mitigate any adverse implications. Identifying risk does not mean that mitigation will take place—for that requires action. But without early recognition, it is unlikely that mitigation efforts will be undertaken, exposing operations to risk.

The implication is that by taking some of the social risks fully into account, the World Bank could capture an opportunity to enhance its development effectiveness. Several other multilateral development banks have broadened their safeguard coverage over the years, making it more likely that priority risks more relevant to client and project needs will be addressed.

Adding these social policies to the World Bank's safeguard framework would not imply an additional burden on the Bank and its clients. It would be possible to consolidate social safeguards under one umbrella policy that provides for an integrated social assessment, which would allow for identifying the most significant social risks relevant to each project, just as identifying the most significant environmental risks is facilitated by policy in that area.

Such policy changes would improve the *relevance* and *efficiency* of the World Bank's environmental and social policies, but to improve *effectiveness* steps are needed to strengthen implementation, supervision, monitoring, and reporting.

Cost–Benefit Analysis Is Underused

One important economic tool for considering project investment is cost–benefit analysis, a method for ensuring that the value of the benefits from a project exceeds its costs. Several measures associated with cost–benefit analysis (such as net present value, net benefits, and the benefit–cost ratio) are useful under different circumstances.

In principle, World Bank policy requires cost–benefit analysis for every project, except where costs and benefits cannot be adequately quantified. In practice, the percentage of projects for which this kind of analysis is undertaken is low—and had been declining for three decades until the early 2000s. In 1970, about 70 percent of projects included ex ante analysis; by the early 2000s, the figure was down to 30 percent.[13] Ex post cost–benefit analysis followed a similar pattern.

Application of cost–benefit analysis varies widely by sector: it is used most in transport (58 percent of projects, 1970–2008) and least in education, health, nutrition, population, and public sector governance (1 percent each). This disparity presumably reflects differences in the sectors' abilities to conduct cost–benefit analysis.

The World Bank produced a series of reports in the 1990s—known as Econ I, II, and III—that tried to address some of these issues. It provided guidance on what should count as costs and benefits, including those that were indirect and unintended. It called for high technical standards, but practice has not lived up to promise.

Not just the use but also the formulation and timing of the analysis have undermined its effectiveness. The estimates of the economic rates of return are often biased and delayed. Of fifty-one project leaders randomly surveyed from projects closed in fiscal 2006–2007 and 2008–2009, only five reported that cost–benefit analysis is given significant weight at the project identification stage, and eighteen reported that it is given significant weight at the preparation stage.

Many cost–benefit analyses were conducted after the decision to proceed. This put the analysis under considerable pressure to reach conclusions consistent with the decisions already taken, rather than provide critical information for decision making about whether the project is justified. In addition, the likelihood that the economic rate of return is recalculated at the close of projects is lower for projects with low outcome ratings.

This decline in the use of cost–benefit analysis is a forgone opportunity. True, the estimates are imperfect. There are always measurement

issues, such as the need to identify all (or most) costs and benefits. These can be difficult problems, both intellectually and practically. But without some form of rigorous, transparent analysis to support proposed projects, it is almost impossible to make rational decisions on their value.

But there is a second sense in which the failure to pursue such analysis is a missed opportunity. Cost–benefit analysis cannot only demonstrate, to a reasonable degree, whether a project provides sufficient net benefits to invest scarce resources in it, but also provide a common measuring stick for comparing alternatives. It can help decision makers choose among alternative uses of resources, providing a firmer basis for choices than is likely to be available otherwise.

***

Opportunities are often missed because of political, bureaucratic, or analytical constraints that make it difficult to identify the most beneficial options available to decision makers. These constraints often manifest themselves in conflicts among policy goals, institutional rigidities, and analytical weaknesses. Overcoming these constraints can shift behavior and improve development results. Seeing the results of capitalizing on opportunities can be the best encouragement for continuing to make these efforts.

Evaluation methods need to adjust to better take advantage of opportunities. Evaluating against stated goals often leads the evaluator to look for lost keys under the lamppost and refrain from asking whether important options have been considered. When there are good examples of how such an investigation makes a difference, it pays to disseminate them to generate support for such investigation.

Notes

1. Margulis and Narain (2010).
2. IEG (2009b).
3. Morgan (2007).
4. IEA (2007).
5. See note 2.
6. Ibid.
7. IEG (2009b); Asian Development Bank (2009b).
8. UK 8020 (2011).
9. See note 2.
10. IEG (2009d).
11. Ibid.
12. IEG (2010f).
13. IEG (2010a).

9

Tarry Not—For Timing Is (Almost) Everything

It's easy to play any musical instrument: all you have to do is touch the right key at the right time and the instrument will play itself.
—Johann Sebastian Bach

The final lesson has to do with evaluation itself: when and how evaluation is conducted makes a big difference in how useful—and how well used—the information is for improving development results. It is not enough to focus on the right issues and to employ sound indicators and methods in monitoring and evaluating development agency efforts. To be useful, evaluations must be conducted at the right time, delivered in an understandable format, and based on collaboration and follow-up with evaluees and stakeholders.

To be effective, evaluations must convey clear messages, supported by persuasive data. They must also consider the limits in using information from the past to inform decisions on future actions. Without close interaction of policymakers and evaluators, development evaluation will gradually lose its relevance in the global system.[1]

First, the information needs to be useful to the end-users—the decision makers.[2] There is much literature on the use of evaluation.[3] Second, the information has to be user-friendly. Piling up undigested data usually does not meet this criterion. Frequently, the users expect some analysis of what happened and why. But "performance data do not, by themselves, tell why outcomes occurred."[4] Translating the data into an explanation for the outcomes is a role that evaluation can play.

Evaluation information needs to be delivered when it can be put to use. Timing may not be everything, as an old adage has it, but it is almost everything. Information delivered too long before it can be used to make decisions is likely to be forgotten or outdated by the time it is

needed. And information delivered after a decision wastes resources. It is hard to predict at the outset when evaluation data will prove most useful, so serendipity comes into play. But depending on serendipity is a poor substitute for making a good effort to target delivery for a decision point when the information can be used.

For high impact it is crucial to learn faster what works and what does not, to focus on results at the right time, and to link evaluation findings to development actions. This means that evaluators need to be cognizant of the political dialogue and prepared to infuse it with relevant evaluative evidence when critical decisions are about to be made.

Traditional Ex Post Evaluation Has Limits

Independent evaluation has focused traditionally on ex post assessments of projects and programs. At the World Bank, for example, the charge President Robert McNamara gave in 1970 was explicit: "The principal task of this unit will be to review past lending operations with the central objective of establishing whether the actual benefits of the completed projects are in accordance with those expected at the time of appraisal and, in case of divergence, the reasons."[5]

Eventually the World Bank developed a system of self-evaluation of projects completed by operational staff, backed by independent assessment. Other evaluation offices adopted similar systems, and in 1999 the systems were codified into Good Practice Standards by the Evaluation Cooperation Group, representing the independent evaluation units in the major international financial institutions.

Ex post project evaluation has strengths—and limitations. It is especially useful for enforcing accountability by focusing on how development agencies have used funds to fulfill promises to their governing authorities. It allows time for results to emerge before rendering judgment. It also provides opportunities to learn from experience—to a point.

Three limitations apply to learning from ex post project evaluation. First, lessons from any one project necessarily are affected by its circumstances, and so are hard to generalize to similar projects, none ever precisely the same in design or context. Second, changes in the aid architecture over time have shifted the unit of account for achieving development results from projects to the country, region, and globe. Project evaluation on its own cannot address these levels of accountability. Third, internal and external pressures have forced results-based

management on development agencies since the late 1990s.[6] This shift has put a demand on evaluation to provide feedback in the form of findings and lessons much farther upstream than traditional ex post evaluation could accommodate.

Evaluation has tried to address these issues by developing new tools and adapting old ones to changing circumstances. One set of changes has involved developing country, regional, global, and sector-thematic evaluations. A second has been adopting upstream evaluation tools to provide more timely information to decision makers, as discussed in the next section. And a third has involved more effective communication of evaluation findings and lessons, as discussed in the subsequent section.

Forward Looking Evaluation Can Provide More Timely Information

The limits to learning from ex post project evaluation were addressed initially by the Operations Evaluation Department (subsequently IEG) in a renewal strategy in 1997. Among the objectives was shortening the feedback loop from evaluation to operations. Early efforts focused on timing country evaluations, introduced in the mid-1990s, to coincide with developing new country assistance strategies in selected borrower countries. The aim was to have available lessons, from an independent evaluation, of work under the previous country assistance strategy to inform deliberations on the new one, thus making evaluation more useful—and (it was hoped) the new program more effective.

The World Bank eventually adopted a system of completion reports for country assistance strategies as self-evaluation tools, subject to validation, intended to ensure that lessons of experience were taken into account in developing country programs. IEG also better timed its sector evaluations to address the expected adoption or early implementation of new sector strategies.

These were the first major steps toward an approach to evaluation focusing less on accountability for past efforts and more on using the tools of evaluation to improve ongoing or proposed development work. The value of this kind of evaluation can be high, particularly for programs expected to be scaled up, where the risk of wasting funds on ineffective programs or failing to support successful ones is acute. But there can be a tradeoff between timeliness and quality, especially where evaluative evidence is thin or incomplete.

Two approaches that some development evaluators have begun to adopt are *real-time* and *prospective* evaluation.

Real-Time Evaluation

Evaluating projects, programs, policies, or other activities in real time is not a new idea among evaluators. Fairly early in the modern history of evaluation, practitioners realized that the underlying model in which a project or program was designed, approved, implemented, and evaluated against planned objectives, outputs, or outcomes was overly simplistic. Instead, they found themselves attending to all aspects of these activities, including the realism and practicality of project or program design, how well the project or program was targeted at intended beneficiaries, the way it was implemented, and whether it was even feasible to evaluate. This is captured in the distinction between *formative* and *summative* evaluation.[7] Real-time evaluation has come to development mainly since 2000.[8]

The reasons for growing interest in real-time evaluation are not hard to discern. First, in some cases real-time evaluation can avoid costly mistakes and prevent real harm. In the United States the Drug Abuse Resistance Education program, intended to deter use of controlled drugs by young people, was assumed to be effective, and eventually was adopted in 75 percent of school districts nationwide. But evaluation found that it was ineffective and thus a waste of financial resources and school time.[9] Another US example, Scared Straight, sought to reduce juvenile delinquency by taking at-risk youths to visit prisons, where convicts tried to talk them out of following a criminal path. But this well-intentioned program led to higher delinquency rates among participants than nonparticipants, suggesting that meeting the prisoners actually made them more attractive as role models to the youths.[10]

Second, in other cases, evaluation can support effective programs. Early results from the evaluation of Mexico's conditional cash transfer program *Progresa* (now *Oportunidades*) showed positive impacts on schooling, health, labor supply, and consumption.[11] Documenting these positive results helped persuade a new administration not only to maintain the program but also to expand it to new areas and extend eligibility to more children in the areas originally covered. The evaluation also encouraged the administration to embrace a program of rigorous impact evaluation more generally in developing its social safety net programs. Similarly, the evaluation of an early childhood development program in the Philippines affirmed the strong support for

the program and played a role in the decision to expand its innovative approaches.

Third, in a rapidly changing environment, delay can greatly increase the eventual costs of inaction. If needed actions are not timely, options may disappear entirely. Or if efforts are scaled up and generalized before their actual impacts can be reliably estimated, scarce resources could be wasted on poor choices, missing better opportunities and adding to the burden tomorrow. In an uncertain world, real-time inputs should inform projects, programs, and policies based on the best available evidence. With the right timing, even minor course corrections based on emerging evaluation evidence can have disproportionate impacts. When tens of billions of dollars are being deployed to tackle the global financial crisis or climate change, traditional evaluation cycles are likely to be too slow, losing opportunities for learning and for achieving better results.

The challenge of being timely, then, is to ensure that evaluation work is available when it can influence decision making. The recent global financial crisis put a premium on the speed of using evaluative findings. Based on the estimate from Global Economic Prospects 2010, the crisis had serious cumulative impacts on poverty, with 64 million more people living in extreme poverty by the end of 2010. Aid agencies and national governments committed huge resources to combat this crisis. For example, the lending commitments of the World Bank Group increased substantially, reaching an all-time high of nearly $130 billion during fiscal years 2009 and 2010 combined. Governments also invested heavily in their domestic economies through economic stimulus packages. Evaluators were called on to provide real-time assessments of how these funds were used, and with what results.

A series of briefs and reports on crisis response aimed to provide timely assessments of the Group's activities.[12] The first (discussed further in the next section) pulled together lessons from past crises to inform the Group's response. It identified seven lessons to consider when developing that response. The second, in 2009, provided early findings on the design and implementation of the response. The third, in 2010, provided some preliminary insights into the results of the Group's lending and nonlending response. The fourth, in 2011, is looking in depth at issues of fiscal, financial, and social protection. Each report suggests directions for future work, making the evaluations part of the dialogue on the underlying development issues.

At the same time, national audit and evaluation offices were conducting real-time evaluations of their own countries' domestic crisis response programs. The U.S. Government Accountability Office conducted its own ongoing assessment of the American Recovery and Reinvestment Act, the stimulus package enacted in 2009. Its reports showed where corrections were needed, and many of its recommendations were quickly put into effect. And the U.K. National Audit Office also looked at crisis response issues in a series of reports, again focused primarily on implementation issues, at least initially.[13]

Another real-time evaluation during a crisis is the World Food Programme's assessment of its Southern Africa Emergency Operation in response to the Southern Africa crisis in 2002–2003. The study evaluated the humanitarian relief operations as they unfolded from inception through implementation to closing. Evaluative lessons captured at several stages aimed to promote corporate lending, assess the relevance of the response, and measure the effectiveness of the mode of implementation and the appropriateness of operational policies.[14] IEG has produced rapid-response evaluations providing relevant findings to help address natural disasters, such as the Haiti earthquake, Pakistan floods, and West Africa floods.[15]

Of course, not all real-time evaluation is concerned with an immediate crisis. On climate-related development work, for example, the problems are long term, but the costs of waiting to evaluate until projects are complete are high. IEG recently conducted two phases of evaluation on climate change, and a third is forthcoming.[16] The evaluation stresses five measures that can offer attractive local benefits while fighting climate change: energy efficiency, forest protection, appropriate project finance, technology transfer, and accelerated learning. Carbon finance has yet to realize its promise of catalyzing large-scale new investments in renewable energy and the World Bank and other multilateral development banks need to help clients move away from coal, using analyses that span entire energy systems to find cleaner, more cost-effective, and financeable alternatives.

Prospective Evaluation

Real-time evaluation increases the timeliness and relevance of evaluation by going beyond its ex post, backward-looking roots. But this approach goes only so far, generally focusing on implementation and early results. By contrast, prospective evaluation attempts to use the tools of evaluation to address planning and designing activities from

the start. There are good reasons to expand the scope of evaluation in this way. Evaluation-based program planning provides an opportunity to improve the chances of success by incorporating oversight mechanisms and program features associated with past successes to provide timely corrective feedback on performance.[17]

Of several methods for conducting prospective evaluations, one developed by the U.S. General Accounting Office in the 1980s is the prospective evaluation synthesis.[18] It is intended for use when three conditions are met:

- A new program or approach is being proposed.
- The most effective approach is not known.
- Similar approaches have been tested in the past.

Assuming these conditions, the evaluator first determines what problem the proposal intends to address and ascertains the logic model that specifies how the proposal will address the problem—that is, the mechanics of how the activities contemplated are expected to ameliorate or eliminate the problem. The evaluator also needs to consider the resources needed or expected to be available for the activity.

The evaluator then reviews and synthesizes existing evaluations or other studies to ascertain what evidence is available to provide insights into the likely results of the proposal. This requires a rigorous review of the quality and relevance of the data underlying the studies, weeding out poorly designed or conducted studies and those where the approach, target population, or other conditions are too unlike those applicable to the proposal to provide useful information. Finally, evidence from the remaining studies is analyzed to assess the extent to which it supports the proposal as likely to be an effective response to the problem.

Success using this method depends on the availability of high-quality studies of similar practices with similar populations. Development agency evaluation offices often have a wide range of such evaluations available to them because of frequent use of similar approaches to various problems and the requirements for project and country completion reports. While quality is not uniform, there are likely to be many cases in which prospective evaluation synthesis is feasible. IEG's first report on the financial crisis response, for example, drew on several previous evaluations to derive seven lessons that helped inform the World Bank Group's response.[19] Earlier, in response to a

request from the Bank's Pakistan country team, IEG drew together lessons from natural disasters to inform the response to the October 2005 earthquake in Kashmir.

Another method is scenario analysis. The evaluator assesses the likely effectiveness of proposals under alternative sets of assumptions or scenarios. For example, in work for the European Commission on Impact Assessment, RAND Corporation evaluators typically assess three options: the preferred proposal, a do-nothing alternative, and a more extreme proposal. The three are tested and compared for potential future effectiveness.[20] Another way to carry out such analysis is to test the preferred proposal against a range of alternative plausible assumptions about the future.

One form of scenario analysis is similar to the use of high, medium, and low assumptions typically found in World Bank country assistance strategies. But another is markedly different and has been used in some development work. For example, in 1990 a scenario analysis was conducted in South Africa to help guide the country's transition of power from the National Party to the African National Congress.[21] Four scenarios were constructed and analyzed by a group of economists, business people, academics, politicians, and nongovernmental organization representatives, positing very different kinds of power transfers and posttransition economic policies. In the end there was a consensus on the best course, which helped guide the actual transition.

A third kind of prospective evaluation is policy transfer.[22] Evaluators compare policies adopted in other countries to assess the policies' likely consequences. This involves three stages. The first is awareness, which involves identifying—usually through international policy networks—examples of policies adopted to address the issue of concern. Such information-gathering requires collecting as wide a range of examples as possible, as well as accurate information on the differences among those specific policies and their political, economic, and social contexts. The second is assessment, in which the evaluator seeks to understand the problems and goals the policies were designed to address and how well they performed, as well as how differences in settings could affect their use. And the third is application, when the evaluator considers whether the information from the case studies is actually used in making decisions.

An example is Canada's adoption of a privacy and information policy. Canada looked at models in Sweden and the United States but determined that neither could be adopted directly. The US model

relied extensively on judicial enforcement, which Canada concluded would be inconsistent with its parliamentary system. And it found the Swedish approach too bureaucratic. In the end it developed a middle-ground approach, setting up a commission that could hear appeals based on privacy against government decisions but that did not have licensing authority.

Real-time and prospective evaluation move independent evaluation upstream in the decision-making process, a major shift in perspective that is not without controversy. But as the examples in this section make clear, it is possible for evaluators to maintain their independence and objectivity, even as they apply the tools of evaluation to assessing the likely successes of proposed projects, programs, policies, and other development actions. In today's fast-changing environment, evaluation cannot remain relevant if it limits itself to the analysis of completed work. The demand for evaluative information that can address current decisions is high, but if evaluation does not respond constructively, that demand will find other sources.

Effective Processes Can Increase the Use of Evaluation

To be effective, evaluation, whether prospective, real-time, or ex post, needs to be used. But such use cannot be assumed. It must be earned instead, and that requires work on the part of the evaluator. Several factors contribute to the effectiveness of evaluations through use, including how well it is aimed at the appropriate audiences, the quality of the work, the communication of evaluation findings, the engagement with stakeholders, and the quality of recommendations.

Several other factors affect whether an evaluation is useful to potential users and whether each is actually used—summed up in three questions[23]:

- *Who* will use the evaluation?
- *What* will they need from the evaluation?
- *When* will they need the information?

Typically, there will be several audiences, including those engaged in the development activity itself. But there will also be government officials, donor agencies, and the broader development community. This means that those involved in evaluation need to identify the relevant audiences for the information being produced and to set priorities among the audiences to ensure that the information is aimed most directly at the needs of those most likely to use it. To the extent

possible they should try to get early buy-in from those stakeholders on the relevance and utility of the information to be collected. But in any case they need to be sure to address those needs.

Undigested data usually do not meet this criterion. Frequently, the users expect some analysis of what happened and why. Translating the raw data into an explanation for the outcomes is the role evaluation can and must play. To be effective, however, the messages must be clear and supported by persuasive and comprehensible quantitative or qualitative data. They must also consider the limits in using information from past events to inform decisions on future actions. There are ways to do this, but they require rigorous application of methods that often are new to evaluators.

So, as a first approximation, an effective evaluation process requires that the evaluator understands and addresses the information needs. While necessary, this is not enough. Evaluation also needs to meet quality standards, an obvious but nontrivial consideration.

Evaluation organizations have developed quality standards. For example, the Evaluation Cooperation Group, representing the independent evaluation units of major international finance organizations, has adopted standards for evaluation of public-sector and private-sector projects, policy-based lending, country programs, evaluator independence, and even the evaluation function itself. The evaluation network of the Organisation for Economic Co-operation and Development's Development Assistance Committee[24] and the United Nations Evaluation Group[25] have also promulgated quality standards, as have national associations, such as the American Evaluation Association.[26] These standards provide useful guidance in how evaluations should be conducted, ranging from technical issues of data collection and use to questions of evaluator ethics and behavior.

Adhering to quality standards certainly enhances the usability of evaluation—but, again, doing so does not guarantee its use. Promoting use also requires effective communication. Messages must be tailored to the needs of the targeted audience and demonstrate the salient points with just enough solid evidence to be convincing, without overburdening the decision makers with inessential information.[27] Understanding what information is needed and useful for the decision maker requires delicate judgment, and perhaps experience, on the part of the evaluator.

Timely dissemination of evaluation results can enable new projects to incorporate the lessons into their designs. Publicly disclosed

monitoring of carbon projects shows the gains such feedback can deliver. Landfill gas projects proliferated with the advent of the carbon market, but monitoring reports soon showed that these projects were underperforming relative to their design expectations. This feedback revealed that the appraisal models were based on US experience, inapplicable to the waste streams of developing countries.

For an evaluation to be effective, engaging clients and stakeholders is essential from the design stage, through the evaluation process, and in the formulation and follow-up to recommendations. Some evaluators mistakenly assume that such collaboration reduces independence, but this need not be so. Indeed, the American Evaluation Association's *Guiding Principles* (2004) explicitly recognizes that evaluators are obligated to communicate with clients and stakeholders. But this does not mean that evaluators must take direction from those stakeholders. Instead, such communication is intended to ensure that the evaluator has a firm understanding of the interests and values at stake in the evaluation. Clearly, engagement requires care to ensure against compromising independence. But such communication helps ensure that the stakeholders develop realistic expectations about the evaluation and an understanding of what was done—and why—by the time results are ready to be reported.

This can be especially helpful when results are negative. To explain why intended outcomes and impacts were not achieved, it is essential to know whether lack of results was due to design failure or implementation failure as activities often are not implemented as planned.[28] For example, on safeguard issues, the effectiveness of the regulatory regime depends not only on upfront risk assessments that the countries and multilateral banks carry out but also on effective implementation and supervision and on the checks and balances provided by monitoring and evaluation, disclosure of findings, and verification of results.[29] Engaging the stakeholders throughout the evaluation can provide the evaluator with the information to make this determination and prepare the stakeholders for the findings, which can help in acceptance and responsive action.

Evaluation recommendations need to be formulated and presented in a way likely to move key constituents or stakeholders to take necessary actions to resolve or ameliorate identified problems or shortcomings. This means that recommendations ought to be realistic and easy to understand so that they can be accepted. They should also consider

the context in which they would have to be implemented; otherwise they may not have traction.[30]

It pays to stay engaged with an issue after an evaluation by keeping track of how recommendations have been incorporated. With a one-off engagement at the completion of an evaluation report, the impact of evaluations would be short-lived. Real change takes place only when findings are accepted and people learn. Indeed, weak follow-up on the implementation of results and recommendations and poor dissemination of findings have been identified as key factors explaining why some technical assistance projects of the Asian Development Bank are less successful than others.[31] Similarly, a review of World Bank economic and sector work and technical assistance over 2000–2006 showed that in addition to high technical quality, close collaboration with clients during the process mattered for effectiveness, regardless of whether clients actually produced part of the task.[32]

One mechanism for continuing to stay engaged with an issue—aside from periodic reevaluations in the area—is tracking the implementation of evaluation recommendations. Some organizations do this through management action tracking systems, which review the extent to which major recommendations have been accepted and acted on.[33] Such systems not only prod response from the organization but also act as a reality check for the evaluators on the extent to which their evaluations have been compelling and persuasive to decision makers.

***

These lessons about evaluation and its use in improving effectiveness shed light on the need for evaluators to attend not only to the technical side of their work but also to the needs of the stakeholders, who ultimately decide whether an evaluation's findings, conclusions, and recommendations are used. Without such use, evaluation would remain merely an interesting analytical exercise with limited practical relevance.

Evaluation systems are not yet in place across the development community to ensure their effective use in development decisions and implementation. Evaluators themselves are often lost in their narrow frameworks designed to carry out evaluations, while many policymakers do not see the scope for evaluations to help shape forward-looking decisions. For better results, the mindsets on both sides should change.

Notes

1. Picciotto (2005).
2. Grasso (2003).
3. Patton (2008).
4. Hatry (2007).
5. Grasso et al. (2003).
6. Gressani and Grasso (2010).
7. Wholey (1996).
8. Sandison (2003).
9. Lynam et al. (1999).
10. Petrosino et al. (2000).
11. Levy (2006).
12. IEG (2009k, 2010h).
13. Shipman (2011); Airey (forthcoming).
14. World Food Programme (2003).
15. IEG (2010i).
16. IEG (2010k).
17. Shipman (2011).
18. U.S. General Accounting Office (1990).
19. IEG (2010h).
20. Ling (2001).
21. Maack (2001).
22. Mossberger and Wolman (2003).
23. See note 2.
24. OECD (2010).
25. UNEG (2005).
26. American Evaluation Association (2004).
27. See note 2.
28. Bamberger et al. (2010).
29. IEG (2010f).
30. Hendricks and Handley (1990).
31. Adhikari (2007).
32. World Bank (2008a).
33. IEG (2010e).

10

Conclusion

We keep moving forward, opening new doors, and doing new things, because we're curious and curiosity keeps leading us down new paths.

—Walt Disney

The development community has been trying to shift the center of attention from financing activities toward ensuring that the funding achieves good results. But linking financing and other inputs to the desired development results requires logical steps related to cause and effect. Understanding the assumptions behind such a results chain of development activities is crucial to improving the effectiveness of interventions.

Linking program inputs to outputs may be straightforward, but it is considerably more challenging to link them to outcomes and impacts. To translate actions into results, the underlying assumptions need to be identified and tested. The influence of factors other than the projects and programs, including the role of government and the external environment, needs to be recognized. Identifying the assumptions behind the links in the results chain and understanding the risks to those assumptions are crucial to understanding whether the intervention has caused the outcomes and to what extent it has made a difference.

On many occasions, evaluation confirms existing knowledge and provides quantitative support validating expected results. But sometimes evaluation reveals overlooked links in the development chain connecting actions and results that get short shrift in decision making. Both types of findings add to the body of knowledge.

This book has focused on vital, underemphasized dimensions that should be given more prominence, thereby bringing crucial elements to the attention of policymakers and development practitioners to help them improve development effectiveness. The lessons follow a logical

path—from what results should be the focus of attention, to how the results should be measured, and to how to best use the information from evaluation to strengthen the cycle of learning.

Development work of the countries and international financial institutions sometimes focuses excessively on producing immediate and visible outputs at the expense of addressing the root causes and underlying issues. This can divert attention away from other important concerns and needs, and obstruct improvements in development outcomes. For example, the urgent distribution of emergency supplies is useful in postdisaster situations, but its contributions to improving preparedness and prevention can be limited. Without them, countries prone to disaster can find themselves in a near-permanent state of recovery.

There is also a temptation to attribute results too readily to the actions that were expected to achieve them, without investigating causality or accounting for the influence of other factors or even considering whether these actions were contributing factors and what conditions are required for the planned actions to work. Positive project outcomes do not necessarily translate to positive country outcomes. Factors bearing on country results can be within the control of the government, donor organizations, or other actors—or completely beyond their control. Narrowly attributing development outcomes to a single area runs the risk of undermining the links or even blocking the realization of the benefits that can accrue from multiple sources.

Monitoring and evaluation are valuable for development effectiveness. If appropriate measures are not in place, wrong signals can be sent, and the allocation of resources could be misdirected. What gets measured is sometimes what is easily measured or is consistent with special interests, and claims for validity and reliability may be overstated.

First, composite indicators may not be well defined. Appealing as it seems to compile information on multiple dimensions of development into a single indicator, such composite indicators could be a poor proxy when there is weakness in their premises and methods. Partial coverage without a balanced view could risk weakening relevance. Lumping together several available datasets and assigning them weights without transparent justification can result in "mashup indices." This can lead to mistaken policies and actions that can impair results.

Second, if only the averages are measured, the target population can be left out or wind up worse off, particularly where the distribution

of benefits is skewed toward the better off. Economic growth can accelerate with little or no improvement in income disparities, or worse, with a greater concentration of wealth. Higher average income can exist alongside reduced wealth or income for the poorest segments of a population in relative or even absolute terms. Looking only at the average is risky if the effectiveness of an intervention differs across the subgroups of the population.

Third, achieving intermediate outcomes may be a poor measure of the likelihood of reaching the final, desired results. Although intermediate outcomes are easier to monitor and can reflect changes in a more timely way, they may not lead to the desired final outcomes when elements in the results chain are missing or links are broken. In some important instances, an exclusive focus on intermediate steps (for example, raising enrollment rates) can come at the expense of the desired results (for example, improving learning outcomes), as resources required for learning get diverted.

Context matters a great deal. In a rapidly changing context, simply replicating what has worked may not guarantee what will work. For water, besides the growing constraints on availability, new challenges such as coastal zone management, pollution reduction, and groundwater conservation have emerged. For transport, besides the demand for roads, increasingly vital links with energy, land use, urbanization, the environment, and climate change require cross-cutting approaches. A given investment need not cover all aspects nor one agency do it all, but it is important to connect the dots. Replicating even highly rated projects is beneficial only if the situation continues to warrant such intervention. It is crucial to focus on underlying conditions and adapt to a dynamic situation.

Evaluation's timing and processes can make a big difference. When and how evaluation is conducted determines how useful—and how well used—the information is likely to be for improving development results. Evaluators need to attend not only to the technical side of their work, but also to the needs of the stakeholders who ultimately decide whether an evaluation's findings, conclusions, and recommendations are used. Timely dissemination of evaluation results should improve the design of policies, instruments, and institutions in the face of powerful interest groups.

Lessons drawn in this book have focused on the key links in the results chain. They have brought out specific elements that are often underemphasized but could add up to large aggregate impacts. For

example, promoting natural disaster prevention and mitigation, not just relief and reconstruction, can save thousands of lives and millions of dollars; and focusing on learning rather than enrollment can make investments in education pay. And keeping an eye on the real beneficiaries amid the general average can ensure that services reach their targets.

Taken together, the book's messages suggest gains from the systematic shifts that evaluation can prompt. Immediate and enduring interests need not conflict, but the political economy may favor the immediate. It is worth capitalizing on big cross-effects, but this may not get attention. There may be pressures to measure and present in ways that do not adequately capture the core concerns, but it is beneficial to get these efforts right. It pays to build in lessons learned in a timely manner—but with agility and flexibility to seize opportunities in a context that is constantly changing.

Too often, easy routes are taken while tough questions are avoided. This book does not address the big questions: Why is this so, and how can this change? But some likely factors surface from experience.

Some links connecting actions to results are overlooked and dysfunctional programs are sustained for too long because rent-seeking and political interests negate the accumulation of findings challenging established views and entrenched policies. The disastrous experience with shrimp fisheries in the 1990s was repeated in East Asia in the 2000s, while the cost of destroying mangrove forests was known long before action was taken. The natural defenses in the coastal areas of New Orleans were systematically devastated despite the scientific literature strongly cautioning against such stripping away of coastal wetlands.

People tend to predict the achievement of final results from intermediate outcomes, even in the absence of compelling evidence for any connection. For example, increases in school enrollment and completion rates sometimes are assumed to improve learning. But to connect the dots across enrollment, completion, and learning requires other crucial factors, inside and outside education. Efforts to focus on these interconnections (targeting learning outcomes in this example) would seem to merit support.

Political pressure may favor the status quo. Distributing urgently needed supplies by helicopter and building temporary homes in record time rightly get headlines; equally important prevention efforts, whose benefits accrue only in the future, do not. Building schools gets

politicians elected; stressing learning outcomes does not. Improving services for the upper classes gets credit; focusing on the poor who lack clout does not. Nor does protecting biodiversity: flora and fauna lack a voice at the decision table. And the length of time required for desired outcomes to emerge is often inconsistent with the short horizon of political interests.

The influence of entrenched special interests also tends to prevent needed behavioral and policy shifts. Mollifying powerful interests often means adopting only cosmetic changes and not addressing root causes, possibly leading to further crisis. For sustainable outcomes, we need to support efforts that look at root causes and to follow up on the implications.

Groupthink is another source of inaction. People are often uncomfortable even looking at issues that might challenge conventional wisdom or investigating whether cases that worked in the past would work in the future. Within the "silo" mentality fostered in many organizations, development professionals often have little incentive to look beyond the projects they manage and at the links among the different areas of concern, even though these are often critical to the desired results.

Evaluators are not immune. If evaluative findings seem to confirm existing knowledge, chances are high that no further steps would be taken to drill down to the rationale behind them. The role of identifying missed opportunities in operational work is often underappreciated. Getting outside old assumptions and received wisdom and looking with a fresh eye could shift behavior or even point to entire new ways of doing business. For example, strategies built around reducing energy subsidies and targeting them to the poor could have contributed to a win–win–win scenario by reducing the strain on government budgets, freeing up resources to allow extension of energy sources to the poor, and promoting more efficient energy use.

The role of evaluators can be to promote or question the status quo. For water projects, it is easier to give ratings that motivate the continuation of repeated projects for access to water that have worked in the past, rather than taking on innovative and risky ones that try to address new concerns such as groundwater, water pollution, or coastal zone management issues. Similarly, for transport, it is easier to encourage replicating past success in building roads than tackling new challenges of cross-cutting, multimodal, and environmentally friendly approaches.

Development processes would seem to benefit especially from evaluation findings that highlight vital but underemphasized or overlooked links in the results chain. It would, therefore, pay to have evaluation frameworks that encourage innovation and risk taking with regard to directing attention to these links and enabling the needed follow up.

References

Adeleye, O. A., and A. N. Ofili. "Strengthening Intersectoral Collaboration for Primary Health Care in Developing Countries: Can the Health Sector Play Broader Roles?" *Environmental and Public Health* (2010): doi:10.1155/2010/272896.

Adhikari, R. 2007. *Lessons from the Evaluations of ADB's Technical Assistance Operations*. Manila: Asian Development Bank.

Admassie, A., D. Abebaw, and A. D. Woldemichael. "Impact Evaluation of the Ethiopian Health Services Extension Programme." *Journal of Development Effectiveness* 1, no. 4 (2009): 430–49.

African Development Bank Group. 2011. Evaluation of Paris Declaration Implementation at the African Development Bank. Operations Evaluation Department. May, Tunis.

Airey, P. 2011. "The United Kingdom Response to the Crisis: Evaluating in Real Time," In *High Impact Evaluations: Exploring the Potential of Real Time and Prospective Evaluations*. Washington, DC: Independent Evaluation Group.

Alderman, H. "Improving Nutrition through Community Growth Promotion: Longitudinal Study of the Nutrition and Early Child Development Program in Uganda." *World Development* 35, no. 8 (2007): 1376–89.

Alves, D. S. "Space-Time Dynamics of Deforestation in Brazilian Amazônia." *International Journal of Remote Sensing* 23, no. 14 (2002): 2903–08.

American Evaluation Association. 2004. "Guiding Principles for Evaluators." American Evaluation Association, Fairhaven, MA. www.eval.org/publications/aea06.GPBrochure.pdf (accessed September 15, 2011).

Araujo, C. M., F. Ferreira, P. Lanjouw, and B. Özler. 2005. "Local Inequality and Project Choice: Theory and Evidence from Ecuador." Policy Research Working Paper 3997. World Bank, Washington, DC.

Asian Development Bank. 2005. "Bangladesh: Jamuna Bridge Railway Link Project." Asian Development Bank Project Completion Report. Asian Development Bank. Manila.

———. 2009a. "Natural Catastrophe Risk Insurance Mechanisms for Asia and the Pacific." Based on Conference by ADB and the Ministry of Finance, Government of Japan. Asian Development Bank. Manila.

———. 2009b. *Evaluation Knowledge Brief: Greenhouse Gas Implications of ADB's Energy Sector Operations*. Manila: Asian Development Bank.

———. 2011. *Equity and Well-Being: Measurement and Policy Practice. November*. Manila: Asian Development Bank.

Bamberger, M., V. Rao, and M. Woolcock. 2010. "Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development." Policy Research Working Paper 5245. World Bank, Washington, DC.

Banerjee, A. 2010. *Making Sense of Microfinance*. Cambridge, MA: Massachusetts Institute of Technology.

Banerjee, A., and E. Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs.

Banerjee, A., E. Duflo, R. Glennerster, and D. Kothari. 2010. "Improving Immunization Coverage in Rural India: A Clustered Randomized Controlled Evaluation of Immunization Campaigns with and without Incentives." Innovation for Poverty Action, New Haven, CT. www.poverty-action.org/provenimpact/vaccines/challenge (accessed September 15, 2011).

Banerjee, A., S. Jacob, M. Kremer, J. Lanjouw, and P. Lanjouw. 2001. *Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials*. Cambridge, MA: Harvard University and Massachusetts Institute of Technology.

Barrera-Osorio, F., and L. Linden. 2009. "The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia." Policy Research Working Paper 4836. World Bank, Washington, DC.

Bekefi, T. 2006. *The Global Road Safety Partnership and Lessons in Multisectoral Collaboration*. Cambridge, MA: Harvard University, John F. Kennedy School of Government.

Bern, C., J. Sniezek, G. M. Mathbor, M. S. Siddiqi, C. Ronsmans, A. M. Chowdhury, A. E. Choudhury, K. Islam, M. Bennish, and E. Noji. "Risk Factors for Mortality in the Bangladesh Cyclone of 1991." *Bulletin of the World Health Organization* 71, no. 1 (1993): 73–78.

Birdsall, N., and H. Kharas 2010. *Quality of Official Development Assistance Assessment Report*. Washington, DC: Center for Global Development.

Birdsall, N., and W. D. Savedoff. 2010. *Cash on Delivery: A New Approach to Foreign Aid*. Washington, DC: Center for Global Development.

Cadena, X., and A. Schoar. 2011. *Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments*. NBER Working Paper 17020. Cambridge, MA: National Bureau of Economic Research.

Cairncross, S., and V. Valdmanis. 2006. "Water Supply, Sanitation, and Hygiene Promotion." In *Disease Control Priorities in Developing Countries*, 2nd ed., ed. D. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, A. Mills, and P. Musgrove. New York: Oxford University Press and World Bank.

Castro, M. F. 2008. *Insider Insights: Building a Results-Based Management and Evaluation System in Colombia*. ECD Working Paper 18. Washington, DC: Independent Evaluation Group.

Castro, M. F., G. Lopez-Acevedo, G. B. Busjeet, and F. Ordonez. 2009. "Mexico's M&E System: Scaling Up from the Sectoral to the National Level." Evaluation Capacity Development Working Paper 20. Independent Evaluation Group, Washington, DC.

Center for Research on Epidemiology of Disasters. 2011. International Disaster Database (CRED EM-DAT database). www.emdat.be/database (accessed July 10, 2011).

Coady, D., M. Grosh., and J. Hoddinott. "Targeting Outcomes Redux." *World Bank Research Observer* 19, no. 1 (2004): 61–85.

Commonwealth Secretariat/World Bank Joint Task Force on Small States. 2000. "Small States: Meeting Challenges in the Global Economy." World Bank, Washington, DC.

Cuny, F. 1983. *Disasters and Development*. Oxford: Oxford University Press.

Daniels, D. L., S. N. Cousens, L. N. Makoae, and R. G. Feachem. "A Case-Control Study of the Impact of Improved Sanitation on Diarrhea Morbidity in Lesotho." *Bulletin of the World Health Organization* 68, no. 4 (1990): 455–63.

Dasgupta, P. 2007. *Economics: A Very Short Introduction*. Oxford: Oxford University Press.

———. "Discounting Climate Change." *Journal of Risk and Uncertainty* 37 (2008): 141–69.

De Mel, S., D. McKenzie, and C. Woodruff. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *Quarterly Journal of Economics* 123, no. 4 (2008): 1329–72.

Devarajan, S. 2010. "Is Our Tanzanian Children Learning?" World Bank blog: http://blogs.worldbank.org/africacan/is-our-tanzanian-children-learning (accessed October 12, 2011).

DeVellis, R. F. 2003. *Scale Development: Theory and Applications*, 2nd ed. Thousand Oaks, CA: Sage Publications.

Development Assistance Committee. 2010. *Quality Standards for Development Evaluation*. Paris: Organisation for Economic Co-operation and Development.

DfID (UK Department for International Development). 2004. *Integrated Coastal Zone Management in Bangladesh: A Policy Review.* Livelihood-Policy Relationships in South Asia. Working Paper 1. London: DfID.

———. 2008. *Water: An Increasingly Precious Resource.* Sanitation: A Matter of Dignity. London: DfID.

Dollar, D., M. Hallward-Dreimeier, and T. Mengistae. "Investment Climate and International Integration." *World Development* 34, no. 9 (2006): 1498–516.

Duflo, E., P. Dupas, and M. Kremer. 2010. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." J-PAL working paper. Massachusetts Institute of Technology, Poverty Action Lab, Cambridge, MA. www.povertyactionlab.org/publication/peer-effects-teacher-incentives-and-impact-tracking-evidence-randomized-evaluation-kenya (accessed September 15, 2011).

ECG (Evaluation Cooperation Group). 2010. *Good Practice Standards on Independence of International Financial Institutions' Central Evaluation Departments.* Manila: ECG. https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/h_9BD8546FB7A652C948257731002A062B/AA95B62CF943F82E4825774B003A82E0/?OpenDocument&Form=h_PageUI (accessed September 15, 2011).

———. 2011. *Evaluative Lessons for Agriculture and Agribusiness*. Evaluation Cooperation Group Paper 3. Washington, DC: World Bank.

Eifert, B. P. 2007. "The Economic Response to Regulatory Reform—2003–06." Draft, September 21. University of California, Berkeley, CA.

Elbers, C., P. Lanjouw, J. Mistiaen, B. Özler, and K. Simler. "On the Unequal Inequality of Poor Communities." *World Bank Economic Review* 18, no. 3 (2004): 401–21.

Esrey, S. A., and J. -P. Habicht. 1985. *The Impact of Improved Water Supplies and Excreta Disposal Facilities on Diarrheal Morbidity, Growth, and Mortality among Children*. Cornell International Nutrition Monograph Series 15. Ithaca, NY: Division of Nutritional Sciences, Cornell University.

European Bank for Reconstruction and Development. 2010. Business Environment and Enterprise Performance Survey. European Bank for Reconstruction and Development, London.

Fan, V. Y. -M., and A. Mahal. "What Prevents Child Diarrhea? The Impacts of Water Supply, Toilets, and Hand-Washing in Rural India." *Journal of Development Effectiveness* 3, no. 3 (2011): 340–70.

Field, E., R. Pande, J. Papp, and N. Rigol. 2010. "Term Structure of Debt and Entrepreneurial Behavior: Experimental Evidence from Microfinance." Harvard University, Cambridge, MA. www.economics.harvard.edu/faculty/field/files/repayment_default_Dec19.pdf (accessed September 15, 2011).

Freeman, P., and K. Mathur. 2008. *The Health Benefits of Transport Projects: A Review of the World Bank Transport Sector Lending Portfolio.* IEG Working Paper 2. Washington, DC: Independent Evaluation Group.

Gaarder, M. M., A. Glassman, and J. E. Todd. "Conditional Cash Transfers and Health: Unpacking the Causal Chain." *Journal of Development Effectiveness* 2, no. 1 (2010): 6–50.

Gakidou, E., K. Cowling, R. Lozano, and C. Murray. "Increased Educational Attainment and Its Effect on Child Mortality in 175 Countries between 1970 and 2009: A Systematic Analysis." *Lancet* 376, no. 9745 (2010): 959–74.

Galasso, E., and N. Umapathi. "Improving Nutritional Status through Behavioural Change: Lessons from Madagascar." *Journal of Development Effectiveness* 1, no. 1 (2009): 60–85.

Gauthier, B., and M. Gersovitz. "Revenue Erosion through Exemption and Evasion in Cameroon, 1993." *Journal of Public Economics* 64, no. 3 (1997): 407–24.

Giles, J., and E. Satriawan. 2010. "Protecting Child Nutritional Status in the Aftermath of a Financial Crisis: Evidence from Indonesia." Policy Research Working Paper 5471. World Bank, Washington, DC.

Gonzales, F., K. Dearden, and W. Jimenez. "Do Multisectoral Development Programmes Affect Health? A Bolivian Case Study." *Health Policy and Planning* 14, no. 4 (1999): 400–408.

Grasso, P. G. "What Makes an Evaluation Useful? Reflections from Experience in Large Organizations." *American Journal of Evaluation* 24 (2003): 507–14.

Grasso, P. G., S. Sulaiman, S. Wasty, and R.V. Weaving, eds. 2003. *World Bank Operations Evaluation Department: The First 30 Years.* Washington, DC: World Bank.

Gressani, D., and P. G. Grasso. 2010. "New Directions in Development Evaluation." In *Vision for Development: Dialogues on the Work of Vinod Thomas*, ed. R. C. Rist. Washington, DC: World Bank.

Guillaumont, P. 2009. *Caught in a Trap: Identifying the Least Developed Countries.* London: Economica.

Hallward-Driemeier, M., and L. Pritchett. 2011. "How Business Is Done and the 'Doing Business' Indicators: The Investment Climate When Firms

Have Climate Control." Policy Research Working Paper 5563. World Bank, Washington, DC.
Hatry, H. P. 2007. *Performance Measurement*, 2nd ed. Washington, DC: Urban Institute.
Hendricks, M., and E. Handley. "Improving the Recommendations from Evaluation Studies." *Evaluation and Program Planning* 13, no. 2 (1990): 109–17.
Hoddinott, J., and B. Kinsey. "Child Growth in the Time of Drought." *Oxford Bulletin of Economics and Statistics* 63, no. 4 (2001): 409–36.
Høyland, B., K. O. Moene, and F. Willumsen. 2010. "The Tyranny of International Index Rankings." Department of Economics, University of Oslo.
Huttly, S. R. A., S. S. Morris, and V. Pisani. "Prevention of Diarrhea in Young Children in Developing Countries." *Bulletin of the World Health Organization* 75, no. 2 (1997): 163–74.
IEA (International Energy Agency). 2007. *World Energy Outlook 2006.* Paris: IEA.
Independent Evaluation Department. 2011. Learning Lessons. ADB's Emergency and Rehabilitation Assistance on Flood-Affected Areas. Asian Development Bank. Manila.
Independent Evaluation Department. 2006. Performance Evaluation Report. Bangladesh: Flood Damage Rehabilitation Project. Asian Development Bank. Manila.
IEG (Independent Evaluation Group). 2000. "Jamuna Bridge Project." Implementation Completion Report Review. World Bank, Washington, DC.
——— 2002. *Social Funds: Assessing Effectiveness*. Washington, DC: World Bank.
——— 2005. *The Effectiveness of World Bank Support for Community-Based and Driven Development.* Washington, DC: World Bank.
——— 2006a. *From Schooling Access to Learning Outcomes: An Unfinished Agenda*. Washington, DC: World Bank.
——— 2006b. *Hazards of Nature, Risks to Development: An IEG Evaluation of World Bank Assistance for Natural Disasters.* Washington, DC: World Bank.
——— 2007a. *A Decade of Action in Transport: An Evaluation of World Bank Assistance to the Transport Sector, 1995–2005*. Washington, DC: World Bank.
——— 2007b. *The Development Potential of Regional Programs: An Evaluation of World Bank Support of Multi-country Operations*. Washington, DC: World Bank.
——— 2008a. *Doing Business: An Independent Evaluation: Taking the Measure of the World Bank-IFC Doing Business Indicator*. Washington, DC: World Bank.
——— 2008b. *IEG Impact Evaluation—The Experience of the Independent Evaluation Group of the World Bank*. Washington, DC: World Bank.
——— 2008c. *Public Sector Reform: What Works and Why—An IEG Evaluation of World Bank Support.* Washington, DC: World Bank.
——— 2008d. *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits: An IEG Impact Evaluation.* Washington, DC: World Bank.

——— 2009a. "*Annual Review of Development Effectiveness: Achieving Sustainable Development*." Washington, DC: World Bank.
——— 2009b. *Climate Change and the World Bank Group: Phase I: An Evaluation of World Bank Win-Win Energy Policy Reforms.* Washington, DC: World Bank.
——— 2009c. *Improving Effectiveness and Outcomes for the Poor in Health, Nutrition, and Population.* Washington, DC: World Bank.
——— 2009d. *Independent Evaluation of MIGA's Development Effectiveness 2009: Enhancing MIGA's Risk Mitigation in IDA and Conflict-Affected Countries.* Washington, DC: World Bank.
——— 2009e. "Morocco: Lakhdar Watershed Management Pilot Project." Project Performance Assessment Report. World Bank, Washington, DC.
——— 2009f. "Protected Area Effectiveness in Reducing Tropical Deforestation: A Global Analysis of the Impact of Protection Status." Evaluation Brief #4. World Bank, Washington, DC.
——— 2009g. *The World Bank's Country Policy and Institutional Assessment: An Evaluation.* Washington, DC: World Bank.
——— 2009h. *The World Bank Group Guarantee Instruments 1990–2007: An Independent Evaluation.* Washington, DC: World Bank.
——— 2009i. "The World Bank Group Program of Support for the Chad Cameroon petroleum Development and Pipeline construction." Project Performance Assessment Report. World Bank, Washington, DC.
——— 2009j. "Vietnam: Irrigation Rehabilitation Project, Water Supply Project, Inland Waterway and Port Rehabilitation Project, Coastal Wetlands Protection Project." Project Performance Assessment Report. World Bank, Washington, DC.
——— 2010a. *Cost-Benefit Analysis in World Bank Projects*. Washington, DC: World Bank.
——— 2010b. *Do Conditional Cash Transfers Lead to Medium-Term Impacts? Evidence from a Female School Stipend Program in Pakistan*. Washington, DC: IEG.
——— 2010c. *Evaluative Lessons from World Bank Group Experience: Growth and Productivity in Agriculture and Agribusiness*. Washington, DC: World Bank.
——— 2010d. "Priorities in Meeting the MDGs: Lessons from Evaluation." IEG, Washington, DC.
——— 2010e. *Results and Performance of the World Bank Group: IEG Annual Report.* Washington, DC: World Bank.
——— 2010f. *Safeguards and Sustainability Policies in a Changing World: An Independent Evaluation of World Bank Group Experience.* Washington, DC: World Bank.
——— 2010g. *Tanzania Primary Education Development Program Project, Secondary Education Development Program Project, Human Resources Development Pilot Project.* Washington, DC: World Bank.
——— 2010h. *The World Bank Group's Response to Global Economic Crisis (Phase I).* Washington, DC: World Bank.
——— 2010i. *Water and Development. An Evaluation of World Bank Support, 1997–2007.* Washington, DC: World Bank.
——— 2010j. "WBG Response to the Haiti Earthquake: Evaluative Lessons." IEG, Washington, DC. http://siteresources.worldbank.org/INTOED/Resources/disaster_note.pdf (accessed September 15, 2011).

——— 2010k. *What Can We Learn from Nutrition Impact Evaluations?* Washington, DC: World Bank.
——— 2011a. *A Portfolio note: World Bank Support to Education since 2001.* Washington, DC: World Bank.
——— 2011b. *An Evaluation of World Bank Support, 2000–2010: Social Safety Nets*. Washington, DC: World Bank.
——— 2011c. *Results and Performance of the World Bank Group: IEG Annual Report 2011*. Washington, DC: World Bank.
Independent Evaluation Office. 2011. "IMF Performance in the Run-Up to the Financial and Economic Crisis: IMF Surveillance in 2004–07." IMF Evaluation Report. International Monetary Fund, Washington, DC.
Inter-American Development Bank. 2011. *Development Effectiveness Overview*. Washington, DC.
International Federation of Red Cross and Red Crescent Societies. 2001. *World Disasters Report 2001: Focus on Recovery.* Geneva: International Federation of Red Cross.
——— 2010. *World Disasters Report 2010: Focus on Urban Risk*. Geneva: International Federation of Red Cross.
Jalan, J., and M. Ravallion. "Does Piped Water Reduce Diarrhea for Children in Rural India?" *Journal of Econometrics* 112 (2003): 153–73.
Jameel Poverty Action Lab. 2009. "Research and Practice: Learning and Doing Together, A Case Study of Seva Mandir and Poverty Action Lab." Background document for PEGNet Conference 2009, "Policies for Reducing Inequality in the Developing World," Institute of Social Studies, The Hague, September 3–4.
Jha, A. 2011. "Time to Wake Up to Disaster Prevention, Asia." World Bank Climate Change Blog: http://climatechange.worldbank.org/blogs/time-wake-disaster-prevention-asia (accessed October 10).
Jorgensen, S., M. Grosh, and M. Schacter. 1991. *Easing the Poor through Economic Crisis and Adjustment: The Story of Bolivia's Emergency Social Fund*. Washington, DC: World Bank.
Jorgensen, S., and J. Van Domelen 1999. "Helping the Poor Manage Risk Better: The Role of Social Funds." Presented at IADB Conference on Social Protection and Poverty, Inter-American Development Bank, Washington, DC, February 4–5.
Kanbur, R., and A. Sumner. 2011. "Poor Countries or Poor People? Development Assistance and the New Geography of Global Poverty." Cornell University, Ithaca, NY. http://kanbur.dyson.cornell.edu/papers/KanburSumnerPoorCountriesOrPoorPeople.pdf (accessed September 15, 2011).
Karlan, D.S., and J. Appel. 2011. *More Than Good Intentions: How a New Economics is Helping to Solve Global Poverty*. Boston, MA: Dutton Press.
Karlan, D. S., and M. Valdivia. 2008. "Teaching Entrepreneurship: Impact Of Business Training On Microfinance Clients and Institutions." Yale University, New Haven, CT.
Karlan, D. S., and J. Zinman. 2009. "Expanding Microenterprise Credit Access: Randomized Supply Decisions to Estimate the Impacts in Manila." Working Papers 68. Yale University, Department of Economics, New Haven, CT.
Keutsch, G. T., O. Fontaine, A. Bhargava, C. Boschi-Pinto, Z. A. Bhutta, E. Gotuzzo, J. Rivera, J. Chow, A. Shahid-Salles, and R. Laxminarayan. 2006.

"Diarrheal Diseases." In *Disease Control Priorities in Developing Countries*, 2nd ed., ed. D. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, A. Mills, and P. Musgrove. New York: Oxford University Press and World Bank.

Kohli, H. S., A. Sharma, A. Sood, eds. 2011. *Asia 2050: Realizing the Asian Century*. Manila: Sage Publications for Asian Development Bank.

Kopp, A. Forthcoming. "More Efficient Transport Will Advance Development and Protect the Environment." World Bank, Washington, DC.

Krug, E., and T. Toroyan. "The United Nations and Road Safety: What is the Connection?" *Journal of Epidemiology and Community Health* 60, no. 12 (2006): 1009.

Kumar, S., and S. Corbridge. "Programmed to Fail? Development Projects and the Politics of Participation." *Journal of Development Studies* 39, no. 2 (2002): 830–51.

Lall, U., T. Heikkila, C. Brown, and T. Siegfried. "Water in the 21st Century: Defining the Elements of Global Crises and Potential Solutions." *Journal of International Affairs* 61, no. 2 (2008): 1–22.

Levy, S. 2006. *Progress Against Poverty: Sustaining Mexico's Progresa-Oportunidades Program.* Washington, DC: Brookings Institution Press.

Lin, J., and C. Monga. 2010. "The growth report and new structural economics." Policy Research Working Paper 5336. World Bank, Washington, DC.

Ling, T. 2011. *Evaluating in Uncertain Environments: Prospective Evaluation and Scenario Building in High Impact Evaluations: Exploring the Potential of Real Time and Prospective Evaluations*. Washington, DC: World Bank.

Linnerooth-Bayer, J., and R. Mechler. 2005. "Disaster Risk Financing for Developing Countries." IIASA Presentation for the World Bank. International Institute for Applied Systems Analysis, Washington, DC, November 17.

Luppino, M., G. Gajewski, S. Zohir, B. Khondker, and D. Crowther. 2004. "Estimating the Impacts of the Jamuna Bridge on Poverty Levels in Bangladesh Using SAM and CGE Models: A Comparative Study." Louis Berger Group, Washington, DC.

Lynam, D. R., R. Milich., R. Zimmerman, S. P. Novak, T. K. Logan, C. Martin, C. Leukefeld, and R. Clayton. "Project DARE: No Effects at 10-Year Follow-Up." *Journal of Consulting Clinical Psychology* 67, no. 4 (1999): 590–93.

Maack, J. N. 2001. "Scenario Analysis: A Tool for Task Managers." In *Social Analysis: Selected Tools and Techniques*, ed. R. J. Krueger, M. A. Casey, J. Donner, S. Kirsch, and J. N. Maack. Washington, DC: World Bank.

Mackay, K. 2006. *Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management.* ECD Working Paper 15. Washington, DC: Independent Evaluation Group.

Margulis, S., and U. Narain. 2010. *The Costs to Developing Countries of Adapting to Climate Change: New Methods and Estimates—The Global Report of the Economics of Adaptation to Climate Change Study*. Washington, DC: World Bank.

Min, S. -K., X. Zhang, F. W. Zwiers, and G. C. Hegerl. "Human Contribution to More-Intense Precipitation Extremes." *Nature* 470 (2011): 378–81.

Morgan, T. 2007. "Energy Subsidies: Their Magnitude, How They Affect Energy Investment and Greenhouse Gas Emissions, and Prospects for Reform." UNFCCC Financial and Technical Support Programme, Bonn.

Morra-Imas, L. 2011. "Embracing Evaluation in Tough Economic Times." In *Evaluation in Stable and Turbulent Times*, ed. J. -E. Furubo, R. C. Rist, and S. Speer. Piscataway, NJ: Transaction Publishers.

Morra-Imas, L., and R.C. Rist. 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations.* Washington, DC: World Bank.

Mossberger, K., and H. Wolman. "Policy Transfer as a Form of Prospective Policy Evaluation: Challenges and Recommendations." *Public Administration Review* 63, no. 4 (2003): 428–40.

Mullainathan, S., M. Bertrand, and S. Djankov. 2006. "Obtaining a Driving License in India: An Experimental Approach to Studying Corruption." WCFIA Working Paper, November. Weatherhead Center for International Affairs, Harvard University, Cambridge, MA.

Multi-hazard Mitigation Council for the Federal Emergency Management Agency. 2010. http://disastersafety.org/conference(accessed July 10, 2011).

Münchener Rückversicherungs-Gesellschaft. 1975. *Fifi, A Study of the Hurricane Damage in Honduras.* Munich: Münchener Rückversicherungs-Gesellschaft.

National Population Commission. 2004. *Nigeria Demographic and Health Survey 2003*. Calverton, MD: National Population Commission and ORC Macro.

Nelson, A., and K. Chomitz. forthcoming. "Effectiveness of Strict vs. Multiple Use Protected Areas in Reducing Tropical Forest Fires: A Global Analysis using Matching Methods." *PLoS ONE*.

OECD (Organisation for Economic Co-operation and Development). 2009. "Managing for Development Results." Policy Brief. OECD, Paris.

——— 2010. *Evaluation in Development Agencies.* Better Aid. Paris: OECD.

Olken, B. 2006. "Political Institutions and Local Public Goods: Evidence from a Field Experiment in Indonesia." Asian Development Bank, Operations Evaluation Department, Manila.

Overbey, L. 2008. *The Health Benefits of Water Supply and Sanitation Projects: A Review of the World Bank Lending Portfolio.* Washington, DC: World Bank.

Pall, P., T. Aina, D. A. Stone, P. A. Stott, T. Nozawa, A. G. J. Hilberts, D. Lohmann, and M. R. Allen. "Anthropogenic Greenhouse Gas Contribution to Flood Risk in England and Wales in Autumn 2000." *Nature* 470 (2011): 382–85. www.nature.com/nature/journal/v470/n7334/full/nature09762.html.

Patton, M. Q. 2008. *Utilization-Focused Evaluation*, 4th ed. Thousand Oaks, CA: Sage Publications.

Petrosino, A., C. Turpin-Petrosino, and J. Finckenauer. "Well-Meaning Programs Can Have Harmful Effects! Lessons from Experiments of Programs Such as Scared Straight." *Crime Delinquency* 46, no. 3 (2000): 354–79.

Picciotto, R. "The Evaluation of Policy Coherence for Development." *Evaluation* 11, no. 3 (2005): 311–30.

——— "The New Environment for Development Evaluation." *American Journal of Evaluation* 28, no. 4 (2007): 509–21.

Primavera, J. H.. "Mangroves, Fishponds, and the Quest for Sustainability." *Science* 310 (2005): 57–59.

Pritchett, L., and G. Sethi. "Tariff Rates, Tariff Revenue, and Tariff Reform: Some New Facts." *World Bank Economic Review* 8, no. 1 (1994): 1–16.

Quisumbing, A. R. 2003. "Food Aid and Child Nutrition in Rural Ethiopia." FCND Discussion Papers 158. International Food Policy Research Institute, Washington, DC.
Rao, V., and A. Ibanez. "The Social Impact of Social Funds in Jamaica: A 'Participatory Econometric' Analysis of Targeting, Collective Action, and Participation in Community-Driven Development." *Journal of Development Studies* 41, no. 5 (2005): 788–838.
Ravallion, M. 2010a. "Mashup Indices of Development." Policy Research Working Paper 5432. World Bank, Washington, DC.
——— 2010b. "Troubling Tradeoffs in the Human Development Index." Policy Research Working Paper 5484. World Bank, Washington, DC.
——— 2011. "On Multidimensional Indices of Poverty." Policy Research Working Paper 5580. World Bank, Washington, DC.
Rawlings, L. B., L. Sherburne-Benz, and J. Van Domelen. 2004. *Evaluating Social Funds: A Cross-Country Analysis of Community Investments.* Washington, DC: World Bank.
Redwood, J. 1993. *World Bank Approaches to the Environment in Brazil: A Review of Selected Projects*, vol. V. Washington, DC: World Bank.
Rist, R. C., M. -H. Boily, and F. Martin. 2011. *Influencing Change: Building Evaluation Capacity to Strengthen Governance*. Washington, DC: World Bank.
Rodrigue, J. -P., C. Comtois, and B. Slack. 2009. *The Geography of Transport Systems*, 2nd ed. New York: Routledge.
Rogers, P. "Facing the Freshwater Crisis." *Scientific American* July 23, 2008.
Rogoff, K. "Equilibrium Political Budget Cycles." *American Economic Review* 80 (1990): 21–36.
Sachs, J. 2008. *Common Wealth: Economics for a Crowded Planet.* New York: Penguin Press.
Sandison, P. 2003. *Desk Review of Real-Time Evaluation Experience*. New York: United Nations Children's Fund.
Schady, N. R. "Picking the Poor: Indicators for Geographic Targeting in Peru." *Review of Income and Wealth* 48, no. 3 (2002): 417–33.
Schuknecht, L. "Political Business Cycles and Fiscal Policies in Developing Countries." *Kyklos* 49, no. 2 (1996): 155–70.
Shipman, S. 2011. "Real Time and Prospective Evaluation in Practice: The Experience of the U.S. Government Accountability Office." U.S. Government Accountability Office, Washington, DC.
Stott, P. A., D. A. Stone, and M. R. Allen. "Human Contribution to the European Heatwave of 2003." *Nature* 432 (2004): 610–14. www.nature.com/nature/journal/v432/n7017/full/nature03089.html.
Swearingen, W. D. 1987. *Moroccan Mirages: Agrarian Dreams and Deceptions, 1912–1986.* London: Princeton University Press.
The Citizen. 2010. "Tanzania Drive to Improve Education Lauded." www.tz.undp.org/docs/Tanzania_education.pdf (accessed September 21).
The 2030 Water Resources Group. 2009. *Charting Our Water Future: Economic Frameworks to Inform Decision-Making*. Washington, DC: McKinsey & Company.
Thomas, V. 2011. *It Is Time to Factor Natural Disasters into Macroeconomic Scenarios*. Economic Premise. Washington, DC: World Bank.

Thomas, V., and J. Tominaga. "Evaluation for Better Development Results." *Journal of Development Effectiveness* 2, no. 3 (2010): 371–86.
Turton, C., and J. Farrington. 1998. "Enhancing Rural Livelihoods through Participatory Watershed Development in India." Natural Resource Perspectives 34. London: Overseas Development Institute.
UK 8020. 2011. "Climate Change Challenge: Causes of Climate Change." United Kingdom. www.climatechangechallenge.org/Resource%20Centre/Climate-Change/3-what_causes_climate_change.htm#5 (accessed July 10, 2011).
UK House of Commons, International Development Committee. 2010. "Appointment of the Chief Commissioner of the Independent Commission for Aid Impact." House of Commons, London.
UN (United Nations). 2009. *2009 Global Assessment Report on Disaster Risk Reduction.* New York: United Nations.
——— 2011. *2011 Global Assessment Report on Disaster Risk Reduction*. New York: United Nations.
UN Daily News. 2011. "End Violence and Free Those Detained, UN Rights Chief Tells Egypt." Issue DH/5832, February 4, 2011. New York. www.un.org/news/dh/pdf/english/2011/04022011.pdf (accessed July 10, 2011).
UN Habitat (United Nations Centre for Human Settlements). 2001. *Cities in a Globalizing World: Global Report on Human Settlements.* London: Earthscan.
UNDP (United Nations Development Programme). 2006. *Human Development Report: Beyond Scarcity.* New York: UNDP.
UNEG (United Nations Evaluation Group). 2005. *Standards for Evaluation in the UN System.* New York: UNEG.
——— 2009. "UNEG Principles of Working Together." UNEG, New York.
UNICEF (United Nations Children's Fund). (n.d.) "Child Info, Statistics by Area/Education." www.childinfo.org/education.html (accessed July 10, 2011).
USAID (United States Agency for International Development). 2010. "USAID Evaluation Policy." USAID, Washington, DC.
U.S. Census Bureau. 2001. "Time Series of National Population Estimates: April 1, 2000 to July 1, 2001." U.S. Census Bureau, Population Division, Washington, DC.
U.S. General Accounting Office. 1990. "The Evaluation Synthesis." U.S. General Accounting Office, Washington, DC.
Vledder, M., J. Friedman, M. Sjoblom, and P. Yadav. 2010. "Enhancing Public Supply Chain Management in Zambia." World Bank, Washington, DC.
White, H. "Theory-Based Impact Evaluation: Principles and Practice." *Journal of Development Effectiveness* 1, no. 3 (2009): 271–84.
White, H., and E. Masset. "Assessing Interventions to Improve Child Nutrition: A Theory-Based Impact Evaluation of the Bangladesh Integrated Nutrition Project." *Journal of International Development* 19, no. 5 (2007): 627–52.
WHO (World Health Organization). 2004. *Family and Community Practices That Promote Child Survival, Growth, and Development—A Review of Evidence.* Geneva: WHO.
——— 2008. "World Health Statistics." Geneva. www.who.int/whosis/whostat/2008/en/index.html (accessed September 15, 2011).
——— 2009. *Global Status Report on Road Safety: Time for Action*. Geneva: WHO.

Wholey, J. "Formative and Summative Evaluation: Related Issues in Performance." *American Journal of Evaluation* 17, no. 2 (1996): 145–49.

World Bank. 1988. *Road Deterioration in Developing Countries. Causes and Remedies.* A World Bank Policy Study. Washington, DC: World Bank.

——— 1997. "FISE-I Nicaragua." Implementation Completion Report. Latin America and the Caribbean Department, Human Development Sector Management Unit. World Bank, Washington, DC.

——— 2001. *World Development Indicators*. CD-ROM. Washington, DC: World Bank.

——— 2006. "Country Partnership Strategy Partnership Report for the People's Republic of China for the Period 2006–2010." World Bank, Washington, DC.

——— 2007a. "Restoring Tarim River Basins Lifts Chinese Farmers Out of Poverty." Mimeo, World Bank. http://go.worldbank.org/1D3CH1FPI0 (accessed September 15, 2011).

——— 2007b. *World Development Report 2007: Development and the Next Generation*. Washington, DC: World Bank.

——— 2008a. *Using Knowledge to Improve Development Effectiveness: An Evaluation of World Bank Economic and Sector Work and Technical Assistance, 2000–2006*. Washington, DC: World Bank.

——— 2008b. *Vietnam's Mangrove Forests Fight Erosion, Improve Livelihoods*. http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:21782087~menuPK:34457~pagePK:34370~piPK:34424~theSitePK:4607,00.html (accessed September 15, 2011).

——— 2009. *Global Monitoring Report 2009: A Development Emergency.* Washington, DC: World Bank.

——— 2010a. "Discovering What Works in Education: Informed Policymaking through Impact Evaluation." World Bank, Washington, DC.

——— 2010b. *EACC study, The Costs to Developing Countries of Adapting to Climate Change.* Washington, DC: World Bank.

——— 2010c. *Economics of Adaptation to Climate Change: Global Report*. Washington, DC: World Bank.

——— 2010d. *Evolving Education Strategy 2020 Emphasizes Learning for All*. Washington, DC: World Bank.

——— 2010e. *Natural Hazards, Unnatural Disasters.* Washington, DC: World Bank.

——— 2010f. *The World Bank Group Mid-Cycle Implementation Progress Report: Sustaining Water for All in a Changing Climate*. Washington, DC: World Bank.

——— 2010g. *World Development Indicators 2010*. Washington, DC: World Bank.

——— 2011a. *Develop Strategic Plan and Roadmap for World Bank Road Safety Activities in China*. Project Concept Note. Washington, DC: World Bank.

——— 2011b. *Education Strategy for 2020: Learning for All*. Washington, DC: World Bank.

——— 2011c. "The Recent Earthquake and Tsunami in Japan: Implications for East Asia." Supplemental note for East Asia and Pacific Region Bi-annual Update, March 2011. World Bank, Washington, DC.

———2011d. "Turkey—Seismic Risk Mitigation Project: P078359—Implementation Status Results Report: Sequence 16." World Bank, Washington, DC.
World Food Programme. 2003. "Summary Report of the Real-Time Evaluation of WFP's Response to the Southern Africa Crisis, 2002–2003 (EMOP 10200.3)." World Food Programme, Rome.
Ziegler, A. D., T. W. Giambelluca, R. A. Sutherland, M. A. Nullet, S. Yarnasarn, J. Pinthong, P. Preechapanya, and S. Jaiaree. "Toward Understanding the Cumulative Impacts of Roads in Upland Agricultural Watersheds of Northern Thailand." *Agriculture, Ecosystems and Environment* 104 (2004): 145–58.

Index