

SCIENCE FOR IMPACT

BETTER EVIDENCE FOR BETTER DECISIONS



© 2018 International Bank for Reconstruction and Development /The World Bank
1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Rights and Permissions

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

Cover photo/illustration: dobrodzei/Shutterstock.com
Cover and interior design and layout: Circle Graphics, Inc.



CONTENTS

ACRONYMS iii

ACKNOWLEDGEMENTS v

1 EXECUTIVE SUMMARY **1**

2 DIME MODEL FOR REAL TIME EVIDENCE-BASED POLICY **7**

2.1 Engagement with the Global Practices 7

2.2 Engagement with the Clients 9

2.3 Collecting Data and Setting-Up Data Infrastructure 16

2.4 Generating Evidence and Motivating Change 20

3 SELECTION & QUALITY ASSURANCE FOR POLICY RELEVANCE & TECHNICAL RIGOR **27**

4 MEETING OUR TARGETS **33**

5 HOW WE CONTRIBUTE TO DEVELOPMENT ISSUES **41**

5.1 Growth 41

5.2 Shared Prosperity 46

5.3 Risk and Vulnerability 50

5.4 Governance and Accountability 55

5.5 Global Public Goods and Externalities 58

6	ECONOMIES OF SCALE IN LEARNING	63
	6.1 Fragility, Conflict and Violence	63
	6.2 Agriculture	72
	6.3 Governance	83
	6.4 Climate Change	88
	6.5 Financial and Private Sector Development	95
	6.6 Transportation	100
	6.7 Gender	107
	6.8 Edutainment	112
7	INNOVATION IN DATA QUALITY & MONITORING OF POLICY INFLUENCE	119
	7.1 DIME Analytics	119
	7.2 Monitoring System & Results	125
	7.3 Policy Influence Verification	128
8	CASES OF POLICY INFLUENCE	145
9	COMMUNICATING LESSONS AND RESULTS	153
	REFERENCES	159
	ANNEXES	<i>See online version of report for full annexes</i>



ACRONYMS

3ie	International Initiative for Impact Evaluation
AADAPT	Agricultural Adaptations
BRT	Bus Rapid Transit
CAPI	Computer Assisted Personal Interviews
CBO	Community-based Organization
CBT	Cognitive Behavioral Therapy
CEGA	Centre for Effective Global Action
CIF	Climate Investment Fund
CCSA	Cross-Cutting Solutions Area
DE JURE	Data and Evidence for Justice Reform
DfID	Department for International Development
DIME	Development Impact Evaluation
DRC	Democratic Republic of the Congo
E&E	Energy and Environment
E4P	Evidence for Peace
EOI	Expression of Interest
ESMAP	Energy Sector Management Assistance Program
EU	European Union
FCV	Fragility, Conflict and Violence
FPD	Finance and Private Sector Development
GAFSP	Global Agriculture and Food Security Program
GEF	Global Environment Fund
GIZ	German Agency for International Cooperation
GoR	Government of Rwanda
GP	Global Practice
HH	Households
i2i	Impact Evaluation to Development Impact
IBRD	International Bank for Reconstruction and Development
IDA	International Development Association

IDP	Internally Displaced Person
IE	Impact Evaluation
IPA	Innovations for Poverty Action
IPV	Intimate Partner Violence
KePSIE	Kenya Patient Safety Impact Evaluation
LHW	Land Husbandry, Water Harvesting and Hillside Irrigation
LIPW	Labor Intensive Public Works
M&E	Monitoring and Evaluation
MATD	Ministry of Territorial Administration and Decentralization
MDB	Multilateral Development Bank
MINAGRI	Ministry of Agriculture and Animal Resources
NORAD	Norwegian Agency for Development Cooperation
OAF	One Acre Fund
PACT	Project for Agriculture Commercialization and Trade
PAD	Project Appraisal Document
PES	Payments for Ecosystems Services
RCT	Randomized Controlled Trial
RFR	Rural Feeder Roads
RWF	Rwandan Franc
SDG	Sustainable Development Goal
SEBRAE	Service of Support to Micro and Small Enterprises
SME	Small- and Medium-Sized Enterprises
STD	Sexually Transmitted Disease
T&C	Trade and Competitiveness
TDD	Transport and Digital Development
TTL	Task Team Leader
UNICEF	United Nations Children’s Fund
USAID	United States Agency for International Development
WASH	Water and Sanitation for Health
WB	World Bank
WDR	World Development Report
WFP	World Food Program
WTP	Willingness to Pay



ACKNOWLEDGEMENTS

This report was prepared by World Bank DIME staff and consultants. It covers the reporting period from March 2017 to March 2018. The main authors of this report included: Arianna Legovini, Guignon Serge Adjognon, Guadalupe Bedoya, Theophile Bougna, Kayleigh B. Campbell, Paul Christian, Aidan Coville, Kevin Croke, Benjamin Daniels, Vincenzo Di Maro, Alice Duhaut, Chloe Fernandez, Marcus Holmlund, Saahil Karpe, Maria Jones, Nausheen Khan, Florence Kondylis, Sveta Milusheva, Florentina Mulaj, Eric Mvukiyehe, Victor Orozco, Caio Piza, Arndt Reichert, Besa Rizvanolli, Daniel Rogger, Michael Roscitt, Bilal Siddiqi, and Astrid Zwager.

Substantial contributions were made by other members of the DIME team including: Avi Ahuja, Livia Alfonsi, Luiza Cardoso De Andrade, Alexis Rivera Ballesteros, Odbayar Batmunkh, Santiago De la Cadena Becerra, Laura Bermeo, Jonna Margareta Bertfelt, Kristoffer Gustav Bjarkefur, Emanuele Brancati, Brooke Alexis Braswell, Julie Claire Moranne Bousquet, Silvia Caroco, Elisa Anne Cascardi, Vanessa Cheng, Emily Crawford, Yuri Lacerda Costa, Laura Elena Costica, Amy Dolinger, Felipe Alexander Dunsch, Sarah Elise Elven, Cecilia Escalante, Ayodele Emmanuel Fashogbon, Miguel Lino Ferreira, Margherita Fornasari, Isabela Brandao Furtado, Aram Gassama, Jules Gazeaud, Steven John Glover, Gogi Grewal, Maria Camila Ayala Guerrero, Rebecca De Guttry, Kelsey Rae Harrison, Loredana Luisa Horezeanu, Josine Umutooni Karangwa, Kerenssa Mayo Kay, Safa Mohsin Khan, Roshni Khincha, Matilda James Kivelege, Sekou Kone, Seungmin Lee, Jessica Mahoney, Robert Marty, Shivang Mehta, Lars Fluge Nordgreen, Gustavo de Padula Novaes, Michael Adedoyin, Sunday Orevba, Ibrahim Ouattara, Lacey Jean Ramirez, Aurelie Melanie Rigaud, Mrijan Rimal, Sushmita Samaddar, Jérôme

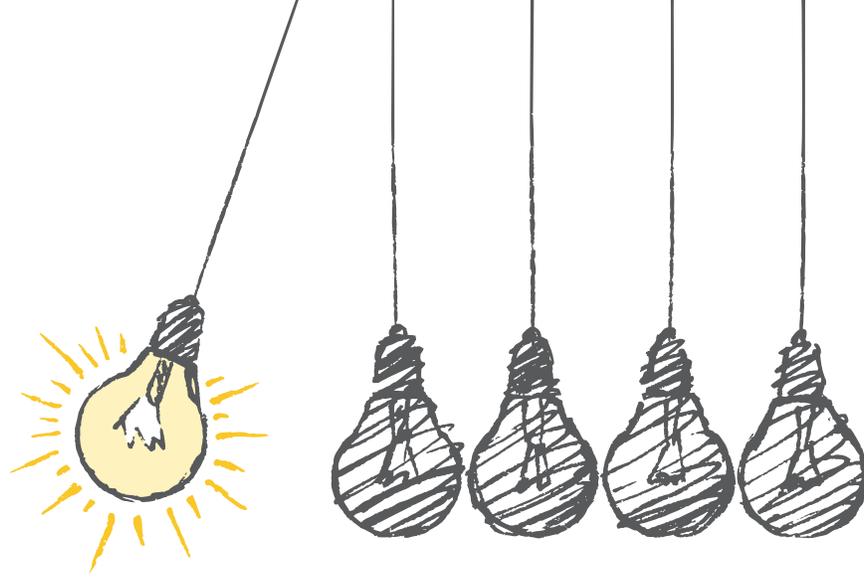
Sansonetti, Catalina Salas Santa, Sakina Shibuya, Marcela Mello Silva, Alexander Matthew Spevack, Kathryn Steingraber, Revaz Surguladze, Tomomi Tanaka, Natalia Tsivadze, Ijeoma Mariet David Ukoko, Lysca Uwamariya, Marco Valenza, Leonardo Viotti, Qiao Wang, Susumu Yoshida, and Wenqing Zhu.

Special thanks go to many staff around the Bank that have been actively contributing to the DIME agenda over the years. Among them we highlight: Jishnu Das and David McKenzie for their contributions to the research agenda as part of the i2i technical committee; David Evans for the Africa and Human Development agendas and for improving the communication of research; Marilou Uy, Christine Qiang, Marialisa Motta, Mary C. Hallward-Driemeier and Ana Goicoechea for the Private Sector and Finance program across five rounds of development; Mark Lundell, Mark Austin, Natasha Hayward, Anuja Kar, and Yasser El-Gammal for mainstreaming IE in Agriculture and Rural Development; Nicholas Menzies, Georgia Harley, Jorge Luis Silva Mendez and Amitabha Mukherjee for their role in the Justice program; Asmeen Khan, Vivek Srivastava, Michael Roscitt, Zaid Hasnain and Jurgen Blum for their role in the Governance program; Pierre Guislain, Nancy L. Vandycke and Javier Morales Sarriera for their role in the Transport program; Christian Borja-Vega for his contribution to the WSS agenda; Stefano Pagiola and Sandra Romboli on

the Climate agenda; Kabir Malik and Juliana Victor for the Energy agenda; Colin Bruce, Nadia Piffaretti, Mattias Lundberg, Jan von der Goltz, Fareeba Mahmood and Laura Ralston for the role in the FCV agenda; Carolyn Turk for the Rwanda program; Diarietou Gaye and Wendy Ayres for the Kenya program; Marie Françoise Marie-Nelly for the Nigeria program; and Martin Raiser and Pedro Olinto for the Brazil program.

Finally, we want to recognize the many contributors from academia, donor and development institutions that have been working with us over the years to change the way we do development. We acknowledge the financial contributions of the following donors: United Kingdom Department for International Development, European Union, Norwegian Agency for Development Cooperation, Global Agriculture and Food Security Program, Gates Foundation, Strategic Impact Evaluation Fund, Jobs Multi Donor Trust Fund, South Asia Food and Nutrition Security Initiative, State and Peace Building Fund, Climate Investment Funds, Korea Partnership Facility, CEGA's Economic Development and Institutions Initiative, International Growth Centre, Japan Policy and Human Resources Development Fund, International Initiative for Impact Evaluation, Umbrella Facility for Gender Equality, Knowledge for Change Program, Stanford Institute for Innovation in Development Economics, Sweden Ministry for Foreign Affairs, and Program for Forests.





SCIENCE FOR IMPACT

BETTER EVIDENCE FOR BETTER DECISIONS





EXECUTIVE SUMMARY

Impact Evaluation to Development Impact (i2i) is a multi-donor trust fund managed by Development Impact Evaluation (DIME) in the World Bank Research Group. It is designed to generate high-quality and operationally relevant Impact Evaluation (IE) research to transform development policy, help reduce extreme poverty, and secure shared prosperity. i2i's purpose is to improve development practice through a greater emphasis on the dialectics of development and the empirical testing of competing hypotheses. Our model embeds learning into each element of the policy cycle, from policy design to operational options and from scope and scale of a program, to the structure of the next phase of learning. Some call this impact evaluation, some, science of delivery, and others, adaptive programming, but the underlying idea is to use scientifically valid operational research to generate useful data and evidence to inform operational decisions in real-time and achieve greater policy effectiveness. In other words, using data and evidence to save and improve lives.

i2i is different from other impact evaluation efforts along some important dimensions: (i) it adopts a programmatic approach to evidence-based policymaking to increase economies of scale in research and learning; (ii) it builds the capacity of a broad set of stakeholders to make systematic use of data and evidence through a collaborative learning-by-doing approach; and (iii) it expands experimental and iterative research into under-evaluated sectors and geographies. Results from a detailed client survey show that the approach secured an enormous amount of policy influence all through the policy cycle, with clients identifying at least four major decisions per project while guided by each IE's data and evidence. Each of these evidence-based decisions helped push policy interventions towards their efficient frontier,



thus increasing costs-effectiveness and value for money. Furthermore, building on the scale and scope of the i2i portfolio, DIME has been heavily investing in public good products (open source tools and resources) to help improve the quality of development research around the world. These products are being made available to local researchers who might not count with the same capabilities.

i2i funds IEs in all sectors, covering 12 operational sectors across four pillars of Shared Prosperity, Governance, Climate Change, and Human Development, and including two cross-cutting themes of Gender and Fragility, Conflict, and Violence. Based on current donor financing, support is not available for the Human Development sectors, specifically for Education and Health, and ongoing discussions are taking place to expand

donor base and funding to cover these sectors going forward. A recent grant from Norway is being directed to the development of a program in Education and ICT.

DIME's contribution, among others, has been to develop the technology and institutional processes and structures to ensure that its IE products are relevant and can influence the decision-making process for development. Internally, i2i works with a cross-institutional council composed of chief economists and directors from operational vice-presidencies, who provide strategic guidance; with working groups in each Global Practice, setting learning priorities and selecting IE cases; and with each operational team for the selected cases, to shape project design and structured experiments to guide project implementation toward greater



effectiveness. Externally, i2i engages with clients from the beginning to set research questions, adapt policy implementation, and agree on entry points to affect policy decisions using experimental results. This model transfers knowledge and tools needed to support evidence-informed policymaking to country institutions through a learning-by-doing approach.

i2i is financed by DFID, EU, and NORAD, but many more have contributed financing to the program including the Gates Foundation and GAFSP, and ongoing discussions are taking place with other bilateral agencies to expand partnerships. World Bank GPs and country units also make a substantial contribution both for IE products and capacity building workshops, and most of the funding for data collection is leveraged from country client governments. This has created the dynamics to

bring a program like i2i to the forefront of the World Bank discussions. Among its achievements, i2i is now present across all the Bank's Global Practices and has helped to affect the operational culture across the Bank and its partners, especially DFID and Multilateral Development Banks (MDBs), as they joined and contributed projects to i2i programs. As a result, i2i is currently evaluating US\$8.2 billion in the World Bank's and US\$1.3 billion in DFID's development financing.

Deliverables

i2i has a portfolio of 145 IEs across 47 countries, covering all the Bank regions and sectors. Progress of portfolio implementation is on track. As of January 2018, 28% of the portfolio has been completed and most of these IEs have already produced a final output, such as working paper, publication, or a project completion report. Fifty-one percent of the portfolio is in implementation phase, having passed technical and policy relevance review and having received both World Bank Country Director and Sector Manager approval. A remaining 21% of the portfolio is in preparation phase, with teams having passed technical and policy relevance and having received preparation grants from i2i to develop full technical proposals. This process usually entails research and Bank teams' travel to the country to meet with the client and to work together towards finalization of the methodology and implementation plan.

The i2i portfolio has been developed through close collaboration with the Bank's Global Practices, operational teams, and client counterparts. Each thematic program was initiated with a launch of an IE workshop for capacity building, aimed at developing strong IE team capacity and collaborations between the operational staff, country governments, and researchers. During

the period of March 2014 through March 2018, i2i completed 21 workshops, covering sectors of i) Fragility, Conflict, and Violence, ii) Agriculture and Rural Development, iii) Energy and Environment, iv) Governance, v) Trade and Competitiveness, vi) Transport, v) Justice, and vii) Edutainment. Many of these programs have completed follow-up workshops to strengthen design of ongoing studies and share results. In addition, i2i completed sector cross-cutting workshops focused on methods, such as on distributional impact analysis and innovations in big data analytics. In all these workshops combined, close to 3,000 people have been trained in impact evaluation methods and around 400 organizations have participated.

The i2i workshops serve three main purposes: they train counterparts in IE methods to provide them with the tools necessary for active participation in the design and implementation of their IEs; they expose participants to the latest available evidence so that it can be incorporated in their own program and policy design; and they pair operational and country teams with highly qualified IE researchers to develop prospective IE proposals. Most importantly, workshops equalize opportunities to access financial support across the development community because by pairing policy teams with researchers, they improve the technical quality of their proposals. On average, proposals submitted to the i2i Calls for Proposals that come from teams that attend i2i workshops receive higher technical and policy relevance ratings from independent reviewers.

Policy Influence

The value of IE as a tool for more effective policymaking is increasingly recognized by governments and donors. Prospective multi-arm

IEs designed to reliably identify cause-effect relationships based on counterfactual analysis can guide decisions over which policies and programs are better able to achieve desired objectives and how to best design these programs for maximum impact. A study using data from Bank projects approved between 2005 and 2011 finds that projects with IEs are better at implementing development activities and do so in a timelier manner. Using disbursements against agreed activities as an objective measure of implementation, the study estimates that IE increases average cumulative disbursements by two-fifths (40.8%) and reduces the planned-to-actual disbursements gap by half (54%). The results suggest that IE is a powerful tool to move projects from design to implementation. In so doing, projects are more likely to achieve their objectives, supporting the idea that project financing and IE research are complementary services provided by the World Bank.

Further, i2i's model aims to generate high-quality data and evidence to motivate real-time policy change via a close collaboration with policy implementers. DIME seek to learn when, where, and how policy is influenced and what we can do to improve the influence of data and evidence. The latest i2i monitoring system survey conducted in November 2017 to understand the impact of IEs included World Bank IE task team leaders, World Bank operations teams, as well as external counterparts/implementing agencies. The results were striking. Respondents report that the IE has influenced their policies and programs across the indicators assessed, and the measure of influence in these areas is relatively high. The greatest influence reported concerns the contribution of IEs to the program/policy M&E function (91%); followed by helping rationalize the design of the program/policies (82%); informing design

through baseline results (66%); the adoption of a treatment arm or another IE element (61%); influencing other projects outside the IE itself (60%); and motivating scale-up/-down of program/policy (47%). Finally, 94% of government and implementing agencies and Bank operational teams report that the IE has added value to their units.

Examples of influence and scale are many. A large-scale evaluation of a school-based financial education program in **Brazil**, for example, found improved financial knowledge, attitudes, and behavior of students, which resulted in a national scaling-up of the program by the Ministry of Education. Another evaluation in **Mozambique** showed that training and placing women in extension delivery positions in agriculture benefits the broader population of women in terms of technology awareness and adoption. These results are being considered in the design of a new agriculture extension policy in Mozambique. In **Nigeria**, evidence and learning from evaluations of maternal and child health interventions indicates that the lives of mothers and babies are being saved, and these results are informing health policy discussions with the incoming Nigerian government. Also in **Nigeria**, the results from the MTV television drama *Shuga* indicate that the treatment group was twice as likely to get tested for HIV, reported fewer concurrent sexual partnerships and reduced attitudes and behaviors related to gender-based violence. These results have contributed to more donor funding for edutainment interventions, such as Gates Foundation scaling up its grants and a new program starting with NORAD under i2i. In **Kenya**, the Patient Safety IE just finalized the rollout of the intervention, and even before impact results are available the IE made significant contributions by enhancing the regulatory

framework for health inspections and strengthening institutional capacity through inspection protocols and monitoring system. Early findings from this IE show large improvements in patient safety compliance which have led to the government's decision to scale-up at the national level.

Way Forward

The last decade has seen a remarkable increase in demonstration of the value for money and effectiveness of developmental programs run by the World Bank and other multilateral lending and grant-making institutions. The standards for demonstrating impact of developmental projects have also been raised significantly. Rigorous and well-designed impact evaluation can help answer the “what” and “how” of economic development and help design better policies. i2i is the largest international initiative designed to systematically learn from development experience based on rigorous impact evaluation. Its model is specifically designed to overcome challenges from traditional approaches to evaluation, including research-capacity constraints, coordination failures and transaction costs to establish researcher–policy maker relationships, and limited understanding of and ability to integrate IE into the implementation of policies and programs at scale. All i2i-supported IEs build on both ongoing and completed work to create virtuous cycles of learning and policy impact. Most importantly, by engaging the government counterparts throughout all phases of IE cycle, i2i empowers governments to test innovations and to scale up solutions. Through this approach and by expanding donor partnerships and increasing financing for IE, the i2i will continue to improve accountability and learning from development interventions and shape the way development work is done at scale.





DIME MODEL FOR REAL TIME EVIDENCE-BASED POLICY

2

2.1 Engagement with the Global Practices

Engagement with Global Practices and CCSA

Engagement with the World Bank Global Practices (GP) and Cross-Cutting Solution Areas (CCSA) is a fundamental ingredient to developing economies of scale in learning and a strategic approach to changing development practice. GP and CCSA engagement serve multiple purposes: (i) knowledge priorities definition; (ii) strategic case selection; and (iii) portfolio learning and feedback. The definition of knowledge priorities is an iterative process that reflects both the composition of the project portfolio and an evolving understanding of what is important for GPs and CCSA to learn. They are usually reflected in selected thematic preferences in each round of program development. Strategic case selection is fundamental to ensure that important policies and investments are included in each IE program (as opposed to researcher-led project selection), and that the cases reflect the priorities of the GP or CCSA. Finally, portfolio learning is secured through a continuous process of knowledge diffusion with GPs and CCSAs that generates demands for summaries, presentations to GP management and knowledge boards and organization of IE events during fora.

Furthermore, GP engagement is used to align the priorities of Bank staff by increasing incentives for investing in operational knowledge. These incentives are both financial and non-financial to include: (i) clearly communicating managerial preferences and incentives for project teams, (ii) developing group dynamics and competition for excellence in learning, and (iii) providing direct financial incentives to conduct analytical work.

In practice, the level of engagement with each GP and CCSA is endogenous to each GP and CCSA history, management interest in impact evaluation, and whether or not i2i has covered a GP or CCSA as part of its program. GPs and CCSAs that have a full level of engagement with DIME and i2i include Governance, Trade & Competitiveness, Transport & Digital Development, and Fragility, Conflict, and Violence (FCV), each with senior management leadership, a dedicated team working closely with the DIME team, regular management consultations, co-sharing of program costs, and active

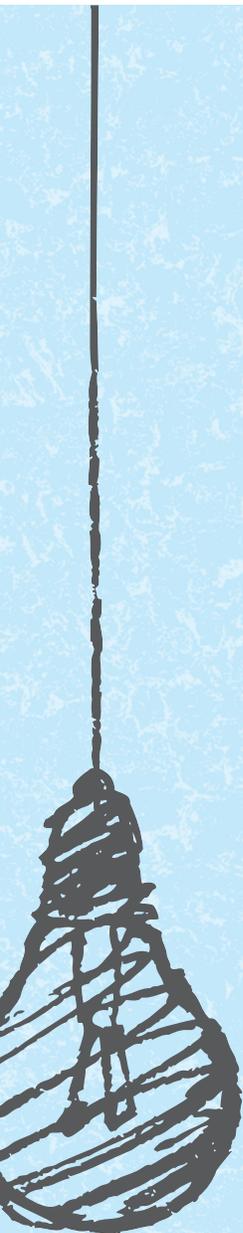
exchanges. The engagements are reflected by the large shares of the i2i portfolio in these areas. At a lesser level of engagement are Agriculture, Energy, Environment, Water, and Finance and Markets. With the exception of Agriculture, a large and striving program supported by the GAFSP, lower levels of engagement are reflected by smaller portfolios of IE. Finally, the Health, Education and Social Protection GPs were not directly targeted for programs and their large presence reflects non-i2i financed work or the intersection with cross-cutting areas, such as Gender and FCV.

DEFINING THEMATIC PRIORITIES IN T&C

The work with Trade & Competitiveness program (T&C) dates back to 2009 in the Africa Region. The Director for Finance and Private Sector recognized the importance of measuring the impact of matching grant and SME service projects. These represented the bulk of the GP's portfolio. Six matching grant projects became the subject of IE, but all failed to materialize for the lack of sample. This was due to the projects failure to attract firms' interest. In 2010, engagement with the business climate group expanded the program in the area of business climate reforms including simplification of firm registration procedures, tax reform, and inspection function. The evaluations of registration procedures resulted in mild effects on registration rates and non-detectable effects on firm growth even among firms that registered their business. The evaluations that resulted from these early engagements produced lots of learning for the

research and operations' communities. In the case of interventions aimed at improving firms' capabilities, such as matching grant and training programs, one learned that transaction costs can be prohibitive for some firms to apply for grants/subsidies, and the type of training interventions (e.g., length, in-classroom versus on-the-job) help explain low take-up rates and consequently programs' failures. The findings generated by rigorous evaluations in investment climate helped reshape the research agenda, as well. For instance, the limited effects of interventions trying to foster formalization rates jointly with the limited effects on firm performance among those that formalized resulted in formalization no longer being seen as a *sine qua non* condition for growth of firms. The new focus shifted towards identifying ways to make informal firms more productive. The 2015 T&C IE Workshop was a turning point as it brought to the center of debate cross-cutting

issues identified in previous evaluations, such as low take-up rates of supply side interventions, the need to target interventions better envisaging to maximize impacts, and the challenges associated with measuring spillover effects. Since then, project teams have been encouraged to rethink the way projects focusing on firm capabilities should be structured, and the importance of identifying firms that could most benefit from those projects. In 2016, the T&C GP identified three areas of work that would move the discourse from evaluating what projects do into stirring project designs into potential areas for firm growth: identification of high-growth potential firms, improvement in firm linkages, and enhancement of regulatory efficiency. The evolution of T&C thematic area reflects an iterative process of learning about operational practice and the need to put emphasis on testing strategies that help design operations in the future.



2.2 Engagement with the Clients

Workshops to Build Capacity and Stimulate Thinking

i2i is characterized by a bottom-up and demand-driven approach, high-quality and policy relevant research, and the use of IE as a strategic and formative tool to generate evidence throughout the project lifecycle. Early and sustained client engagement is therefore a core element of work under i2i. Engaging with government agencies and other clients early and often in the IE design process ensures that clients are active participants in the definition of research questions and design from day one. Furthermore, this sets the foundation for building client capacity and empowering policy makers and practitioners to exert control on their local environment through, for example, using data generated during an IE to make mid-course corrections or using final results to inform scale-up decisions. Ultimately, clients become educated producers

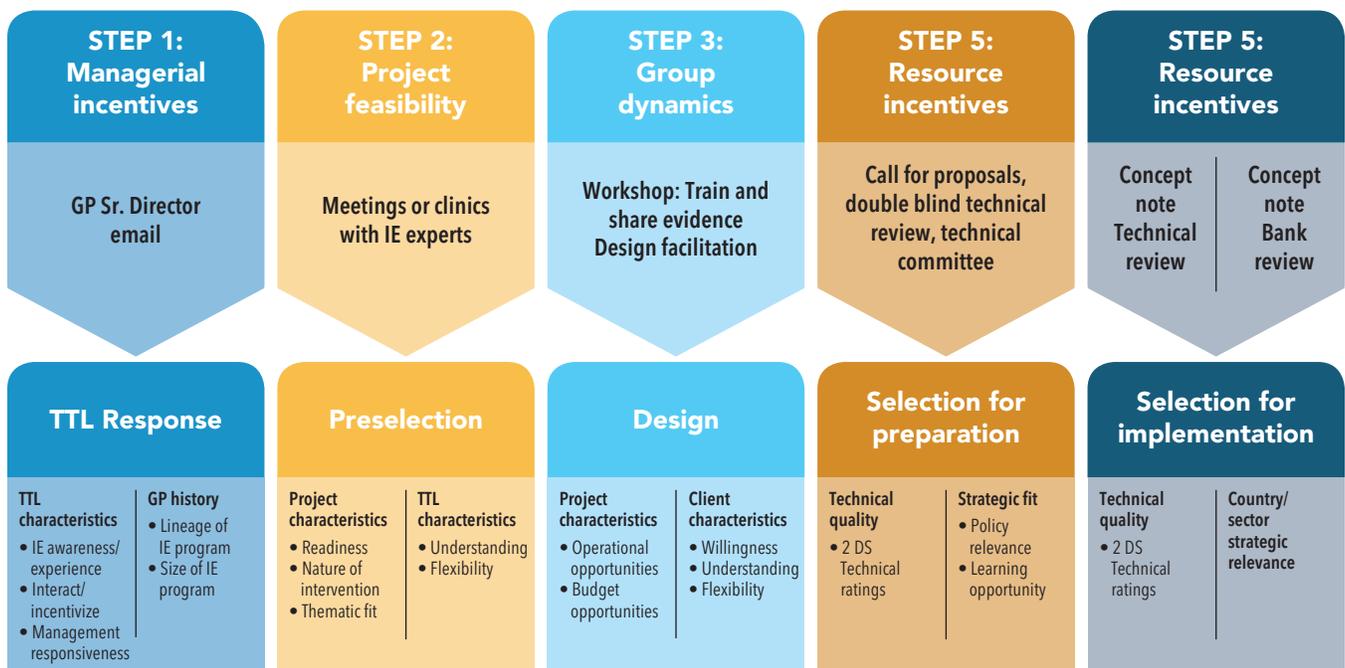
“DIME has helped the Ministry to understand better the use of strong evidence in decision-making.”

IE CLIENT, RURAL FINANCE EVALUATION, RWANDA

and consumers of evidence, whether from IEs or other types of research, and local capacity for the systematic use of data, experimentation, and evidence for policymaking is created. A 2017 survey of World Bank operations, TTLs, and government counterparts participating in DIME-led IEs confirms the value of the client-centered model, with 94% of respondents reporting that the IE added value to the project implementation unit and 91% reporting that the IE contribute to program and policy design.

IE researchers also benefit from the early building of a relationship with policymakers and

■ ■ **Figure 2.1 Engagement with GPs and implementation**



practitioners. This affords researchers a better understanding of the unique contexts where policy decisions are made and programs are implemented, allowing them to better tailor research questions and design. Early engagement between researchers, policy makers, and practitioners bridges the gap between theory and practice and kick starts a process of feedback loops between the various stakeholders. This is the foundation

for iterative learning. The IE workshop is the vehicle for initiating, stimulating, and strengthening this process. Thematic workshops are carried out periodically by each i2i program and unite policy makers, practitioners and operational staff, and World Bank and external researchers to advance knowledge on the research agenda set by the respective steering group. Table 2.1 contains a list of workshops and the number of persons

■ ■ **Table 2.1** Workshops and number of people trained, FY14–18

Workshop	Year	City	Sector	# of People Trained
Evaluating for Peace	March 2014	Lisbon	Fragility, Conflict and Violence	95
Innovations for Agriculture	June 2014	Kigali	Agriculture	126
Annual Conference on Measurement Technology	August 2014	Berkeley	Energy and Environment	60
Local Solutions to Global Problems	October 2014	Lisbon	Energy and Environment	71
ieGovern program launch	January 2015	Istanbul	Governance	143
Trade & Competitiveness	May 2015	Istanbul	Trade and Competitiveness	139
Global IE Workshop on Transport and ICT	June 2015	Rio de Janeiro	Transport & Digital Development	148
Energy and Environment Research Design follow-up	July 2015	Chicago	Energy and Environment	20
Understanding Bureaucracy through Anthropology and Sociology	February 2016	Washington, DC	Governance	179
Edutainment/ Narrating Behavior Change	May 2016	Mexico City	Edutainment	168
Using Evidence to Improve Policy and Program Designs	June 2016	Nairobi	Transport & Digital Development	66
Transport and ICT follow-up	June 2016	Nairobi	Transport & Digital Development	52
Evidence for Agriculture	November 2016	Washington, DC	Agriculture	58
Distributional Impact Analysis	December 2016	Washington, DC	Methods	33
IE target for Growth	February 2017	Mexico City	Trade and Competitiveness	97
Beyond the Status Quo: Using Impact Evaluation Research to Drive Innovation and Improve Outcomes in Health	May 2017	Lagos, Nigeria	Health	84
ID4D	May 2017	Washington, DC	Governance	49
De Jure, Data and Evidence for Justice Reform	June 2017	Washington, DC	Justice	80
ieConnect for Impact	July 2017	Lisbon	Transport & Digital Development	141
Development Impact Evaluation Evidence for Agricultural Transformation	November 2017	Rome	Agriculture	40
Impact Evaluation Workshop of Mass Media Entertainment to Improve Development Outcomes	February 2018	New Delhi	Edutainment	85



trained over World Bank fiscal years 2014-2017. This includes first and second-generation sector-specific workshops which launch and revisit, respectively, a set of interrelated IEs, methods workshops, and dissemination workshops which focus on making accessible data and evidence from ongoing and completed IEs.

Workshops to Build Capacity and Stimulate Thinking

First, government counterparts and other partners receive instruction in IE methods and tools to equip them to participate actively in the development and implementation of their own IE and to become better-informed consumers of knowledge, whether generated through IEs or other methods. This would be valuable even if the engagement were to stop here as policy makers will already have been exposed to multiple sources of information and acquired a better

understanding of the type of evidence needed to understand causal links, both of which are vital to better decision-making.

Second, participants are presented with the latest rigorous evidence in each focus area with a view to incorporating relevant findings in their own program and policy design. For many policy makers and practitioners, this is a rare opportunity to update technical knowledge in their area and to interact directly with top academics and peers from other agencies and countries. Presentation and discussion of existing evidence enhances cross-country learning, stimulates create thinking, provides ideas that may be incorporated and tested in one's own IE, and illustrates the feasibility of conducting high-quality and often experimental IEs at scale in complex settings.

Third, each country team is matched with one or more researchers based on compatibility of

interests. Just as country teams are invited to workshops based on preselection conducted in collaboration with the relevant GPs, researchers are likewise selected based on their knowledge, interest, availability, and capability to work on long-term research collaborations with governments. Researchers and country teams begin their work together on the first day of the workshop, thereby setting the foundation for collaboration and compromise between what is feasible

and what is ideal on both sides. This ensures that the resulting product is of high technical quality, addresses pressing policy questions directly, and contributes to answering broader development questions. While efforts are made to get this matching right from the outset, in those cases where this is not fully successful, workshops serve as a marketplace where both researchers and government teams alike can find a suitable match. Successful matching is critical for IEs as

Figure 2.2 Projects which attend a workshop consistently achieve higher technical scores in their i2i proposals (technical ratings of proposals by program on a 0 to 3 scale)

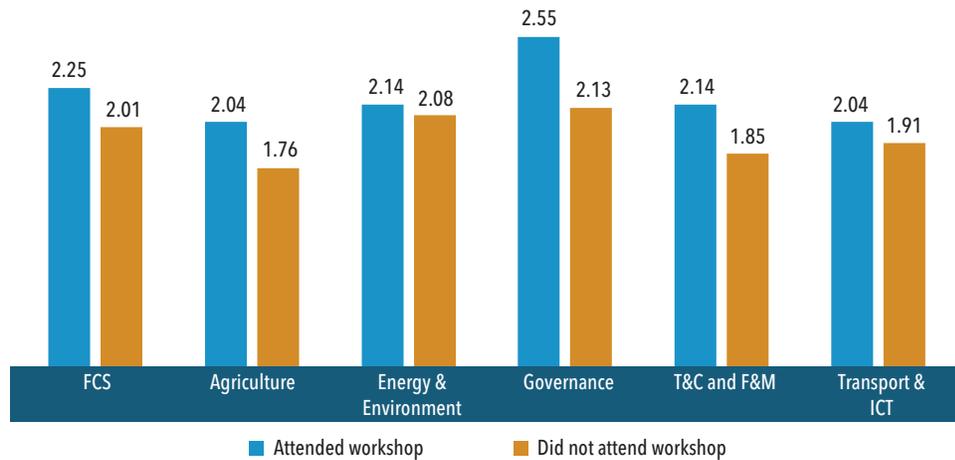
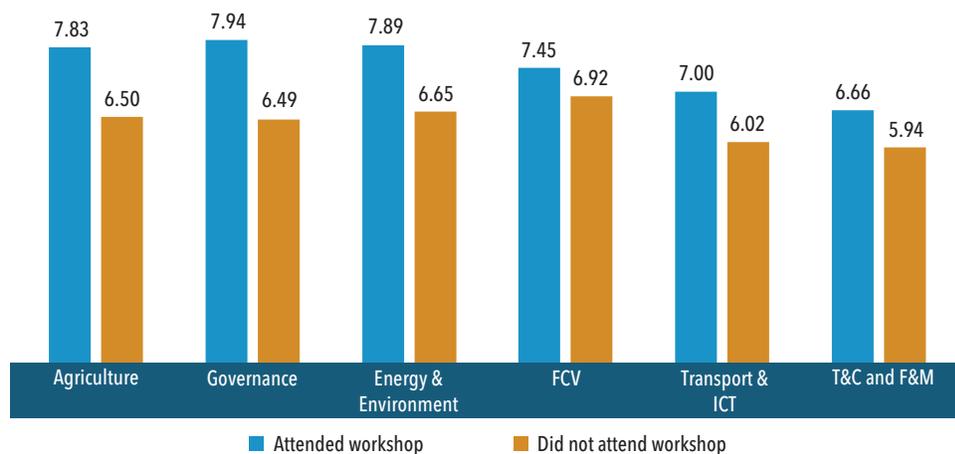


Figure 2.3 Projects which attend a workshop consistently achieve higher policy scores in their i2i proposals (policy relevance ratings of proposals by program on a 0 to 10 scale).



IE 4 PEACE WORKSHOP, LISBON, MARCH 2014

The first workshop under i2i launched the IE 4 Peace program which focuses on Fragility, Conflict, and Violence (FCV). This was co-convened by three World Bank Group teams: DIME, the FCV Group (at that point known as the Center for Conflict, Security, and Development), and the Latin America and Caribbean Region's Citizen Security Team. The workshop brought together practitioners, subject experts, and researchers to promote the strategic use of evaluation to inform policy and program design and advance knowledge on key issues related to FCV under four themes which then became the focus areas for the first two years of program implementation under i2i: jobs for resilience; public sector governance, urban crime

and violence, and gender-based violence.

In total, 22 teams from across the world attended the workshop, representing diverse countries such as Democratic Republic of Congo (DRC), Colombia, Honduras, Russia, and Papua New Guinea. Additionally, researchers from the World Bank and leading universities, such as Harvard and MIT, contributed their expertise and worked with individual project teams on initial design concepts. Development partners including DFID and USAID also participated, as did World Bank FCV and operational staff and research organizations, such as Innovations for Poverty Action.

Following the workshop, the first i2i call for proposals was

launched. 33 eligible EOIs were received and 14 projects were selected for funding, including several that had benefitted from training, experience and evidence sharing, and researcher matching through the IE 4 Peace launch workshop. Several of these are now under implementation and incorporate ideas first developed at the workshop. For example, after learning about the Becoming a Man program in Chicago, the Honduras Safer Municipalities Project decided to incorporate Cognitive Behavioral Therapy (CBT) in a training and labor market insertion program for at-risk youth in high-violence municipalities. The project is now testing variations in the delivery of CBT through its IE.

CASE STUDY



these are often multi-year engagements. Using the workshop as a vehicle to achieve this further equalizes opportunity for i2i funding calls opened after workshops and for other funding opportunities where both technical research quality and policy relevance count. This matching function of workshops is thus a core element of DIME's/i2i's model for the co-production of knowledge.

Fourth, IE teams develop the initial concept for a prospective impact evaluation to answer questions of direct policy relevance to their program. Providing dedicated "clinic time" for this work during each workshop day encourages teams to debate and incorporate relevant evidence, methods, and IE design options discussed throughout

the workshop week. This also allows for key concepts and ideas to be reinforced each day, contributing to workshops' capacity-building function. Oftentimes, the IE process achieves its first policy influence at this early stage by encouraging the incorporation of lessons learned elsewhere and the testing of alternative policy designs (or "variations in treatment"). IE teams are then charged with presenting their preliminary IE design on the final day of the workshop, which ensures that teams go home with a clear concept and roadmap. Furthermore, the presentation is always carried out by a government team member, which reinforces ownership and empowers practitioners and policy makers alike to present this work within their governments.

From Design to Implementation

DIME IE workshops play the role of a collective brainstorming: they serve to create ideas and provide the basis for further thinking. Following the workshop, consistent follow-up is needed to consolidate ideas and to turn these into a series of concrete actions which, ultimately, will lead to the successful completion of one or more IEs.

A critical first step is to secure buy-in from a broader range of stakeholders than those that participated directly in the workshop. This includes decision makers in government, World Bank operational and country management staff, and the IE research team. The objective is to form a coalition where everyone has a clearly defined role, including rights and responsibilities. This is essential if the IE is to correspond directly to country-specific and broader policy learning priorities, be carried

CASE STUDY

ATTRACTING & DEPLOYING TALENT TO REFORM CIVIL SERVICE IN THE DEMOCRATIC REPUBLIC OF CONGO (DRC)

Following the participation of the World Bank-assisted DRC Public Service Reform and Rejuvenation Project in the Evaluation for Peace workshop held in Lisbon in March 2014, the IE team and government continued working together to refine the IE concept and move from design to implementation. The IE, *Attracting & Deploying Talent to Reform Civil Service in the DRC*, focused on interventions targeting the civil service and the young professionals working within it.

Initially, the IE design included one intervention and sought to evaluate the impact of civil service retirement and recruitment reform on the age structure of civil servants, qualifications, and skills of young professionals, and perceptions of fairness in the recruitment process. However, following the workshop, the evaluation was expanded to include treatment arms to assess: i) recruitment strategies to attract and deploy the

most competent and pro-social young professionals; ii) the impact of recruiting young professionals on the performance of their units; and iii) the additional impacts of mentoring the young professionals. This was done through a Meeting of the Task Force, which took place in August 2017.

When implementing the project in the field, one of the challenges that the IE ran into concerned data collection in fragile contexts. The availability of high-quality and reliable survey firms proved to be a challenge, as did the scarcity of enumerators with knowledge and experience of working in the project's contexts. To address this challenge and ensure meticulous data collection for rigorous results, the IE team decided not to recruit a firm, but rather build the capacity of the government statistical agency. Training workshops were held in the DRC for government representatives working in relevant government units and

offices, who were trained on running a survey that would be flexible and adaptable to the local contexts amid security threats. For example, they were trained on the need to have both electronic and paper versions of the survey instrument, should there be a need to switch from an electronic format to a paper based one or vice-versa; and were also trained on developing a data management system to monitor and manage field activities. Through this type of capacity building, DIME enables partners and client agencies to take ownership of data collection and other research processes, thus setting the framework for the take-up of results and improving in-country know-how.

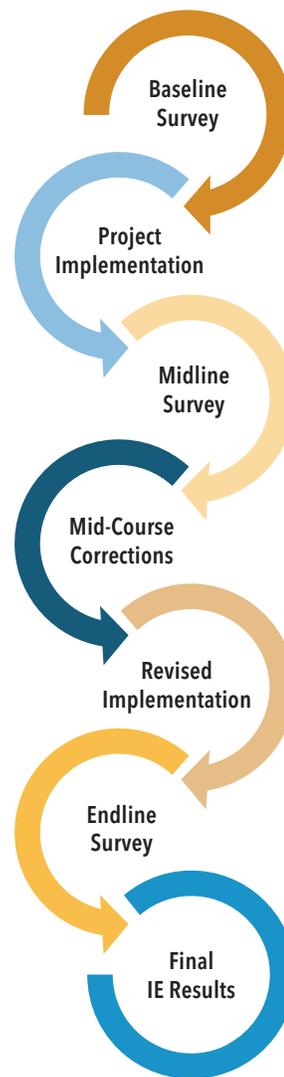
This initiative was possible due to the close cooperation between the DIME research team and the DRC government project implementation unit. IE implementation is currently ongoing.

out at scale in the context of a government program, and to facilitate the use of its intermediate and final findings as policy and program management inputs.

An important impetus for building such a coalition is the possibility of seed funding from i2i. In preparing their expressions of interest (EOI), teams build on the initial concept developed at the workshop through refining key details of IE design including the types of interventions and number of treatment arms to be tested, identification strategy, sample size, key outcomes, and budget and timeline. The i2i EOI is often the first attempt at situating the IE within a framework of existing knowledge and defining how it will contribute both in its immediate context and more broadly. It is also often the first signal of policy influence of the IE research process: 82% of IEs contribute to rationalizing policy design through informing the design of a particular intervention (or “treatment”) based on existing knowledge, or through introducing variation in existing interventions or entirely new interventions to be tested. While these things may have been initially discussed at the workshop, their inclusion in the EOI signals a broader consensus to go ahead.

Once seed funding is secured, IE and operational teams work towards fully defining the research design and documenting this in a concept note, which is subject to peer review for both technical quality and policy relevance. This document serves as a blueprint for the IE. Following quality clearance by i2i reviewers, a review meeting is held to discuss the technical, operational, and policy implications of the proposed work which is chaired by the relevant World Bank Country Management Unit. This validates the policy relevance of the proposed IE, informs a broader set of stakeholders, and contributes to highlighting

■ ■ **Figure 4** Data in the IE lifecycle



the IE as an important part of the Bank’s technical assistance to its client countries.

The completion of the concept note review is the final step in the IE design process and represents a formal commitment by all parties—the government, World Bank, external researchers, and development partners—to continue working on the IE through to completion. It is often followed by the deployment of a field coordinator (if such a person is not in place already), who serves as an in-country liaison for the research team and whose role is to support



the government on all aspects of IE implementation including operational planning, supervision, data collection, and dissemination. With the field coordinator in place, work continues to develop the operational aspects of IE interventions to coordinate their implementation and to plan for data collection.

2.3 Collecting Data and Setting Up Data Infrastructure

High quality data is a hallmark of the i2i portfolio. i2i research teams provide technical assistance on data collection throughout the lifecycle of the impact evaluation. At the initial IE workshop, teams design a data strategy aligned with their project cycle and discuss key points of influence. A typical data strategy includes both in-depth surveys and routine monitoring data.

The foundation of the impact evaluation analysis is high-quality microdata from multi-module surveys of a representative sample of the target population. These surveys are completed at key influence points during the IE lifecycle, typically before project implementation (baseline survey), midway to project completion (midline survey), and at project closure (endline survey). In developing a data strategy, the researchers discuss with the project team how to align data collection to positively impact project design and implementation.

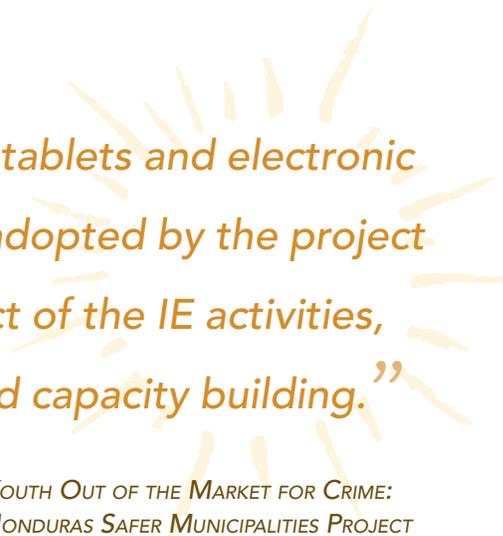
Providing data early and often throughout the project cycle is essential to the i2i data strategy, which is in sharp contrast to a project evaluation model in which a team of evaluators arrives after a project is completed to assess whether or not the project was successful. The i2i model aims

to provide real-time feedback and actionable information on how to improve. If results are not as expected at midline, there is the opportunity for mid-course corrections and additional learning-by-doing.

Effecting Decisions Through Better Data

Baseline surveys. While not technically necessary for randomized control trials (the majority of the impact evaluations in the i2i portfolio), baseline surveys provide an excellent opportunity to provide government ministries with high-quality, sector-specific data, which is almost never otherwise available. If timely, this data can provide valuable input to project design and implementation.

IE EXAMPLE: *The Impacts and Sustainability of Irrigation IE in Rwanda* provides a useful example of how baselines can influence project design. A baseline survey conducted on a sample of farmers cultivating within the irrigation areas provided the project team with detailed data on farm practices in the targeted area. This was well-timed to have influence: the team was in the midst of designing interventions to complement the irrigation infrastructure. For example, the project team had planned to collect fees by automatic deductions from sales to cooperatives. However, the data revealed that fewer than 2% of farmers made any sales to the cooperative, which forced a change in strategy. In addition, the data provided representative statistics on vegetable cultivation, which the project sought to promote, such as the realization of how few farmers had experience cultivating vegetables (5.3%) influenced the structure and intensity of agricultural extension.



“The use of tablets and electronic surveys was adopted by the project as a product of the IE activities, trainings, and capacity building.”

IE CLIENT, MOVING YOUTH OUT OF THE MARKET FOR CRIME: INTERVENTIONS IN THE HONDURAS SAFER MUNICIPALITIES PROJECT

Midline surveys. These are implemented mid-way through project implementation, at a point when the project expects to have achieved initial gains. The midline survey is a critical mechanism for improving project implementation and assuring that projects meet or exceed their development objectives. Discussion of findings of the midline survey hinges around potential mid-course corrections, and possible new experimental implementation variations.

IE EXAMPLE: *Rwanda Land Husbandry, Water Harvesting, and Hillside Irrigation Project IE.* There are two primary agricultural seasons in Rwanda, known as Season A (long rains) and Season B (short rains). A midline survey found that the project had met and, in fact, exceeded its objectives on increasing agricultural productivity, commercialization, and farm incomes during Season A. However, the project was significantly behind its commercialization target during Season B. As a result of this finding, the team increased their focus on Season B, shifting agricultural extension efforts and developing complementary interventions, such as an experiment testing new savings products for farmers.

Endline surveys. These are completed at the end of project implementation. The objective is

to capture the full lifecycle impact of the project, and measure gains from mid-course corrections by comparing indicators to the midline. The research team presents preliminary findings to the project team in a dissemination mission, and after incorporating feedback prepares an impact evaluation report and policy brief. A primary objective of the dissemination mission is to discuss policy implications of the impact evaluation findings, particularly opportunities for scale-up or scale-down. The next chapter discusses policy implications in detail.

Building a Data Infrastructure

The i2i data strategy aims to create a comprehensive data infrastructure by integrating with the project monitoring system, utilizing existing administrative data, and testing new measurement technologies. The end objective is to create a data infrastructure that is informative, allows for timely responsive action, and is sustainable beyond the duration of the specific project being evaluated.

Monitoring & (Impact) Evaluation. Rather than having parallel systems of the project's M&E and the impact evaluation data, i2i IEs strive to fully integrate the two. Monitoring systems are designed to use the same identification codes as the impact evaluation, so that data can be easily merged and compared, and key indicators (e.g., agricultural yield, firm profits, household income) are constructed in a consistent manner. Each i2i IE team includes a field coordinator, who is based in country and typically sits with government M&E staff. The field coordinator is primarily responsible for technical assistance on both in-depth surveys and monitoring data, and frequently conducts trainings on data management in Excel and

statistical software (e.g., Stata, SPSS) for government counterparts.

IE EXAMPLE: *The Kenya Patient Safety IE* provides an excellent example of creating a data infrastructure, fully integrating monitoring and impact evaluation data, and filling an important data gap. The monitoring system includes: (i) data on planning and progress of the inspection pilots (e.g., Are inspections taking place?); (ii) inspection results at the facility and aggregate levels for each pilot (e.g., How are facilities performing in each intervention?), and (iii) third-party monitoring indicators to assess quality of intervention and protocol adherence (e.g., What is the quality of the inspection delivered?). This customized solution then leads to the availability of timely and actionable information to identify challenges in the implementation and to enhance accountability to make mid-course corrections, without the intensive use of resources, expertise, or equipment commonly absent in poor-resource contexts.

Administrative data. Many of the i2i impact evaluations also incorporate existing administrative data into their data infrastructure. Governments typically have large amounts of existing data, but it is often in hard copy only, or lacks identification codes, inconsistently structured, or not centralized at the national level. As a result, integrating this data requires trips to field offices, digitization, and painstaking efforts to merge on available variables. However, the gains can be substantial, and it has the positive externality of creating a useful data infrastructure sustainable beyond project completion.

IE EXAMPLE: *Building a Supportive Environment for Operation and Maintenance in the Tanzanian Rural Water Supply Subsector*

IE. To understand the extent of interactions between government employees (and particularly water technicians) and village citizens, the team obtained records from visitor logbooks in rural villages in the project area. After consultation with the village chief, surveyors took photos of each page of the logbook, which were then digitized. This sort of data collection provides a rich source of information on engagement of rural communities with the government and NGOs, which had not been previously utilized or available beyond the village.

Big data. The next frontier in building responsive and sustainable data infrastructures is to establish the most cost-effective ways to monitor project outcomes by testing various indicators and measurement technologies. Household surveys are time-intensive and very expensive, limiting the potential for large-scale or high-frequency data collection. Big data has the potential to be a more cost-effective alternative. Recent technology developments have dramatically increased data availability and processing capacity. The task at hand is to apply these to program evaluation to rigorously test how well big data can be used to measure key outcomes of interest by comparing to gold standard methods to then carefully compare costs.

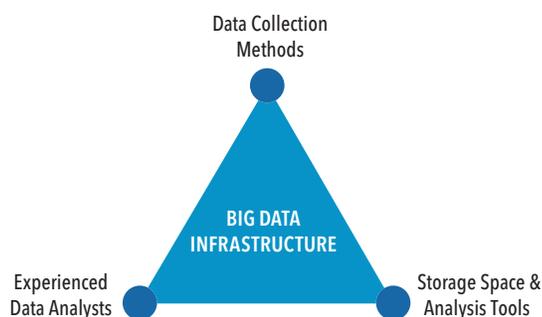
IE EXAMPLE: the *Nairobi smarTTrans IE* will develop a validated measurement framework for driving and road safety, and a set of instruments and crowdsourcing methodologies that can be deployed in Kenya and other low-income settings. The IE builds an ICT-based monitoring system leveraging the high penetration of mobile phones and low-cost technologies, including: 1) vehicle-based technology (sensors, GPS trackers); 2) app-based interfaces for owners and drivers to learn about road safety standards; 3) direct feedback mechanisms for drivers on performance and ways to improve their driving; and 4) a hotspot and entertainment box combined with SMS and Android platforms for consumer feedback on drivers' behavior. The technologies will be integrated into a big data infrastructure that will provide information to improve decisions by all actors in the urban transit system, from policy makers and regulators to private insurers, operators, drivers, and riders.

Ensuring High Quality Data

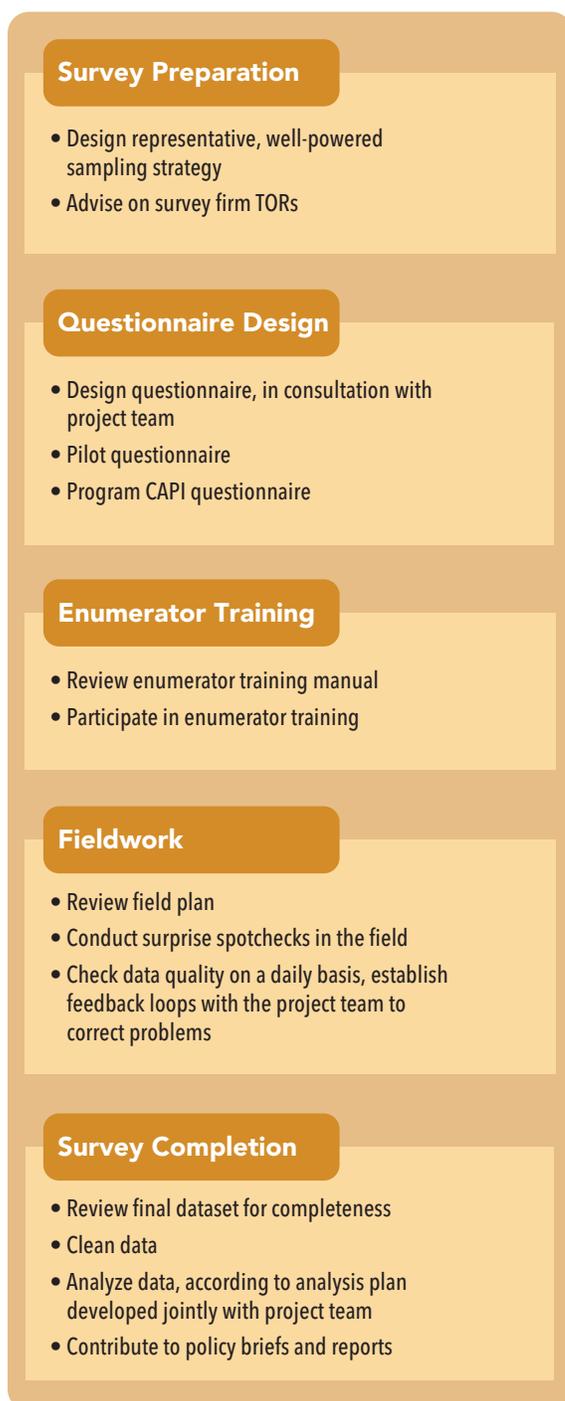
In the i2i portfolio, project teams finance all data collection. i2i provides technical assistance to ensure high quality data at each stage of the process: developing the terms of reference for a survey firm, designing the survey instrument, participating in enumerator training, supervising field work, and checking data quality and alerting the team of any issues in real time. In this way, i2i not only achieves high quality data, it builds capacity for high-quality data collection within the government ministry. Figure 2.6 outlines the main areas for technical assistance in a typical survey.

i2i surveys are typically Computer Assisted Personal Interviews (CAPI), with exceptions only in case of concerns of enumerator safety or extreme

■ ■ **Figure 2.5** Elements of a big data system



■ **Figure 2.6** i2i technical assistance during data collection



infrastructural limitations (e.g., electricity, internet, and transport). The i2i team typically program the questionnaire to ensure that programming meets the highest quality standards. CAPI technology has the potential to greatly increase data quality, but only if carefully programmed. All i2i surveys incorporate automated skip codes, range restrictions, internal consistency checks, pre-populated identification information for follow-up rounds of panel surveys, and validation before submission (i.e., that all expected fields are completed).

The field coordinators are the primary bulwark for data quality. They work closely with the project teams and survey firms on a day-to-day basis with support from the IE researchers. Recognizing their critical role, i2i supports an annual Field Coordinator Workshop in which all field coordinators come together as a community of practice to be trained on managing high-quality surveys, learn the latest developments in best practice protocols, and to build skills on relevant survey programming and data analysis software packages.

2.4 Generating Evidence and Motivating Change

DIME's ambition is to use rigorous evidence to motivate policy change in the world's poorest countries. At a minimum, this requires connecting lessons from our evaluations to new policy decisions. DIME's evaluation model aims to embed learning into each element of the project cycle, from defining policy, through the decision as to whether to continue a program or not, to the structure of the next phase of learning. Given its continuous nature, our model is a form of "real-time" learning.

In traditional evaluation models (as summarized in figure 2.7) the evaluation process is done after the program is completed. Policies inform designs, which are then evaluated, and the results are used to decide whether the program should continue or not. In the best cases, evaluation reports aim to distill wider learning from the efforts of evaluation, but there is no strategic linkage between that learning and the policy formation process. This could even be true if an impact evaluation is embedded in the project. If the evaluation team is seen as separate from the implementers, independently receiving data and reporting results in the final stages, the model is that in Figure 2.7.

Real-time learning implies that we should aim to undertake a circular model, where learning is an integral part of each stage of policy development. Figure 2.8 connects learning from previous project cycles to new ones. However, DIME’s ambition is to embed circularity and feedback loops in each stage of the project cycle. As exhibited in figure 2.8 by the anti-clockwise arrows, not only do we see learning as feeding into policy design, but feeding into the decision as to whether to continue a program and how to modify plans for project evaluation. Similarly, discussions around the continuation of the project become ones about optimal design: each stage of the project cycle can be designed so as to feedback into any other.

Embedding such learning in each stage of the cycle requires strong partnerships between ourselves and our colleagues at the World Bank,



other multilateral agencies, and our government counterparts. Chapters 2.1 and 2.2 explore how we build those partnerships with our colleagues and clients. This chapter will explore how we work with those partners at each stage of the project cycle to effectively generate evidence and motivate evidence-based change.

■ ■ **Figure 2.7** The traditional evaluation model



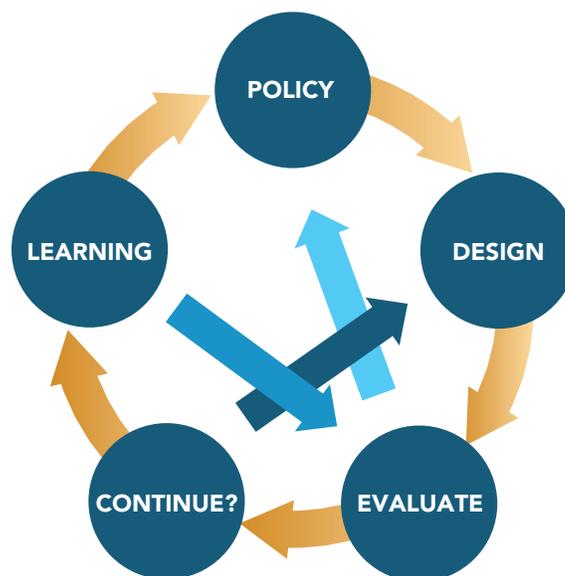
Policy

As discussed in chapter 2.1, we aim to engage with projects at an early stage of their development. Preferably, the policies under which projects fall will still be under development, which allows us to ensure that policies reflect frontier evidence from academic and policy research. This evidence identifies both what is known and what still needs to be tested so that evaluation can be embedded in policy itself. Ninety percent of DIME projects have facilitated the embedding of frontier research into government policy. This happens in both an *ad hoc* way and more systematically. For example, DIME's Fragile and Conflict States group worked with a series of senior academics and Bank staff to write white papers summarizing the frontier literature of the sector. These white papers ensure ready access to frontier research lessons and corresponding gaps for project teams working in the areas they touch upon.

Sometimes the relevant evidence will not be available in the research literature. DIME works with government counterparts and our colleagues to develop policies that reflect these ambiguities. The change we hope to inspire is a recognition in country policies that there is not currently an obvious right path for a country, and so different interventions will be experimented with until there is sufficient evidence to choose a single path or to expand multiple interventions to those communities as they best fit.

We also work to generate evidence to inform policy. In Tanzania, DIME is working with the UK's Department for International Development (DfID) on improving the maintenance of water infrastructure in rural areas. DfID's approach for this project is to pilot an "adaptive model": rather than contract DIME to undertake a single

■ ■ **Figure 2.8** DIME's evaluation model



evaluation formulated at the start of the project, they have written into the contract multiple stages of review, where, for example, their policy on water maintenance in Tanzania can be updated. Inspired by academic models of policy development, such as *Problem Driven Iterative Adaptation*, this model ensures that DfID's policy constantly reflects the most up-to-date information and insights from the field.

Design

Arising from policy is the design of interventions. Whilst policy sets the broad parameters of intervention, there are typically large areas of ambiguity in how to design a particular project or program. This is where a circular or real-time learning model facilitates evidence-based decision-making. Working with researchers allows operational staff access to frontier research on project design, and, similarly, working closely with operational staff allows researchers to identify research that is "operationally-relevant." It is in jointly designing a project that both sides of

the research-policy divide can best understand the perspectives of one another: this is what DIME strives to do.

For example, working closely with the Rwandan Government had given DIME's agriculture team a reputation for being easy to work with closely. The Government were keen to operationalize their agricultural strategy and requested DIME's support. DIME worked with the Government to design, introduce, and test innovative farmer feedback tools. Where it was not obvious how to design the tools, DIME helped the government trial multiple modalities. DIME's research background also allowed for an additional twist to be added to the project so that the evaluation could learn about the underlying mechanisms driving the results, finding that particular feedback tools increased attendance in agricultural extension trainings and increased adoption of superior farming techniques amongst non-users. The most cost-effective feedback mechanism (a hotline) was adopted and scaled up throughout Rwanda. In addition, the satisfaction data collected as part of the evaluation convinced the government to continue supporting a public-private partnership that had supported the implementation. The ongoing policy of the Government reflected frontier research.

Evaluation

Embedding an evaluation into any part of the project cycle allows for learning that can strengthen the project and provide broader lessons for policy around the world. Such evaluations can take many forms, both across and from within projects. In Ghana, DIME worked with the Office of the Head of the Civil Service to survey all civil servants on the bottlenecks they faced to improved service delivery. As discussed in chapter 2.3, the rigorous collection of carefully designed indicators

can be sufficient evidence to motivate change by itself. From the results of the civil servants survey, there were a series of obvious reforms that could be implemented immediately. For example, officials complained that organization heads were not being monitored as was intended by the Public Service Rules. The Head of the Civil Service immediately fixed the monitoring system. Similarly, in Kenya, DIME worked with the judiciary to build a 'Daily Court Returns Template' to scientifically gather and organize daily court output. This was both legislated into national law, as well as being a useful data collection tool for the impact evaluation.

“The most efficient study arm was scaled up. The Senegalese government replaced HIV community sensitization campaigns with peer-led campaigns after the IE. The implementing NGO intensively trained peer leaders for these to train in turn other community members (versus inviting everyone in the community to attending community sessions). The scale-up was national.”

IE CLIENT, LONG TERM EFFECTS OF HIV COMMUNITY SENSITIZATION CAMPAIGNS, SENEGAL



DIME frequently embeds a preliminary evaluation in its first year of engagement with a project. Using the best available data or that which can be collected immediately, DIME researchers aim to identify whether there is an empirical basis for an evaluation approach. In the Ghana case, there were areas that all civil servants stated were working well and with no need for reform. On balance, these didn't seem the first places to test for significant constraints to productivity. There were others that required further investigation, focusing the scope of our evaluation and changing the reform policy of government.

Using a multitude of impact evaluation methods allows DIME researchers to shine a light on areas of reform that government has been considering, along with those they had not yet conceptualized. These discussions are then used as the basis to identify a series of reform options that DIME typically evaluates using a randomized control trial. In 88 percent of our Randomized Controlled Trial evaluations, DIME looks to go beyond a simple

understanding of which flavor of intervention works best to why it is most successful. For example, DIME is now using an RCT to build a more effective training system for public officials in Ghana, something officials had highlighted as a major constraint in the survey.

Continuation

Many of our government and implementation partners require an answer as to whether an intervention should be continued or not. Public sector funding is constrained across the world, and so allocating program budgets efficiently is of importance to the effectiveness of the state. DIME aims to provide inputs to answer this question, but with a focus on how to generate the largest gains from a program. By isolating the best way to deliver an intervention, we allow ourselves to reformulate its design and evaluate it on its greatest strengths.

In Kenya, DIME worked on supporting regulatory reform in the health sector and supported a



review of the health inspections regime for the country. Responding to the government's question of whether health inspections were worthwhile, DIME researchers supported the passing of an enhanced regulatory framework for health inspections and the design of a toolkit of instruments to measure patient safety. The database created through these efforts allowed the team to identify multiple inspections regimes that could be trialed in an impact evaluation framework. Thus, the project is now looking at how to maximize the impact of patient monitoring.

Learning

This subchapter of the report has outlined how evaluation teams can organize projects to learn in "real-time." There is often a point in an evaluation cycle seen as an opportune time to reflect substantively on the learning process and aggregate the lessons to date. In the traditional model, this would be the point at which the evaluation report had been produced. In the DIME model, this can

occur at various points throughout the lifetime of a singular evaluation, or at the completion of multiple complementary evaluations. DIME's organization around thematic groups allow it to strengthen the lessons of any single evaluation by relating it to the learning from a range of others.

Working closely with the Nigerian Government on their healthcare sector, DIME undertook real-time learning in partnership with the Ministry of Health for almost a decade. We introduced variations in policy design informed by existing evidence on interventions in the sector. Generating a wealth of microdata, DIME supported the delivery of routine program elements (such as working to improve the timeliness of payments to frontline health workers) and fed the resulting data into decision-making across the health sector. To capitalize on the multitude of evaluations in the sector, DIME organized an "Evidence and Action" workshop that presented senior stakeholders with the results of all the evaluations in one go. This facilitated a sector-wide approach to evaluation that enhanced the learning of any single evaluation by allowing evidence from one project to be cross-checked and validated by another. DIME also benefited from this exercise by allowing us to conceptualize the structure of learning that would best fit our next tranche of work with the Nigerian Government.

Undertaking policymaking in the developing world can be a daunting experience. DIME strives to enable our partners to have access to better data, knowledge, and learning at each stage of the project cycle. Providing feedback to implementers with whom we continuously work also means that we ourselves are gaining knowledge in real time. Our model is changing and evolving with each new partnership and evaluation that we undertake.



SELECTION AND QUALITY ASSURANCE FOR POLICY RELEVANCE & TECHNICAL RIGOR

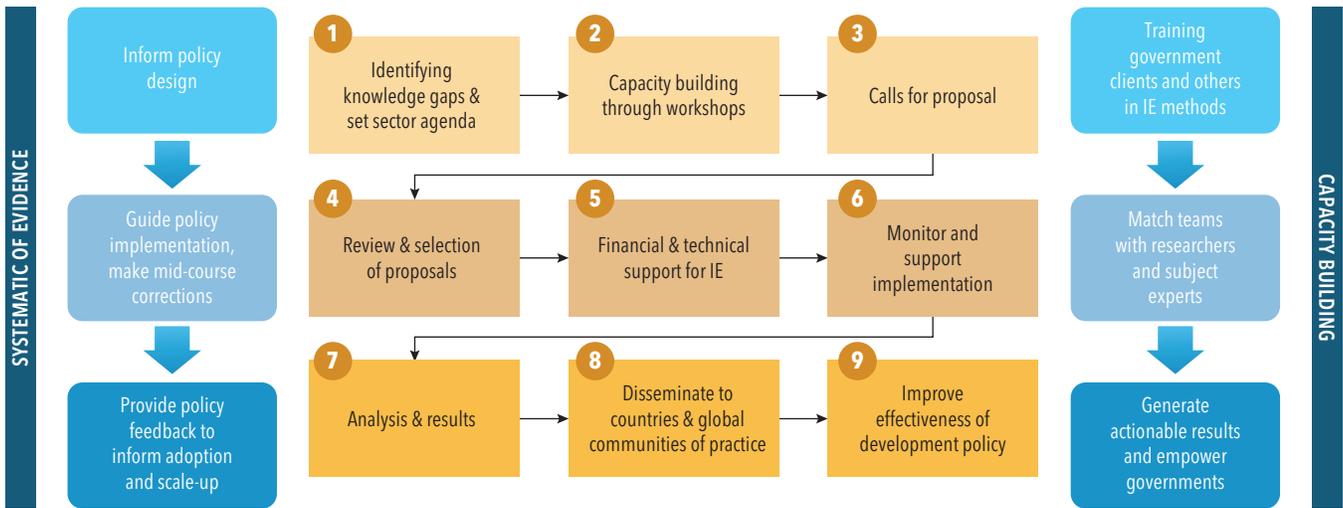
3

As previous chapters describe, DIME's goal is to increase the use of Impact Evaluation (IE) evidence in the design and implementation of public policy across the developing world. It works with programs at scale on game-changing issues to answer policy questions that clients identify and increase development outcomes. The highest technical quality and policy relevance is ensured at all stages of the project cycle while maintaining a flexible environment for adaptation based on real-time evidence and changes in context and client needs. How does DIME achieve this? It does so through an innovative operating model with a bottom-up approach, transferring IE knowledge and tools to clients and matching them with internationally renowned technical experts to deliver the highest quality products of policy relevance (see chapters 2.1 – 2.4 on DIME's model). Client engagement from the early phases of the design ensures relevance of policy questions, government buy-in, feedback loops, and policy action through almost all phases of project cycle. DIME projects report policy influence at baseline, during implementation, and scaling-up or down based on final results (see chapter 8 on Cases of Policy Influence). This approach, which is significantly different from the conventional development research, has defined DIME's model for the last 12 years.

DIME's approach also overcomes a critical challenge from traditional evaluation and research. That is, the disconnect from research and practice. To do it well, IE requires strong technical expertise, which is why most competitive Calls for Proposals reward applications submitted by researchers, as opposed to operational teams or country counterparts. As a result, there is often a lack of collaboration and client engagement in the design of the study and implementation, and results are often questionable in terms of their policy relevance



■ **Figure 3.1** DIME's operating model



and ability to generate policy action. DIME was built to overcome this barrier and serve as a bridge between researchers and policy makers. Its thematic programs start by defining boarder research priorities through a close consultation with the research community and development partners, including World Bank's global practices and regions (see chapter 2.1). It then identifies operationally relevant programs that enable impact evaluations around critical policy issues, and engages operational teams and government counterparts starting with a workshop to build capacity, raise awareness, and form partnerships with technical experts, setting the foundations for longer-term collaborations and policy influence.

At the project level, DIME's model generates evidence throughout policy lifecycle. As chapter 2.4 describes, at the design phase, DIME evaluations strengthen economic theory of interventions and make improvements based on existing evidence. During baseline and follow-up, data is used to stimulate policy dialogue and make mid-course corrections. During implementation, the IE strengthens clients' M&E systems and develops high quality survey instruments and data

collection protocols. During analysis and results, teams fine-tune country policies and programs based on evidence, often motivating scale-up or down. Results are also shared more broadly with the international development community through publications, seminars, workshops, and face-to-face interactions.

Applications for IE work are submitted to DIME through Calls for Proposals, usually announced after DIME workshops and targeting both teams participating in workshops, as well as the development community more broadly. All submissions both at the EOI and Concept Note stage undergo a rigorous technical and policy relevance review. External technical experts, identified for their IE and subject matter expertise, score proposals on a set of technical criteria through a blind review process. Internal Bank global practices and regional focal points score the proposals on policy relevance and feasibility of implementation. A technical committee, comprised of senior and lead economists from the Bank's Research Group, makes selection decisions based on technical and policy score, overall feasibility of implementation, and capacity to

EXAMPLE OF A THEMATIC PROGRAM DEVELOPMENT: ENERGY & ENVIRONMENT

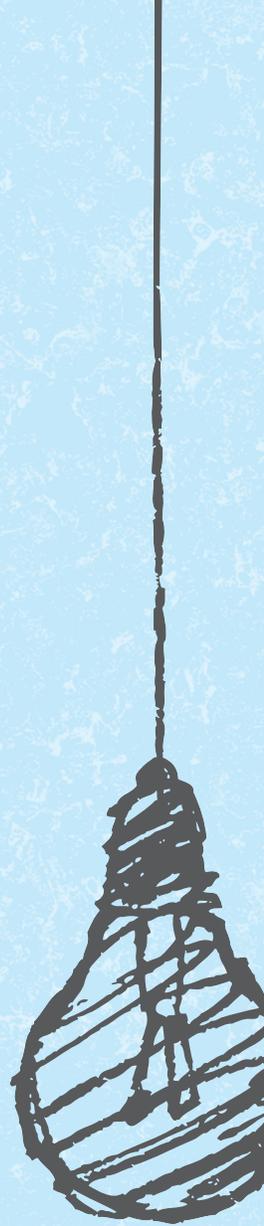
DIME's Energy and Environment program began in June 2014, following an iterative approach to balance the objective of addressing prioritized knowledge gaps with operational realities. The process included direct engagements with the Bank's Climate Change Cross-cutting Solutions Area, Energy & Extractives Global Practice, Environment and Natural Resources Global Practice, Water Global Practice within the World Bank, the Global Environment Facility (GEF), the Climate Investment Funds (CIF), the Energy Sector Management Assistance Program (ESMAP), and DFID (including the evaluation department and climate change teams). These partners actively contributed to the identification of the evidence gaps mostly from an operational standpoint and portfolio of potential projects aligned with those gaps. The main research partnership was with the University of Chicago led by John List, chairman of the Department of Economics. Other collaborations on measurement and technology include partnerships with the University of Berkeley and the Centre for Effective Global Action (CEGA). These partners contributed to the knowledge agenda by summarizing and discussing the current stock of academic

literature and future research priorities.

A kick-off workshop held jointly with CEGA in August 2014 in Berkeley brought together engineers, economists, and World Bank counterparts to explore how to leverage new technologies to improve measurement in energy and environment projects and research. A follow-up measurement workshop with the focus on innovative measures for climate resilience was held in June 2015. The main workshop and official launch of the program of work in Energy and Environment was held in Lisbon in October 2014 and brought together 19 project teams (financed through DFID, the GEF, CIF, IDA, and IBRD) and 28 researchers from 11 academic institutions to refine research opportunities based on project interest and operational feasibility. Representation from governments typically included high-level policymakers, project implementers, and/or M&E specialists. For instance, from the Bangladesh Rural Electricity Transmission and Distribution project, the Project Director, the Project Deputy Director, and the Executive Engineer of the government's Rural Electrification Board participated. From the Argentina Renewable

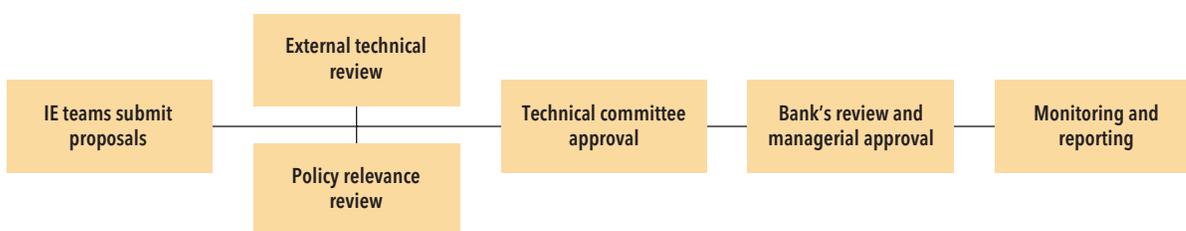
Energy in Rural Markets Project, representation included the coordinator general, electrician engineer, and an M&E specialist from the Energy Secretariat of the Ministry of Energy and Mining. During the week-long workshop, project teams spent time with researchers to develop proposals relevant to their specific projects and presented their achievements on the final day of the workshop.

Two months after the workshop in July 2015, researchers participating in the program were invited to present initial designs and then receive feedback during a session chaired by John List and supported by Michael Greenstone, director of the Energy Policy Institute at the University of Chicago. The first round of projects presented concepts after initial scoping missions that have led to the refinement and finalization of concept notes. Out of this program, three studies have now reached the completion stage, and six are under implementation. Client engagement continues throughout all stages of the impact evaluations. Examples include joint analysis and discussion of the baseline data to develop technical capacity in the government and inform intervention design.





■ ■ **Figure 3.2** i2i Review process



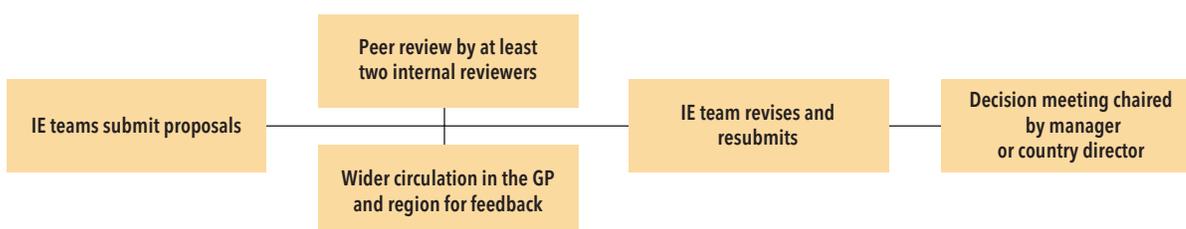
target important knowledge gaps. The teams that pass the EOI stage receive a preparation grant of US\$25,000 to develop a full technical proposal. The teams that pass the Concept Note stage receive an implementation grant of US\$150,000 (over three years).

After Concept Notes receive i2i approval, they undergo Bank's internal quality assurance process, involving a separate review meeting, chaired by country or global practice manager, and incorporating review feedback from at least two peer reviewers, usually a subject matter expert and an operations expert. This process of combining i2i's

and Bank's review processes ensures technical quality, buy-in from the client, and ongoing relevance to World Bank and country policy priorities. In cases where Bank regions rank a proposal highly on policy relevance and there is a strong commitment from the client to do the evaluation, but the technical evaluation from the external review does not meet i2i technical standards, DIME provides technical expertise to build capacity and revise the design.

During implementation, i2i supported IEs report annually on progress and potential channels through which they affect capacity and influence

■ ■ **Figure 3.3** Bank's review process



REVIEW CRITERIA FOR EXPRESSIONS OF INTEREST

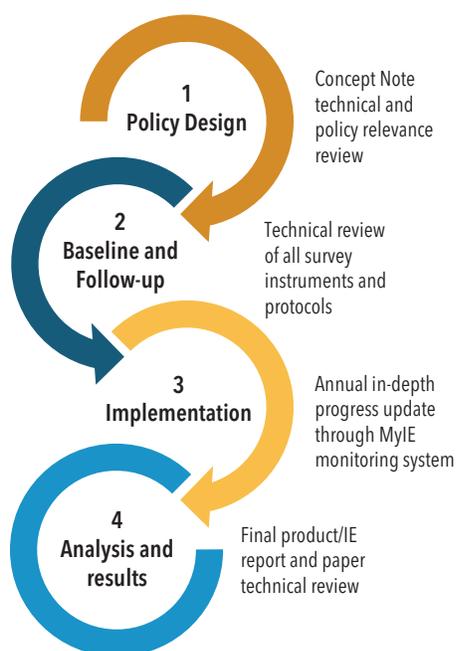
- Clarity of the research questions and potential to contribute to evidence gaps.
- Prioritized project components/interventions have logical pathways to intermediary and final outcomes.
- Credible identification strategy for each research question.
- Potential for learning, e.g., by including multiple treatment arms.
- Targeting and recruitment of participants clearly defined, and the number of participants adequate to implement proposed analyses.
- Feasibility of implementation (sample size, intervention, selection of beneficiaries, and country context).
- Evidence of partner engagement and support.
- Potential to influence the design and/or prioritization of current and future development interventions.
- Potential to influence policy design and/or scale-up.

REVIEW CRITERIA FOR CONCEPT NOTES

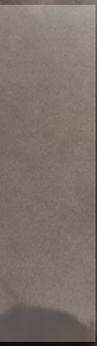
- Hypotheses and research questions clearly linked to the theory of change and have relevance to important research and/or policy questions.
- Main outcomes of interest are relevant to answering research questions and are feasibly gathered.
- Evaluation Design and Sampling Strategy:
 - identification strategy well explained and defines a credible counterfactual
 - design presents no ethical issues or, if so, mitigation measures are highlighted
 - sufficient detail on sample size/power calculations is provided for each of the primary research questions, given available data
- Details on data collection instruments; the data collection strategy is thought out and feasible; and includes information on ethical clearance.
- IE management: research team and implementing partners have sufficient capacity to carry out the proposed research and proposed budget is realistic and represents research value-for-money.



■ ■ **Figure 3.4** Quality assurance during project cycle



policy decisions throughout the cycle of implementation (see chapter 7.2 on monitoring). At completion, final product reports and working papers are submitted to i2i for technical review, following a similar review process from the Concept Note stage. After i2i clearance, teams proceed with Bank review of its final product and then delivery to client and publication. Crucially, all data is expected to be made available for public use through the World Bank’s Microdata library.

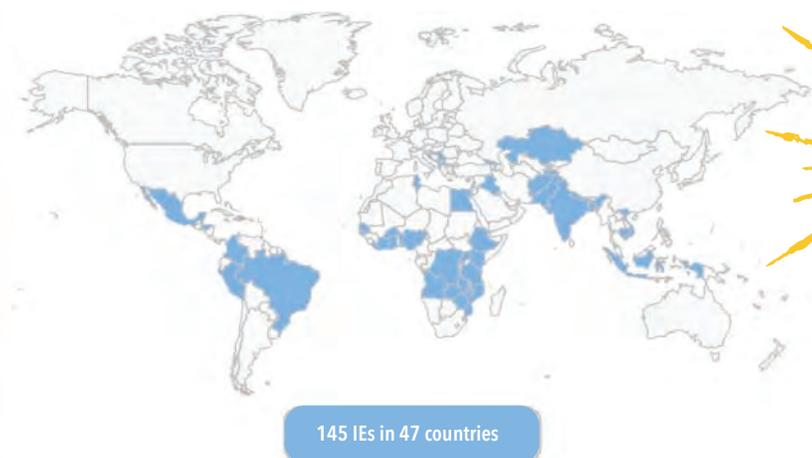


MEETING OUR TARGETS

The i2i program has met and over-delivered its targets for FY18. The results framework in appendix 1 shows progress towards all the targets, while the rest of this chapter provides a summary of some of the main deliverables for impact evaluation (IE) products, capacity-building workshops, and dissemination. In addition to the deliverables defined in the results framework, the i2i has gone beyond to deliver a number of additional outputs that provide great value to maximizing the impact of our program and benefiting the development community more broadly, such as DIME Analytics (discussed in chapter 7.1) and MindMyIE monitoring system, including a third-party verification on policy influence (chapter 7.2, 7.3).

The i2i portfolio has 145 IEs across 47 countries, covering all Bank regions and sectors. Figure 4.2 shows the organization of i2i topics

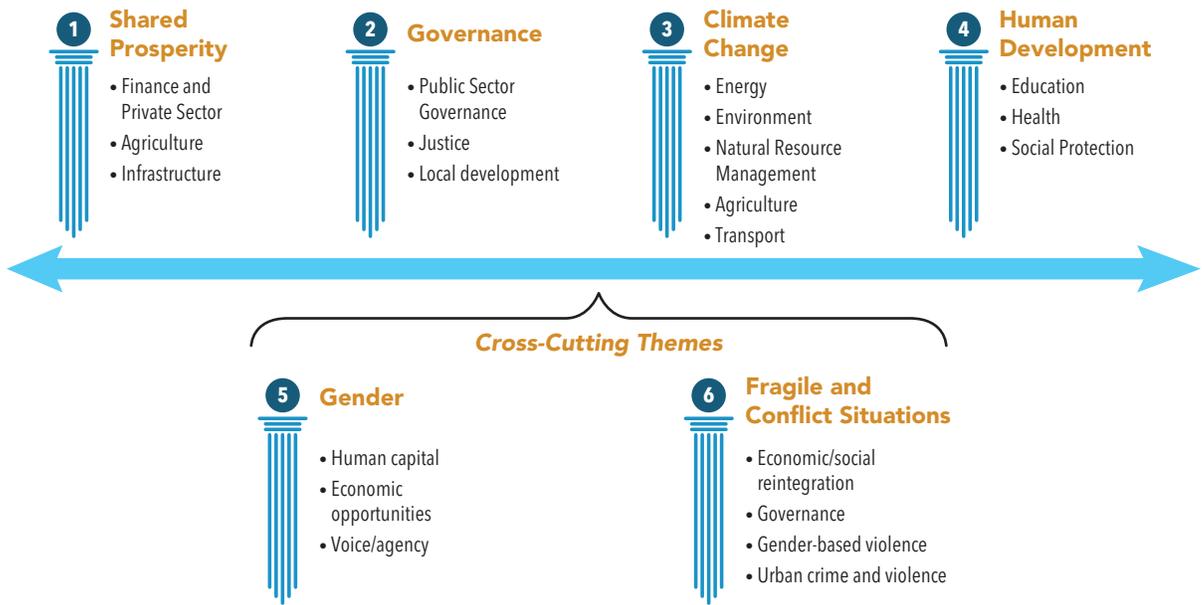
■ ■ **Figure 4.1** The worldwide i2i portfolio



Last Updated: January 2018

*Due to world wide status, the IE "IE Helps Deliver Development Projects" is not included in this report

■ ■ **Figure 4.2** i2i thematic pillars

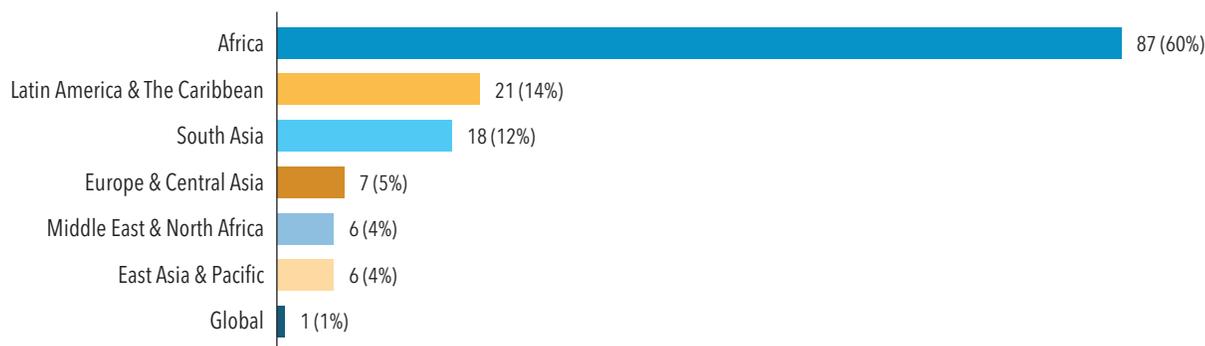


around thematic pillars. The broader thematic agenda was developed through close collaboration with the Bank's Global Practices and projects identified through global workshops and competitive Calls for Proposals (see chapter 2 and 3 for more background on thematic program development, workshops, and selection of proposals).

Between FY14 and FY18, the program completed 11 Calls for Proposals in the areas of: (i) Fragility, Conflict and Violence, (ii) Agriculture, (iii) Energy and Environment, (iv) Governance, (v) Trade & Com-

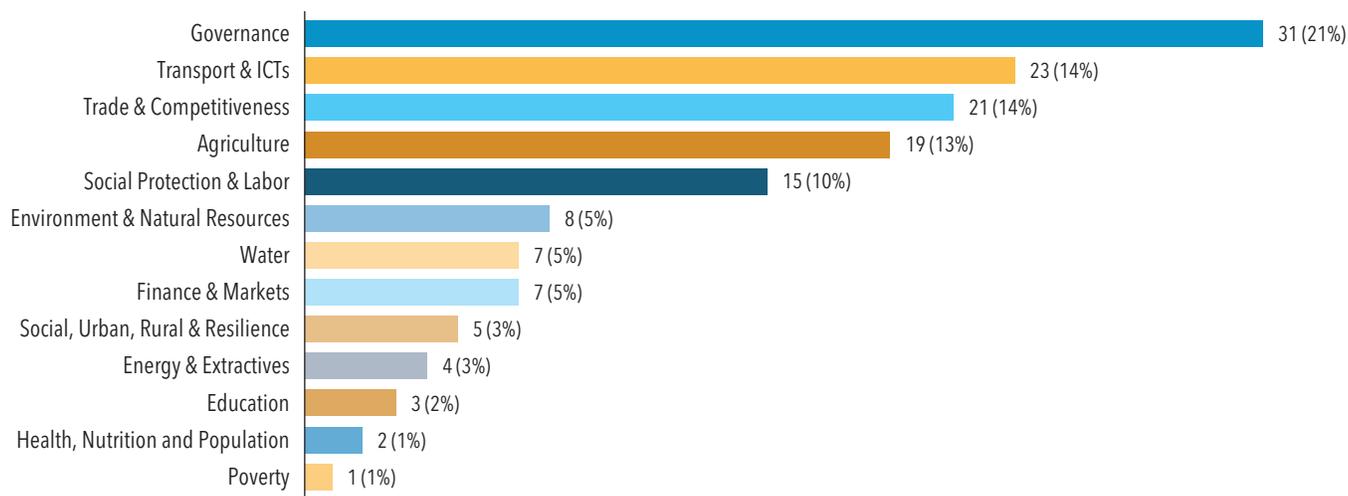
petitiveness, (vi) Transport & Digital Development, and (vii) Legal Identity, and a cross-cutting window for ongoing impact evaluations. Most of these programs have moved into the second and third rounds of program development, increasing the number of IEs through follow up workshops and Calls for Proposals, and expanding the donor base. Much of this work has reshaped Bank's Global Practice approach, providing evidence on what works, what does not, and helping inform future operations. Figures 4.3 and 4.4 show the distribution of the i2i portfolio by region and thematic areas.

Figure 4.3 Distribution of the i2i portfolio by region



Last Updated: January 2018.
The sum percentages may not be equal to 100% due to rounding off.

Figure 4.4 Distribution of the i2i portfolio by Global Practice



Last Updated: January 2018.
The sum percentages may not be equal to 100% due to rounding off.

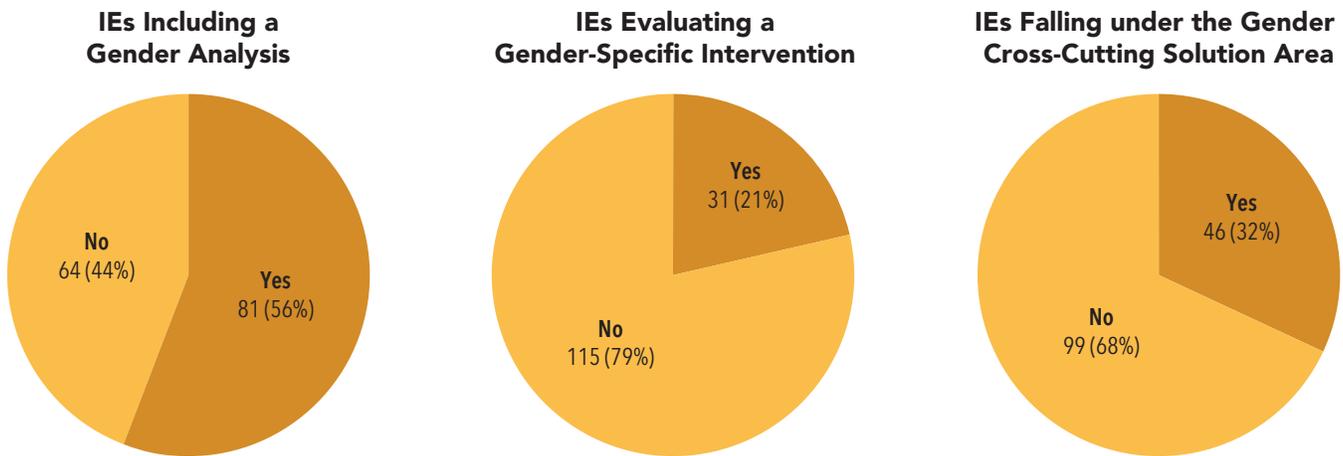
Targets for gender and fragile and conflict affected situations (FCS) have also been met. Gender and FCS are core themes of i2i work. As figures 4.5 and 4.6 show, at least 21% of the portfolio evaluates a gender-specific intervention and 56% of the portfolio conducts disaggregated gender analysis. Thirty-two percent of the portfolio falls under the Bank's Gender Cross-cutting Solution Area, 17% take place

in FCS countries, and 27% in FCS-affected situations.

In terms of design, 68% of the IE portfolio adopt experimental methods, 16% use both experimental and non-experimental, and another 16% have non-experimental design.

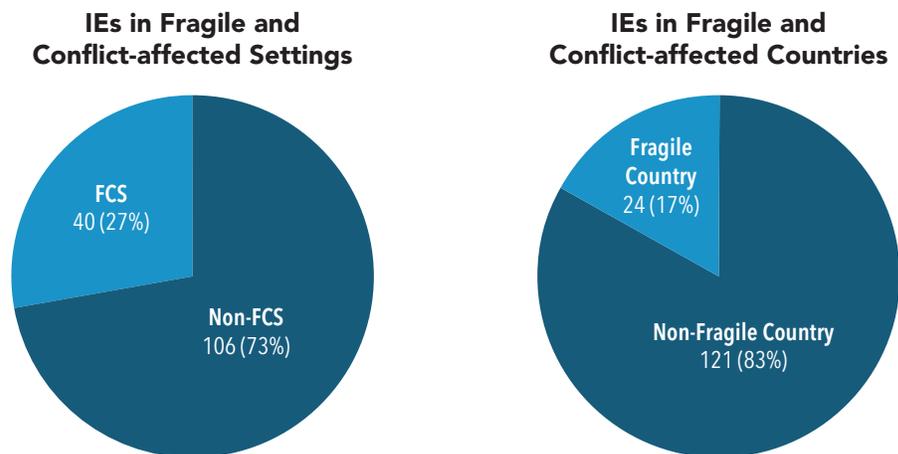
Progress of portfolio implementation is also on track. As of January 2018, 28% of the portfolio is

■ ■ **Figure 4.5** Distribution of the i2i portfolio by gender



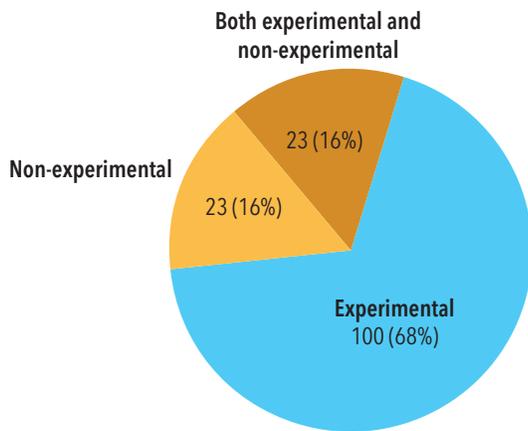
Last Updated: January 2018.
The sum of percentages may be not equal to 100% due to rounding off.

■ ■ **Figure 4.6** i2i portfolio in fragility, conflict, and violence



Last Updated: January 2018.
The sum of percentages may not be equal to 100% due to rounding off.

Figure 4.7 Distribution of the i2i portfolio by evaluation method



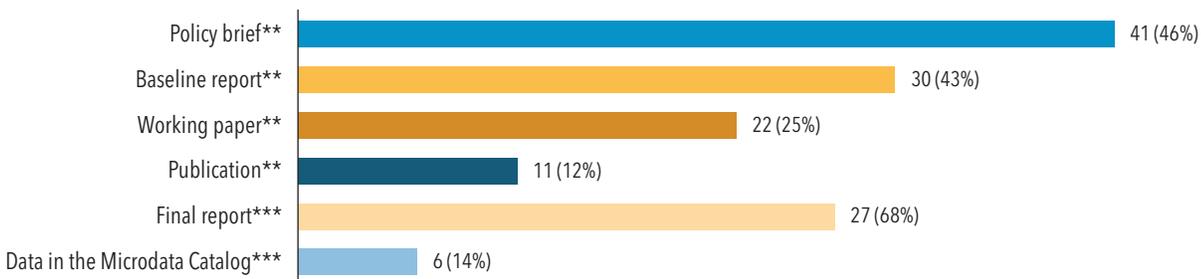
Last Updated: January 2018.
The sum of percentages may not be equal to 100% due to rounding off.



completed (41 IEs), 51% is under implementation (75 IEs), and a remaining 21% is under preparation (30 IEs). All IEs are expected to produce working papers when they complete, as well as a short policy brief that summarizes study results and policy implications. Depending on counterpart demand, some IEs choose to generate additional products, such as baseline reports and final IE reports. The figure below shows the availability of these products at this stage of the portfolio.

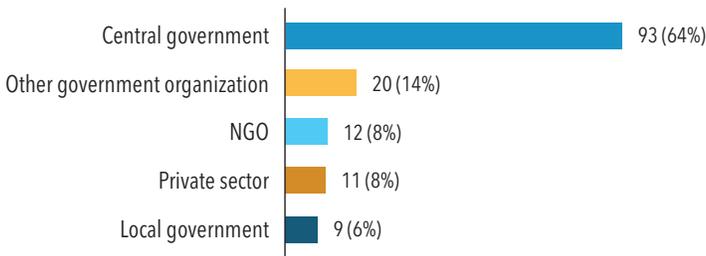
Close counterpart collaboration from early phases of IE design is at the heart of i2i's operating model. Figure 4.8 shows that the central government is the main counterpart in over half of the portfolio, and other local government agencies represent another 14%. Engaging counterparts to work together with development researchers empowers countries to find their own solutions to the most critical policy questions and helps build feedback loops to

Figure 4.8 Portfolio products



Last Updated: January 2018.
* Applicable after CN review.
** Applicable after baseline results are available.
*** Applicable after final results are available.
The sum of percentages may not be equal to 100% due to rounding off.

■ ■ **Figure 4.9 IEs by type of main counterpart** —



Last Updated: January 2018.
The sum of percentages may not be equal to 100% due to rounding off.

influence policy. As chapter 7.3 on policy influence shows, this model has achieved great impact on the ground, generating evidence used to improve project design and support project adoption and scale-up, as well as influencing other projects outside of the IE itself.

In addition to supporting IE products, the i2i program has committed to delivering 14 IE workshops for the period of FY14-19. The i2i has reached and overdeliver this target, completing

21 workshops in close collaboration with the Bank's Global Practices and other external partners (see chapter 2.2 for more detail on workshop design, partner collaboration, and a list of Impact Evaluation workshops). During the first two years of the i2i program, workshops focused on thematic program development, capacity building, and portfolio selection. Audience in these workshops comprised of researchers and subject experts from leading universities, government policy makers, operational staff from the World Bank, and representatives from other MDBs and donors. Following workshops, Calls for Proposals were organized to invite teams to submit proposals for IE in select thematic areas. During the last two fiscal years, the nature of workshops has shifted from program development and capacity building of government counterparts to strengthening the research designs of the selected portfolios in each thematic area and reviewing and disseminating evidence. The target audience, therefore, shifted towards the



research community, subject experts, and MDB operational staffs.

Further, as the table below shows, all targets have been met for people trained, participation of organizations, and dissemination to policy makers. People trained in i2i workshops and by i2i project teams form the i2i network, which as of today includes over 2,000 representatives and over 400 different organizations (donors, MDBs, government agencies, academic institutions, and NGOs). Teams also put a strong emphasis in sharing results with government counterparts as they become available. The latest data from

the i2i monitoring system shows that 90% of baseline results and around 80% of the final results from IEs have already been discussed with the clients. In addition, around 193 dissemination events at the country level have taken place to share results and align counterparts and broader set of stakeholders. Client engagement continues throughout all stages of project cycle, including joint analysis and discussion of the baseline data to develop technical capacity in the government and inform intervention design. These events are highly valuable by clients and there is demand for more in-country training and knowledge sharing.

Table 4.1 Training and dissemination targets

Indicator	Target FY18	Actual
Number of people trained in i2i capacity building events or by i2i-supported teams	800	2,785
Number of organizations, universities, research centers, and government agencies participating in i2i networks	165	322
Dissemination events to policy makers	50	193
Seminars in the World Bank	45	73







HOW WE CONTRIBUTE TO DEVELOPMENT ISSUES

5

5.1 Growth

Productivity is the key driving factor for long-term and sustainable growth. The empirical research, specifically at macro level, has mostly emphasized the determinants of productivity growth. Empirical research has mapped innovation, human capital, and institutions as some of the key determinants of total factor productivity (Syverson, 2011). However, much less is known about how to increase productivity. In other words, what policies work to improve skills of the working force? What reforms and regulations enhance business environment? And what incentives are needed to spur innovation? The *how to design effective policies* question is key, given that many developing countries struggle with low levels of labor and total factor productivity (The World Bank Enterprise Surveys, January 2013).¹

Designing effective interventions that increase productivity requires us to (1) identify ways to increase efficiency of factors of production that are either idly used or misallocated and (2) find alternative combinations of factors of production or technologies that result in higher growth potential.

Productivity growth can occur via both efficiency gains (the “catch up effect”)—changes in the production process to help firms move closer to the efficient production frontier—and shifts in the production frontier. In many cases, adoption of new technologies is required to overcome inefficiencies due to resource misallocation and/or behavioral biases.

¹ <http://www.enterprisesurveys.org/>

“CONEF [the National Committee for Financial Education for its acronym in Portuguese] decided to scale up the [financial literacy] program to public high schools in Brazil.”

IE CLIENT, BRAZIL FINANCIAL LITERACY

In this context, *technology change* encompasses any shock in the production process that leads to higher output, given the inputs available. That shock could be caused by, for instance, better use of land, better trained employees, better managerial practices, reduction in red-tape costs, and change in organizational culture. Technology change and, thus, productivity growth go hand in hand with technology adoption. Shifts in the production frontier assume factors of production are already used optimally, but empirical findings suggest that this is more the exception than the rule. DIME’s portfolio supported by the i2i program reflects that.

The DIME team uses rigorous evaluation methods to test different policies and interventions aimed at increasing productivity and growth. Even though several impact evaluations in DIME operate at a micro level, the close partnership with both operational teams at the World Bank and government counterparts help with scaling up good practices. DIME’s work then plays a crucial role in building knowledge, improving program designs inside and outside the Bank and the quality of policy recommendations.

Currently, there are 50 IEs in 26 countries that address issues related to productivity growth. US\$1.7 billion has been allocated to this agenda through World Bank lending operations and US\$20 million has been allocated to fund IE research. The current portfolio tests different strategies aimed at increasing productivity in both rural and urban settings and among micro-, small- and medium-sized firms.

With farmers in Bangladesh, Malawi and Mozambique, DIME tests gender and social network dimensions of agricultural extension to increase dissemination and adoption of new technologies. With small- and medium-sized firms, DIME tests different ways of improving management practices in multiple settings to affect the industry of service provision to private enterprise. The Brazil portfolio ranges from Rio de Janeiro, where DIME partners with SEBRAE to work with 2,000 small firms (of 5 to 20 employees) to understand whether lack of information is a barrier to adoption of best business practices, and whether a more intensive “handholding” approach is needed to spur adoption; to Paraná, where we work with 1,200 SEBRAE small client firms to test different ways of delivering benchmark information to affect adoption of good managerial practices and improve performance outcomes; and to the North-East where, in partnership with Banco do Nordeste, medium-sized firms are randomly assigned to receive feedback information and intensive on-the-job training and monitoring visits from consultants for six months to spur adoption of good business practices. At the level of individuals and households, we test the use of financial education to affect intertemporal decisions of the population and improve financial planning. A large financial literacy pilot program on knowledge and adoption of improved financial decisions by high school students revealed the potential of high



quality financial literacy interventions on youth behavior and the enhancing effect of co-production when parents were made aware of the material. A portfolio of financial literacy work continues in collaboration with financial regulators and the Central Bank.

Work that focuses on shifting the frontier is quite diverse and focuses on infrastructure investments, from broadband, to roads and corridors, electricity network, and large-scale irrigation schemes. Access to broadband may be associated with enhanced market access, and we are testing this idea in Georgia, where we evaluate the rollout of broadband internet across the country to test the impact of the broadband expansion on firm performance, and whether an intervention combining training on e-commerce and

demand shocks on first online orders increase firms' access to markets.

Major roadway investments have the potential to improve market access and create economic opportunities in remote areas. In Ethiopia, Kenya, Rwanda, Tunisia, Iraq, and Pakistan, DIME is investigating how major transport investments are affecting production, trade, and export of manufacturing. As reduced commuting costs may also encourage population movements, capturing the full welfare implications of road infrastructure development requires careful documentation of these dynamics. In Tunisia, DIME is looking specifically at these dynamics and at which complementary interventions support economic growth through improved market connectivity. The project in Pakistan seeks to measure

the impacts of a new expressway and its complementary investments on transport, trade, private sector investment, market integration, and household welfare. In Ethiopia, DIME is studying the impact of the construction of an expressway which will connect Ethiopia's largest industrial park to export markets, reducing travel times and transport costs for the international garment manufacturers that have invested there, while also evaluating complementary labor market interventions for worker recruitment and training

around the industrial park. In addition, rural transport infrastructure with complementary interventions is linked to the transition from subsistence agriculture to market integration, an area we are studying in Mozambique and Rwanda.

Two-thirds of Nepalese firms identified frequent electricity outages and reliance on back-up options as constraints to productivity growth. DIME examines how long-term improvements in hours of uninterrupted supply through improved



IN MOZAMBIQUE, PROTOCOLS FOR ASSIGNING IRRIGATION SCHEMES REDUCED ELITE CAPTURE

An experiment in irrigation shows that when communities are at freedom to choose where the infrastructure will be placed, they choose more of the larger plots and plot owners (map on the left). When communities are assigned to a protocol that prioritizes small holders, elite capture is reduced (map on the right).



Legend

- Plots
- Kit area
- Eligible area

0 0.05 0.1 0.2 Miles



Legend

- Plots
- Kit area
- Eligible area

0 0.05 0.1 0.2 Miles

Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GISUser Community

Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GISUser Community



physical energy infrastructure promotes firm entry, investments in technology and capital goods, and productivity. We also analyze where the productivity gains of improved energy reliability are largest (e.g., SME versus large firms) and what type of energy supply improvements have highest returns (more hours of supply versus fewer interruptions during hours of supply). This evidence can be incorporated in energy sector assistance strategies around the world to achieve development objectives faster.

In Kenya, Mozambique, Nepal, and Rwanda we are testing the impact of different irrigation

interventions on returns to on-farm investments and land productivity. A special focus of this work is the development of value chains for high-value crops and the agenda of rural transformation. Early results, using spatial regression discontinuity analysis, show that investment in riskier cropping choices sharply increases in the presence of irrigation, as exposure to climate vagaries is reduced. DIME is testing complementary interventions such as subsidies to induce experimentation with new crops, as well as market facilitation interventions on the output and land markets, to ensure farmers can realize gains from trade on the path towards a rural transformation.

5.2 Shared Prosperity

DIME's research contributes to the theme of shared prosperity in several distinct ways. Shared prosperity is a key goal of the World Bank, which has twin objectives of ending poverty and raising the incomes of the bottom 40 percent. The rationale for this is based on the idea that we care not just about the mean of a distribution or the percent below a certain threshold (such as the poverty line), but also about the distribution across the population. A given mean level of income can be consistent with divergent levels of poverty depending on the distribution of income within a society. Moreover, issues of inequality have reemerged on the intellectual agenda in recent years in both developing and developed societies (Piketty 2014; Milanovic 2016). In the development sphere, shared growth is now seen as critical, not just for poverty reduction, but also for the stability and legitimacy of institutions. And inequalities are persistent, not just across income groups, but between marginalized groups, castes, ethnicities, and across gender.

DIME's contribution to this research agenda can be organized conceptually around two fundamental research questions:

- *Pervasive market failure suggests opportunities to improve efficiency AND promote equity—but which interventions work?*
- *When are the objectives of promoting equity and efficiency/growth complements and when are they substitutes?*

We further break this down into a series of sub-research questions:

- Sub-research question 1: How can we reduce physical (transportation) barriers for household and firms to overcome social and spatial inequalities?

- Sub-research question 2: How can basic household services be targeted to sustainably expand access to marginalized populations?
- Sub-research question 3: How do social norms contribute to gender inequality and what are the growth and equity opportunities in overcoming these inequalities?

Theme 1: Improving Access to Jobs and Social Services for All

The importance of physical linkages to markets and jobs is a central issue in development economics. Recent contributions have underlined the magnitude of this in developing countries: for example, Wantchekon and Stanig (2015) estimate that transport costs are the main drivers of poverty in Africa. In last year's annual report, we focused on rural roads and increased access to markets, but inequality in access can be just as important in urban settings. While public transit is usually presumed to benefit the entire urban population, in many cases disadvantaged groups miss out on the benefits. This is often the case for women, whose daily mobility patterns may not be conducive to using public transport, or who face other constraints that prevent them from being able to utilize the public transit system. It can also be the case for socio-economically disadvantaged individuals who are unable to afford the transit fares and can miss out on job opportunities further away. Additionally, new public transit systems can lead to increased inequality if housing prices around the transit system increase and poor households are pushed further away, leading to decreased access to jobs and services for the displaced communities. Without complementary investments to enhance equity, even badly needed urban transport investments can entrench inequality across groups.

DIME researchers are exploring these issues around access to the benefits of urban transport through evaluation of new Bus Rapid Transit (BRT) systems, which are being constructed in many cities across Africa and throughout the world. In addition to studying the direct impact of the BRTs, DIME teams are testing policy responses and interventions that could maximize the inclusiveness of new transit systems, particularly for women and low-income households.

Specifically, DIME researchers and external partners are evaluating the first operational BRT in East Africa, in Dar es Salaam, Tanzania. This project aims to understand how the BRT has affected the lives and welfare of residents and the impacts on urban neighborhoods, development and growth. It further seeks to understand the mechanisms through which these impacts happen, and whether it is through greater mobility or through relocation of people and firms. This assessment is done using a spatial triple-differences approach to isolate the causal effect of the BRT. In addition, experimental methods are being used to study policies that could complement the BRT investment to ensure equal access. For example, in one experiment, cash transfers to encourage some residents to remain in rapidly developing neighborhoods despite rising prices and rents will generate causal evidence on the BRT's impacts on relocation, urban development, and access to employment.

Women and low-income households often face additional barriers that prevent them from accessing urban opportunities. Women who do not have access to a safe means of transport are limited in their ability to access health, educational, and employment services. DIME is exploring potential evaluations of the impact



of complementary interventions in the BRT system on gender-based harassment in public transportation and also evaluating interventions to reduce gender-based violence in public transport in Brazil, as discussed in more detail under Theme 3. We are also studying equitable access across socio-economic groups through changes in fare pricing for a BRT that will be constructed in Dakar, Senegal.

While this section focuses on urban transport accessibility, DIME continues to expand its work on rural roads. For example, DIME research teams are broadening the scope of an ongoing impact evaluation of rural roads in Rwanda to cover roads rehabilitated across the majority of the country, while other teams are continuing to study the impact of road rehabilitation in Peru and are beginning a new study of the impact of road rehabilitation, with a particular focus on women's transport access, in Guinea-Bissau.

Theme 2: Access to Basic Household Services

A second theme relates to access to basic household services. Most current evidence focuses on the impacts of service provision (water, sanitation, electricity, broadband), but there is relatively limited focus on efficiency/equity issues. At the same time, huge inefficiencies exist (for example, US\$89 billion worth of electricity and 32.6 trillion liters of water are estimated to be lost each year to theft and low-quality infrastructure). Private delivery of basic services may be more efficient than public delivery (for example, Galiani et al 2005), but this raises equity concerns. The marginal costs of expanding services to urban and rural poor is often higher than willingness to pay (WTP), and utilities face fiscal

pressure and uncertainty. Even when services are delivered to poor areas, the transitory nature of urban slums means that this may simply result in crowding-in higher income populations, rather than providing targeted support to the neediest. Therefore, three key policy questions are: (i) *How can we effectively expand services to the poor?*; (ii) *Is public intervention justified/required?*; and (iii), *Even when services are expanded to poor areas, are the intended recipients the ones who benefit?*

The energy sector, for instance, is experiencing a technology-driven revolution which is influencing how policy makers and practitioners think about expanding electricity. Low-cost, adaptable off-grid technologies, typically offered through the private sector are providing a solution to the



fact that expanding grid services to remote poor populations is often not financially viable. As the capacity of off-grid solutions increase, this may provide cheaper, cleaner alternatives even in areas with grid accessibility, further increasing access to geographically and economically marginalized populations. Yet, regulations and government support still favor traditional grid expansion with prohibitively high connection costs that exclude the poor. i2i-Supported work in Kenya has helped to estimate the gap between grid expansion costs and willingness to pay to highlight the fact that public subsidies are required to expand access, while also measuring the impact of rural electrification on welfare to assess whether such public intervention is justified. In contrast, ongoing work in Senegal is helping identify

interventions that can support the private sector to more rapidly expand high-quality solar technologies to poor households and explore whether market failures limiting widespread adoption can be overcome.

Other ongoing DIME projects in this area provide subsidies to promote water/sanitation services (Kenya) and broadband (Georgia). Varying subsidy levels allows each study to trace out the demand curve for each service and the resulting benefits that accrue through access to the service. This provides the information needed to estimate the implied weighting of benefits/equality needed to justify intervention. Other work in Kenyan slums takes this a step further by measuring changes in rent, as well as in- and out-migration resulting from provision of sewerage connections to households, in order to estimate the effect of the intervention on housing markets and resulting gentrification.

Theme 3: Gender Equality and Social Norms

A final area of focus is gender inequality and social norms. In this realm, taste-based discrimination and social norms limit the participation and productivity of marginalized groups (gender, race, ultra-poor, etc). There is strong underlying theory and empirical evidence across multiple sectors that this leads to both inefficient and inequitable outcomes (for example, Sen 1992, 2001). DIME IEs focus particularly on gender and ask: Where and how big are the market failures resulting from discrimination and what interventions work to reduce them?

An impact evaluation of extension services in Malawi (Benyishay et al. 2016) shows that women adopt new technologies better and retain



knowledge as well as (or better than) men, but women are less trusted as teachers. This leads to a perception among communities that female extension agents are less effective than they actually are.

A second IE aims to quantify the costs harassment of women on public transport, in the context of gender-segregated train cars in Rio de Janeiro. This study uses a novel approach to estimate i) compliance with the women-only law, ii) willingness to pay for female-only cars and iii) levels of wellbeing in the different cars. Women who commute on Rio de Janeiro's Supervia, a suburban train system, report about their rides on a smartphone app in return for compensation. The IE team randomly assigned them to ride on a train carriage designated for women only to test whether this reduced their experience of harassment. In a separate stage of the experiment, women were offered the option to choose which carriage to ride, varying the payment across types of cars to estimate their willingness to pay to ride the women-only car. A potential unintended consequence is the reinforcement of the social norm that women are responsible for avoiding harassment. This could potentially act as an alternative mechanism for willingness to pay if women are willing to pay to avoid social stigma. The study will capture potential social stigma through a platform survey, including Implicit Association Testing on these issues.

Beyond social norms, inequality along gender lines can be exacerbated through various supply- and demand-side constraints. DIME's program of work also explores these broader underlying factors and their influence on gender disparities across multiple settings. For example, an impact evaluation in Rwanda provides promising evidence of how simple, low-cost innovations can improve

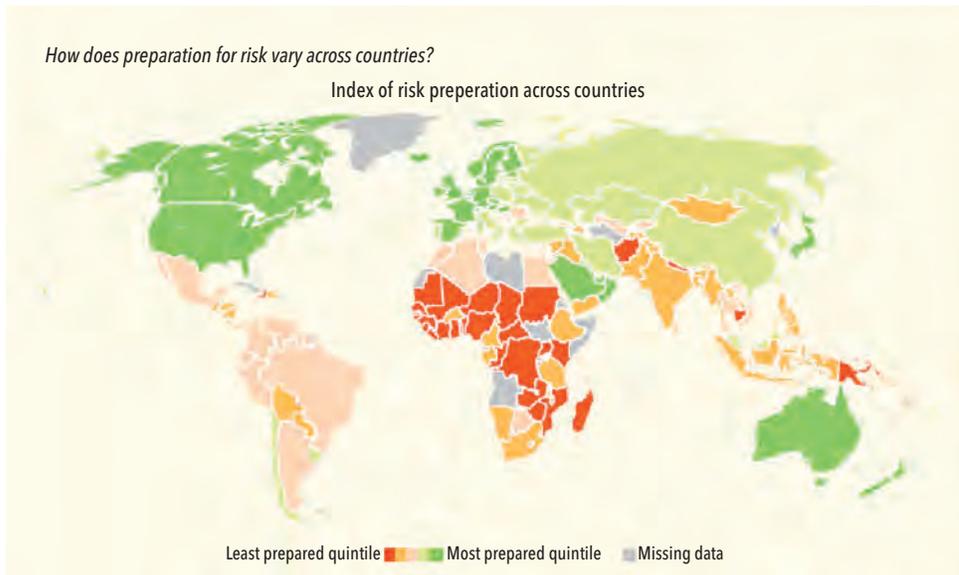
gender equality in terms of access to service. The impact evaluation tested the introduction of simple feedback tools for farmers. At baseline, only one in five women participated in extension trainings; asking for feedback substantially improved perceptions of service quality and more than doubled participation. This may result from a combination of demand-side constraints, as women perceive the benefits of training to be low or face higher opportunity cost and supply-side inefficiencies, as providers have little information as to what training more marginal farmers require and, therefore, face higher acquisition costs in reaching this.

5.3 Risk and Vulnerability

In addition to targeting poverty, we consider the extent to which the circumstances of individuals, households, communities, and countries fluctuate over time. Indeed, variation in the face of an otherwise sustainable, even comfortable livelihood can mean the difference between wellbeing and destitution. Dercon and Shapiro (2007) estimate that 30 to 60 percent of the world's poor live in transitory poverty, meaning that they may move in and out of poverty according to fluctuations in their situation. For the chronically poor, unmanaged negative shocks have long-term negative consequences (Dercon et al. 2005).

The importance of mitigating risk and reducing vulnerability in promoting global development is emphasized across the Sustainable Development Goals. For example, the first goal—to end poverty in all its forms everywhere—includes the target of building the resilience of the poor and those in vulnerable situations and reducing their exposure and vulnerability to climate-related extreme events and other economic, social, and environmental shocks and disasters.

■ ■ **Figure 5.1** Global risk preparedness



Source: Risk and Opportunity, WDR 2014.

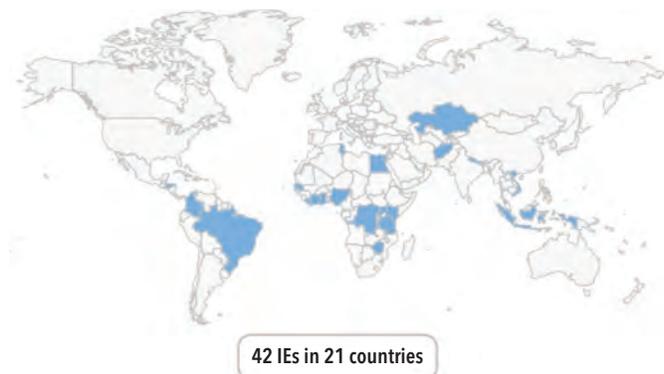
This highlights the many sources of risk and the interconnectedness of poverty and vulnerability. The relationship between poverty and vulnerability is further emphasized by the 2014 World Development Report on “Risk and Opportunity, which finds that countries with high incidence of poverty are some of the least prepared to deal with the risks that threaten so many livelihoods: Sub-Saharan African countries are the least prepared to manage risk, followed by southern and southeast Asia and Latin America (see figure 5.2).

Sources of Risk

Risk comes from multiple sources, some unique to individuals and households (idiosyncratic) and others common to a larger area or population (covariate). Crises due, for example, to health or weather shocks, loss of livelihood, or violence contribute to pushing households below the poverty line and to keeping them there. These negative shocks are likely to be most detrimental to

those living just within their means. Additionally, the anticipation of a potential negative shock and realization that this may lead to destitution may constrain decision-making and may bias individuals and households towards lower-return investments and livelihoods. In other words, the potential for crisis can affect aspirations, preferences, and behavior in education, labor, and social markets, and similarly constrain firms towards pursuing activities with low but relatively predictable

■ ■ **Figure 5.2** DIME Risk & Vulnerability Portfolio



returns. In this way, even the presence of risk, when not properly managed, can be a source of stagnation and poverty.

Constraints to Risk Management

Some barriers to risk mitigation and management relate primarily to individuals, while others relate primarily to broader issues, such as missing markets, absent institutions, and social or economic externalities. In this chapter, we focus more on the former issue: helping individuals to mitigate and manage their idiosyncratic risk over time (see chapter on Governance and Accountability and chapter on Global Public Goods and Externalities for descriptions of work that seeks to tackle sources of covariate risk). Research within DIME seeks to release the following three categories of constraints:

- Reducing **information constraints and cognitive-behavioral biases**.
- Reducing **resource constraints**, both to avoid destitution and to unlock the productive potential of poor and vulnerable groups.
- Resolving **collective action and coordination failures** resulting from perverse or incompatible incentives between multiple individuals.

“The impact evaluation has inspired the client, MISFA, to conduct its own impact evaluations of the TUP program in other provinces not covered by the study.”

IE CLIENT, TARGETING THE ULTRA-POOR IN AFGHANISTAN

Fifteen DIME economists currently contribute research to issues of risk and vulnerability through 42 impact evaluations in 21 countries across Africa, Asia, and Latin America. The distribution of DIME’s risk and vulnerability research is concentrated in the geographies with lowest risk preparedness, as diagnosed by the 2014 World Development Report (see figure 5.2). The risk and vulnerability portfolio represents around US\$33 million in research across projects of a total value of US\$2.1 billion.

Theme 1: Limited information and cognitive-behavioral biases

Limited information leads to sub-optimal risk management, since individuals must base their decisions on expected returns (e.g., to investing in a new skill) or expected constraints (e.g., a health shock) without knowing ex-ante the state of the world that will materialize. When expectations deviate greatly from reality or when the possible outcomes are very variable, this can lead to sub-optimal individual decisions and socially inefficient outcomes. On the other hand, individuals may also act inefficiently despite seeming to have all the information they need to choose the most appropriate action. This may be a result of poor interpretation of available information or of a cognitive-behavioral bias (such as time inconsistent preferences), making the most efficient solution less appealing at the time of decision-making (Duflo 2006). Individuals may also hold cognitive-behavioral biases that predispose them towards certain (re)actions or automatic behaviors, which may be maladaptive (Tversky and Kahneman 1992).

Poor and vulnerable populations may be especially exposed to risk due to limited information and cognitive-behavioral biases and DIME research, therefore, studies mechanisms to

release these constraints. With regards to information constraints, we explore questions such as, “How can we harness mass media to provide information to empower individuals to make better decisions relating to sexual behavior, human rights and democracy promotion, and citizens’ rights to transitional justice in post-conflict contexts?” Results from Nigeria show, for example, that viewers of the MTV Series Shuga, which sought to reduce risky sexual behavior, had lower incidence of STDs, fewer sexual partners, and were more likely to be tested for HIV.

DIME has recently launched a new series of studies of mass-media interventions in fragile and conflict states. In northern Nigeria, we are studying the effectiveness of social norms marketing campaigns aimed at changing attitudes and behaviors around primary education, especially for girls. In Brazil, DIME is studying a school-based intervention that includes video clips aimed at increasing aspirations and reducing teenage pregnancies among vulnerable youth. And in Mexico, a new study will explore complementarities between entertainment-education screenings and school-based information sessions in reducing consumption of alcohol and drugs among middle school youth.

With regards to mitigating cognitive-behavioral biases, our research asks, for example, “How can we leverage investments in developing vocational and soft skills for the labor market to reduce at-risk youth participation in violence and crime?” and “Can psychosocial interventions such as Cognitive Behavioral Therapy reduce impulsive and risky behaviors, like substance abuse or delinquency?” In Mexico, for example, we are testing whether CBT, alone or in conjunction with mindfulness practice, can alter life outcomes for youth gang members, and in Afghanistan, we are testing the



effectiveness of this form of therapy for internally displaced persons. This research explores low-cost, scalable delivery models for mental health interventions in low-resource and low capacity settings which are relevant for a range of vulnerable populations, including those recovering from humanitarian crises, refugees, migrants, and internally displaced persons who may suffer from elevated mental health disorders such as depression, anxiety, and post-traumatic stress disorder.

Theme 2: Resource Constraints

Negative shocks can push households into poverty traps (Sachs 2005). For example, in the case where an indivisible investment is needed to reach a certain level of productive capacity, households with resources below this critical amount are unable to transition from poverty. Resource injections to move individuals over critical thresholds and social safety-net schemes to prevent falling below these are well-known approaches to resolving poverty traps, though whether they do so in a sustainable manner and how best to design and target such schemes remain open questions.

DIME’s work in this area considers both approaches and asks, for example, “How can we support ultra-poor households living in the



context of high-unemployment to pursue trajectories which break poverty traps?” and, “How can we best design social safety-net programs to provide immediate relief while fostering longer-run sustainability, thus preventing poverty traps from occurring in the first place?” In Afghanistan, for example, we are testing the ultra-poor graduation model which provides highly vulnerable households with an integrated package of interventions designed to address several resource constraints, including income and human and productive capital.

Furthermore, DIME’s work in agriculture and food security contributes to greater risk management and resilience to drought, disaster, and malnutrition by testing strategies to increase the adoption of climate smart agricultural practices and assessing the linkages between agricultural productivity and households’ food security and resilience. Using a common set of indicators on food security allows DIME to compare successful strategies implemented both within and across countries. For example, finding that when women are trained to be contact farmers for

promoting sustainable land practices (in Mozambique), other women are significantly more likely to implement these practices. This is an example of how DIME’s agriculture work is advancing the conceptual agenda in how to measure food security and its drivers, producing research on how to measure exposure to seasonality and adoption of sustainable land practices.

Theme 3: Collective Action and Coordination Failures

Coordination failures and collective action problems may occur in contexts where the external consequences of one’s actions do not factor in individual decision-making or in the case where the action chosen by other individuals is uncertain. For example, in the case of disease prevention, an individual deciding whether to have a particular vaccination may not consider the possible contagious nature of an illness in the event of suffering from it. By ignoring the effects their illness could have on others, individuals underestimate the cost of no vaccination and fewer are vaccinated than might be socially optimal.

DIME research evaluates strategies to overcome negative externalities by conveying the marginal social cost of actions and strategies which help individuals collaborate to reduce risks in social, political, and environmental spheres. We ask questions such as, “How can informal institutions be mobilized to help individuals internalize the social cost of their actions?”; “To what extent can information sharing, through public recognition, help individuals cooperate to reduce risk?” and, “How can interest-groups be harnessed to promote welfare beyond their immediate members?” Recent results from Senegal, for example, show that local community-based organizations (CBOs) can be incentivized through non-monetary means to promote environmental protection, which reduce flooding and associated illness and income-loss. And in Burkina Faso, DIME is testing whether CBOs can be motivated to pressure their municipal governments to achieve improved service delivery in the municipality as a whole.

5.4 Governance and Accountability

A society’s institutions or rules of the game are critical determinants of its development outcomes. The design of a society’s governance and accountability structures leads to huge differences in development.² As such, large-scale development efforts require an improved understanding of institutions and the policy interventions that can change them. Governance and accountability have taken centerstage in the search for institutions that underpin development. For example, property rights have frequently

been argued as the fundamental building block to an effective economy. Such institutions span the public and private sectors and are present at each level of socio-economic organizations.

At the theoretical level, a rich body of micro-theory forged around contracting and principal-agent models provide a framework for understanding governance relationships. In its simplest formulation, the principal-agent problem is about designing contracts that can induce an *agent* (an employee, a child, a subordinate division, or an aid contractor) to perform a task as required by a *principal*. This could be within a firm or government office, or between a community and their health workers. When effort and/or outcome are observable and there is no uncertainty on how effort is transformed into outcomes, it is relatively straightforward to design effective contracts. In real world applications, problems arise when neither effort nor outcomes are directly observables. Theoretical contributions have analyzed these trickier situations at great length, including the seminal work of Nobel Prize winners Bengt Holmstrom and Oliver Hart.

However, micro-empirical evidence has not kept up with theory. The evidence-base on many governance institutions is limited. The approach of the DIME research program in Governance and Accountability is to start from the classical principal-agent framework to produce evidence on the impact of governance and accountability mechanisms that are key to development outcomes. As these mechanisms can allow self-interested agents and principals to reach a more cooperative equilibrium, our approach ultimately also explores novel areas of collective action.

We focus on information, regulation, monitoring and enforcement mechanisms. Also, we study governance and accountability problems

² See for example North (1991) and Acemoglu and Robinson (2012).

“KePSIE helped show us

that the system can work.

In one sense, joint inspections

are a good idea and there’s

no question about that.

In theory, it should work,

but in practice, can they

actually work based on the

policy environment?”

IE CLIENT, KENYA PATIENT SAFETY IMPACT EVALUATION

with an approach that recognizes their systemic nature. Finally, we study the principal-agent model through multiple-lenses, which span theory, measurement and intervention.

Information, Regulation, Monitoring and Enforcement

These mechanisms act as commonalities across our research program. We study them in a variety of contexts to assess whether they can be effective in addressing one (or more than one) dimension of the principal-agent problem. Shifting societal rules and norms can affect the observability of effort and outcomes and the type of agents that enter contracts with the public sector. Regulatory reforms that introduce new procurement processes (in Bangladesh, Brazil, Colombia, and Philippines) and introduce minimum patient and building safety (in Kenya and Peru) are studied to ascertain whether they change the outcome

(in terms of quality and efficiency of the service delivered) of the principal-agent game.

Providing information to key actors is predicted to reduce information asymmetry, which is one of the main issues within the principal-agent framework. The research program is testing the impact of providing information on court efficiency to judges, registrars, clerks and users (in Azerbaijan, Kenya, and Senegal), as well as information on public works performance to politicians, senior and junior bureaucrats (in Pakistan). More direct mechanisms to increase observability of effort and outcomes work through monitoring and enforcement devices. Several ways of tracking provider performance for bottom-up accountability are being tested as part of the research program (in Angola, Burkina Faso, Cambodia, DRC, and Nigeria).

System-wide Analysis of Governance and Accountability

The systemic nature of collection action problems requires a system-wide approach to analyze them. Predicted solutions from theory (increase information or monitoring) typically apply only to specific elements of the system. There is growing evidence on these elements of the system (e.g., audits and monitoring, rewards and warnings), but still scarce evidence on the system and market equilibrium (price, quality, supply) effects of interventions.

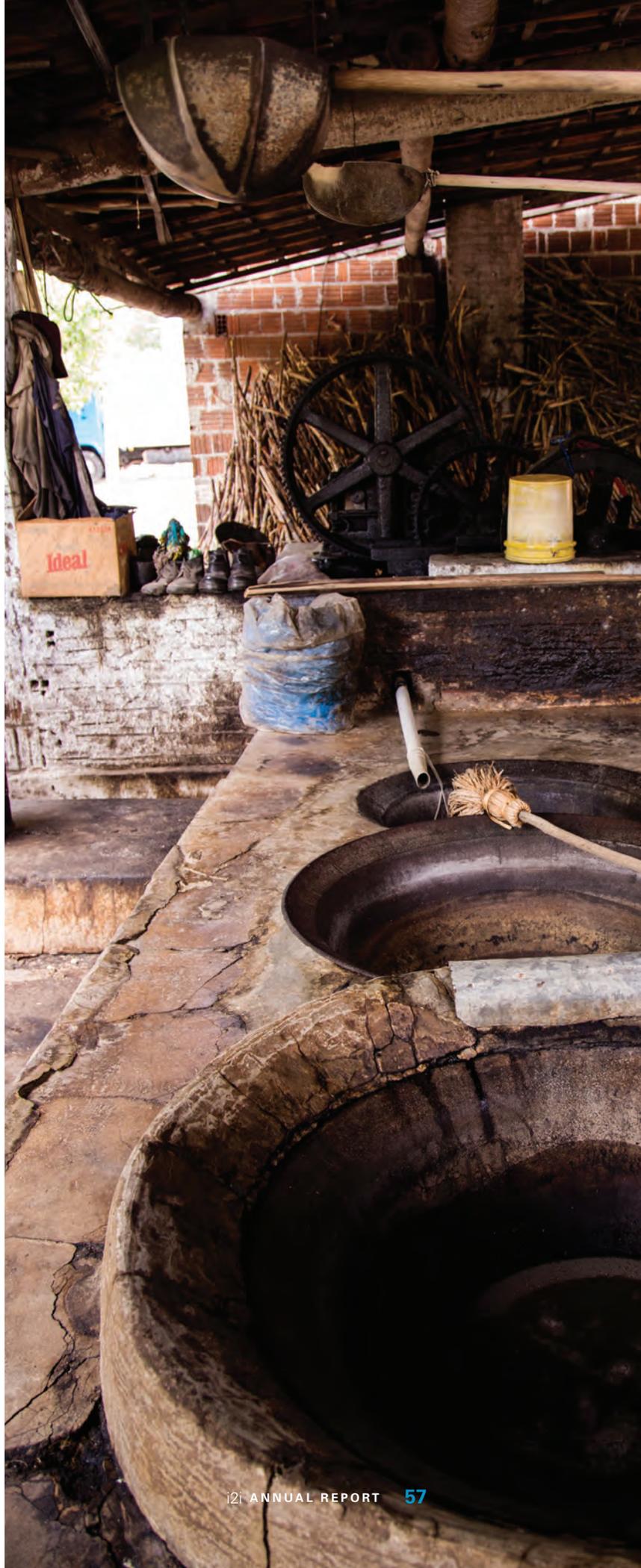
A system-wide analysis is studying how to understand and align incentives in markets for road safety and patient safety in Kenya. In particular, the Kenya research program is studying, within the same market, the impact of shifting formal and informal rules (regulation and enforcement of patient safety and road safety) of providing information (on clinic

performance to clinics managers and patients, and on driver safety to drivers, transport company owners and riders) and of enabling bottom-up and top-down accountability devices for clinics and patients, and bus owners and drivers. In addition to partial equilibrium impacts on service provider performance, the research program will study market outcomes in terms of prices and supply indicators.

Theory, Measurement, and Intervention

Our research program explores the principal-agent model through multiple lenses. These include: (i) theory, which provides frontier predictions that can be tested by our research program; (ii) a focus on measurement to improve observability of efforts/actions and performance and provide data for monitoring/diagnostics; and (iii) an emphasis in designing interventions/treatments that are embedded in theory and test mechanisms.

We test theory predictions in Tanzania as we study the role of characteristics of organizations' structure on solving incomplete contracting in infrastructure maintenance and the impact of performance-based incentives to improve sub-national service delivery under weak monitoring capacity. Specific treatments underlying these IEs stem from the same theoretical framework. They include co-production organization in infrastructure maintenance (to solve incomplete contracts) and performance based-grants to subnational governments (to solve observability of agent's outcomes). On the measurement side, our research program is producing globally comparable micro-data on civil servants, developing systems to track subnational government capacity, and court measurement systems to track judiciary performance (Kenya, Senegal, and Tanzania).



5.5 Global Public Goods and Externalities

Addressing the global challenges of natural resource depletion, climate change, and health epidemics requires the provision of global public goods. Theory and experience show these are often undersupplied due to their inherent consumption characteristics. Public goods whose consumption (1) is non-rival and non-excludable, like TV broadcasts, prevent producers from experiencing the full benefit in the presence of free-riding, (2) is rival but not-excludable, like water from an unregulated river, can lead to overuse by farmers who don't bear the full cost of their overuse, thereby (3) create negative externalities borne by others, such as car pollution, can lead to overuse of private transportation.

Economic theory shows that coordination mechanisms, often in the form of financial incentives or subsidies for producers or consumers, are needed to expand the provision of public goods. Development efforts to encourage coordination mechanisms are particularly important for global public goods, where agents are less likely to interact and therefore less likely to jointly invest in global solutions. Take the example of mobile apps or educational TV programs: despite their potential to reach millions of individuals in developing countries at low marginal costs, these information goods remain undersupplied as producers often incur in high production costs and free-riding is common in developing countries. Failures to produce global public goods hampers efforts of governments and development partners to address today's global challenges.

In this theme, DIME evaluations are studying interventions and coordination mechanisms to promote investments in global public goods. For

example, impact evaluations address important knowledge gaps on effective ways to use mobile technologies that monitor and promote efficient use of resources and the cost of under-monitoring; the relative effectiveness of different mass media outlets (for example, TV, radio, print) to promote adoption of new technologies; and the role that entertainment-education can play in addressing global challenges and epidemics, such as HIV/ AIDS and gender-based violence.

The knowledge agenda aims to unpack the causal mechanisms linking media exposure to individual and community-level impacts on knowledge, attitudes, and behaviors. DIME evaluations of agricultural extension systems demonstrate that while advice about agricultural practices may be freely shared between farmers, ensuring that such information actually reaches targeted populations requires careful decisions about how to design information dissemination strategies. For example, the identity of extension workers appears to influence the likelihood of new methods being adopted as shown by an example in Mozambique, where women were 9% more likely to try a new farming practice if they were randomly matched to a female extension worker. Means of disseminating information that entertain the audience are also especially effective. An evaluation in Nigeria revealed that viewers of an MTV-produced soap opera were twice as likely to get tested for HIV and half as likely to be infected with chlamydia 8 months after seeing the show when compared to non-viewers. Chapter 6.8 provides greater details about this evaluation and DIME's Edutainment research program.

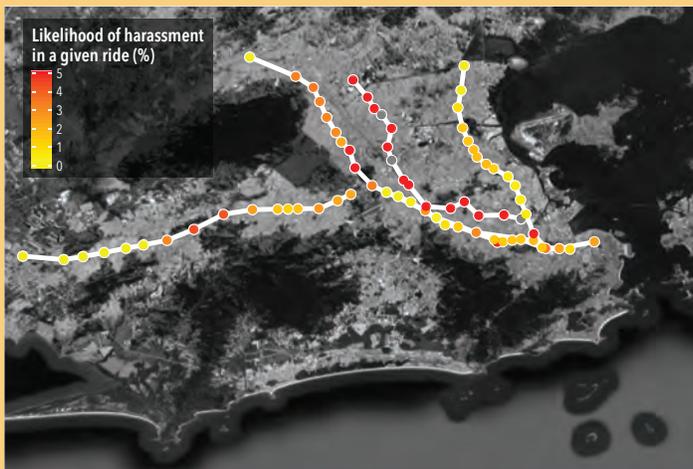
Managing the Commons

Recognizing that many crucial economic goods are common pool resources, DIME evaluations

CARIOCA WOMEN ARE HARASSED ONCE A WEEK WHILE RIDING THE METRO



Sexual harassment affects women's public transport, as well as school and job choices. Sixty to eighty percent of women around the world say they have been harassed on public transport, but most fail to report it to authorities. Thus, we count with little data on the incidence of sexual harassment and its economic costs. We use a crowdsourcing app to measure the incidence of harassment in the Rio de Janeiro train network and explore the correlation between presence of men and incidence of harassment in mixed and women-only carriages. We show that in practice the women-only carriages still contain a substantial share of men – especially in the outskirts of the city where enforcement is low (map on the left) – and that harassment varies with the share of men in the cars (map on the right).





demonstrate and test the strategies to creating and maintaining institutions that manage common pool resources.

DIME's work on this area seeks to test and demonstrate which institutional structures and incentives are effective in managing common goods. There are many examples where individuals either over- or under-perform an action relative to the social optimum because they do not take into account the fact that their own use reduces the available resources for others.

In contexts where a given resource is known to be subject to the problems arising from managing common pool resources, such as irrigation, DIME is testing the question of why existing institutions established to manage the problem fail. For example, in Mozambique, projects generally prioritize the inclusion of smallholder farmers, but when communities can choose who will receive irrigation investments, they tend to choose farmers with larger land size who would have been excluded, suggesting that these wealthier farmers may have a role in improved management.

In Rwanda, a project varies the user fee paid by irrigation users to assess whether these fees are set too high for sustainable participation in water user groups. Finally, DIME works to design and test innovations to institutions to manage the commons to improve these institutions. For example, a project in the Democratic Republic of the Congo is testing accountability mechanisms for service providers working in conflict-affected areas to assess whether these reforms change the private incentives to participate in conflict.

Private Incentives for Adopting Pro-Social Activities and Conservation Technologies

DIME research focuses on uncovering and mitigating the behavioral constraints that prevent people from adopting technologies that would be privately profitable to them, but also have non-rival attributes that create large public externalities. This situation is particularly apparent in the context of production or consumption decisions that may be just profitable or incentive-compatible for an individual or firm to adopt, while at the

same time generating large public benefits. Take the case of environmental conservation, when governments may need to provide incentives to individuals or communities to adopt conservation technologies with large social benefits. This could apply even when technologies may be profitable, due to market failures (for example, imperfect information about the technologies' benefits, unclear property rights), financial constraints faced by households or firms, or psychological limitations that are increasingly being studied by behavioral economists. For example, present bias may prevent people from properly weighing the long-term benefits or savings of adopting new technologies, even when they can afford it and would be better off in the long term by adopting these technologies.

A highlight of DIME's work on these questions is a body of experiments across contexts on the optimal design of payment incentives for sustainable adoption of tree-planting and avoidance of deforestation—a type of Payments for Ecosystems Services (PES) system. For example, the first evaluation in this cluster found that the Mexico PES program, among the largest and most mature programs in the world, reduced the loss of tree cover by approximately 40% within areas at high risk of deforestation. The program also led to a 50% increase in management activities to protect land cover (e.g., patrolling for illegal conversion, building fire breaks, combatting soil erosion). Interestingly, this incentives program had positive but modest improvements in pro-social work effort and in measures of social capital at the communal and household level.

The results of this first study are promising, however, open questions remain on whether gains from the same investment could be larger, or

whether they can be sustained in the long term. With this aim, each new evaluation in this cluster after the Mexico IE not only measures the overall effect of payments in terms of avoided deforestation, but also aims to shed light on these open questions. One evaluation in Uganda assesses upkeep of pro-environment actions after the initial policies encouraging take-up are withdrawn. Two other experiments investigate how best to make offers to adopt conservation technologies. For example, when group dynamics are involved, as in community-based PES, coordination failure or social dilemma can negatively affect the overall outcome. To understand whether group or individual incentives perform best, DIME is implementing an IE embedded in the Burkina Faso Forest Investment Program, a REDD+ project, and testing the impact of alternative contract design options on group performance measured by tree survival rates. A second experiment in Ghana focuses on how offers of payments should be made, finding that virtually 100% of farmers sign-up to be paid to plant trees on their land when directly offered a price, but sign-ups are around 10% lower when the farmers are first asked to think about how much they would need to be paid before being told what the offer is. These results suggest that the way offers are framed can influence whether the most enthusiastic participants are the ones included from a fixed budget for payments.

The knowledge agenda in this theme also covers operational research questions, such as optimal subsidy schemes for purchasing new technologies (e.g., testing a money back guarantee against a rent to own plan for solar lamps in Senegal) and selecting the most cost-effective methods to diffuse new information (e.g., new cultivation strategies among farmers in Bangladesh).



ECONOMIES OF SCALE IN LEARNING

6

6.1 Fragility, Conflict, and Violence

Around 2 billion people live in countries affected by fragility, conflict, and violence (FCV). Poverty rates are 20 percent higher in countries affected by repeated cycles of violence. While trends show that poverty is declining across much of the world, countries affected by conflict are falling behind, with conflicts reducing growth by two percentage points per year on average. The share of extreme poor living in conflict-affected situations is expected to rise from 17 percent of the global total today to almost 50 percent by 2030. Development challenges in FCV settings transcend national boundaries through the displacement of populations, spread of disease, reduced trade, and increased organized crime and terrorism.

The international community is committed to assisting FCV countries to restore social and political order, sustain peace, and resume economic growth. The record US\$75 billion commitment under the World Bank's International Development Association's 18th Replenishment (IDA18) marks a strategic shift in making reducing the risk of fragility and conflict a top development priority for donors across the world. The evidence base for designing such programs is, however, sparse, especially with regards to rigorous evaluations aiming to identify what works and how to prevent and reduce fragility, conflict, and violence. The unprecedented resources for FCV-related work highlight the need for more rigorous evaluation of policies targeting FCV issues. Not only is developing this evidence a priority, but experience to date shows that, even with the amplified challenges of working in FCV environments, rigorous evaluation in such settings is possible. For example, IEs have been conducted

in diverse FCV settings including in Afghanistan, DRC, Colombia, Côte d'Ivoire, Sri Lanka, Rwanda, northern Nigeria, and Liberia.

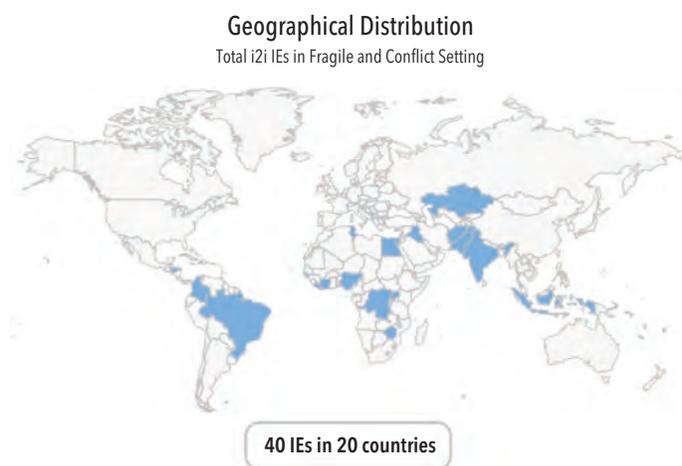
FCV Impact-Evaluation Program

In March 2014, DIME and partners inside and outside of the Bank launched the Evidence for Peace (E4P) program.³ Its overall goal is to assess evidence gaps in FCV responses and to generate improved knowledge about how to best support FCV clients to deliver the results so critically needed for citizens to gain confidence in the path out of conflict. Today, the program includes

40 IEs across 20 countries as highlighted in Figure 6.1. The portfolio represents around US\$30 million of research across projects of a total value of US\$2.1 billion. Further, a series of white papers synthesizing the state of the evidence has been completed in each of the four target themes.

In 2017, much of the work focused on continuing to support the implementation of several IEs in the portfolio and program-level analytical work. This included technical support to the implementation of at least two dozen ongoing IEs and the dissemination of three of the four thematic white papers aiming to summarize the FCV evidence base in the prioritized strategic areas of the program and to disseminate the knowledge through a series of learning events that brought together academics, policy makers, and operational staff of the Bank and other donors to close the gap between evidence and policy. The DIME commissioned white papers on *i) Gender-Based Violence*, *ii) Employment in FCV*, and *iii) Civil Service Reform* were each presented at a dedicated dissemination event.

■ ■ ■ **Figure 6.1** DIME FCV Research Program

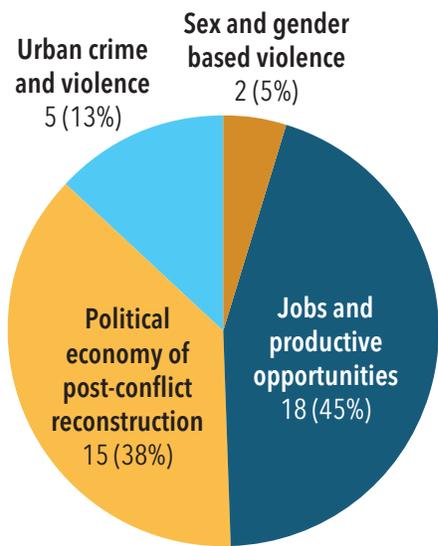


³ Internal World Bank partners included the Fragility, Conflict, and Violence (FCV) Group-CCSA and the Latin America's Citizen Security Team. External partners included the International Initiative for Impact Evaluation (3ie) and Innovations for Poverty Action (IPA). The original initiative included: (i) development of a scoping paper, based on an evidence "gap map" that identifies the status of the evidence base to highlight priority questions for future research; (ii) design, implementation and dissemination of impact evaluations funded within the Bank and through a new external funding window; (iii) development of a framework to improve the quality of analytical work on FCV and of tailored methodologies for evaluation/data collection in FCV settings; and (iv) hands-on training in impact evaluation and the creation of communities of practice for knowledge sharing.

Theme 1: Basic Service Delivery in Weak States

The inability of the state to implement policies, collect taxes, and govern the national territory is a fundamental and pervasive problem in developing countries, especially in FCV states where institutions are likely to be the most eroded and dysfunctional. An important contribution of the i2i agenda is to invest in strengthening the accountability of service providers. The research under this agenda focuses on civil-service reform and on the rebuilding of government capacity and accountability systems and covers the recruitment and deployment strategies, patronage networks, misalignment of incentives, and capacity

Figure 6.2 DIME FCV Research Program by Number of IEs, Status and Topic



issues that hinder the effective delivery of services. This program area links to DFID’s support to the Open Government Partnership that congregates 75 countries to make governments more accountable. The knowledge generated under the program is providing clear policy guidance to build and strengthen government monitoring of the delivery of critical services such as health, education, and justice, a function largely underdeveloped in low-income countries that are DFID’s priority. Work towards service provision in weak states is ongoing in Liberia, Zimbabwe, and the DRC, among others.

In Uganda, large-scale CDD interventions contributed to greater monitoring of programs and increased complaints to local and national officials, and these efforts also improved project quality, as reflected by improvements in health of animals of livestock projects, which were sustained 18 months after the project ended. Meanwhile, in Colombia, receipt of legal services through Mobile Victims Units has increased

the advancement in the administrative process of reparations, as well as increasing the rate at which victims receive assistance, attention and reparation from the state. At the community level, work is also ongoing to evaluate the DFID-funded Tuungane project, which aims to improve the DRC’s state capacity to provide basic services. Results for the first two phases of the program were made available in 2017, which suggest that the program continues to have an impact on infrastructure quality and service delivery, however there is no evidence of an impact on the downstream social, economic and accountability-related outcomes. This set of IEs will generate knowledge on a 10-year CDD program, which will provide policy-ready guidance to lending institutions and policy makers across the world.

“The impact evaluation research team supported MISFA [Microfinance Investment Support Facility for Afghanistan] in developing a framework to monitor the implementation of the project. MISFA is using this framework in all of its Targeting the Ultra Poor projects, also outside of the study areas.”

IE CLIENT, TARGETING THE ULTRA-POOR IN AFGHANISTAN



Theme 2: Improving Job Opportunities for At-Risk Youth

Youth in FCV contexts are often left with few marketable skills and little opportunity to cultivate a sustainable livelihood. They can become vulnerable to involvement in conflict, illicit activities, or violent crime. A second strand of research targets present bias and labor market constraints that affect the crime versus employment choices of youth-at-risk in settings ranging from post-conflict countries to high-violence urban areas. So, in addition to traditional skills and apprenticeship

programs, we also look at novel approaches such as soft skills, psychosocial therapy, and other labor market insertion schemes targeting youth who lack basic foundational skills.

Our program is currently seeking to expand the opportunities for at-risk youth in Côte d'Ivoire, Honduras, Liberia, Nigeria, and the DRC. In Côte d'Ivoire, the Labor Intensive Public Works (LIPW) program had impact on employment, earnings, expenditure, and savings in the short run which dissipated in the medium run 12-15 months after project completion. In northern Nigeria, the DFID-funded Mafita program, which is in the implementation phase, explores the effectiveness of vocational training and apprenticeship programs for marginalized young boys and girls isolated from mainstream society, both in terms of improving economic prospects and in reducing violence and support for extremism in the volatile and crisis-prone Sahel.

Theme 3: Breaking Poverty Traps and Cycles of (Gendered) Vulnerability

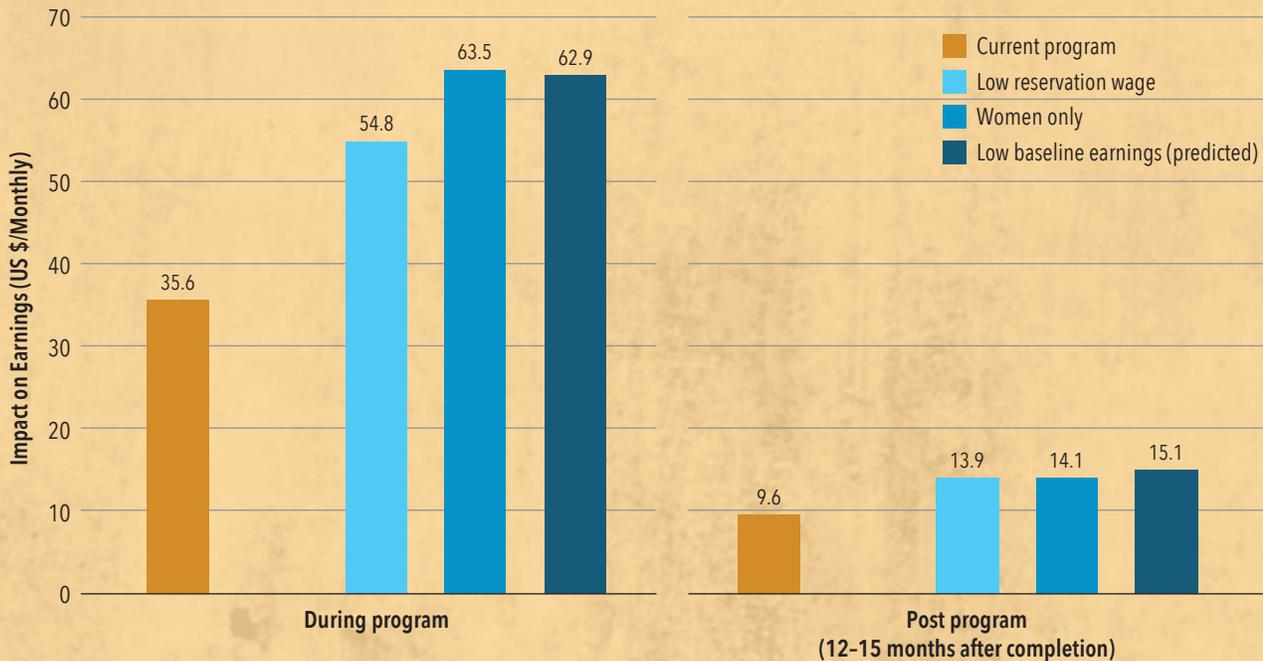
In FCV contexts, support for vulnerable groups is often lacking with both their immediate and long-run needs overlooked and their productive capacity ignored. This can lead to a perpetuation of poverty traps and cycles of vulnerability. Further, evidence shows that poverty-induced vulnerabilities tend to disproportionately affect women and children. A third strand builds on the ultra-poor literature to understand the potential of interventions geared toward breaking poverty traps and addressing systemic vulnerabilities. The research investigates the potential of integrated approaches and effectiveness of social safety net programs to break poverty traps in a sustainable manner unlocking

DIME RESEARCH SHOWS THAT TARGETING MATTERS FOR EFFECTIVENESS OF LABOR INTENSIVE PUBLIC WORKS PROGRAMS IN CÔTE D'IVOIRE



DIME research in labor intensive public works programs shows that who you target really matters for effectiveness both during and after the program. In Côte d'Ivoire the research tested alternative targeting approaches and found that targeting women or people with low baseline earnings maximizes the effect on monthly earnings relative to the control group. The orange bar on the left in each graph shows impact on earnings under the current program specifications (in both the short- and medium- term). The bars on the right show the impact on earnings if the program is targeted differently (three test case scenarios: paying a lower daily wage, targeting women only, and targeting the poorest).

Short- and Long-term Impact on Earnings of Various Targeting Mechanisms



the labor market potential of the ultra-poor. It also explores norms-shifting and targeted interventions to eradicate child labor and address gender-based violence.

Research is currently ongoing in Afghanistan, Comoros, the DRC, Egypt, and Tunisia. This set

of experiments will provide compelling evidence on the basket of features (e.g., the size of cash payments, targeting of populations, and modes of delivery) and on complementary interventions (e.g., capital transfers and mentoring) to secure deeper and longer-term impacts. For example, the LIWP in Tunisia and Community Social

Services LIPW in Egypt had short term positive impacts across economic outcomes of interest including employment, income, consumption, and savings, among other areas. Results from Egypt show that the community social services intervention component yielded positive labor market impacts, leading to a re-design of the US\$400 million scale-up of the program.

Theme 4: Political Economy of Post-Conflict Reconstruction

Postwar societies are often confronted with a wide range of issues which prevent a rapid return to stable social and political equilibria. The political economy of post-conflict reconstruction research area in an effort to understanding the drivers and perpetrators of conflict and on evaluating strategies designed to address

these obstacles studies information asymmetries between elites and masses, social dislocations, and security and mobility constraints. Some of these are root causes of the conflict in the first place or conflict drivers that sustain the cycle of fragility, conflict and violence. Work in post-conflict reconstruction currently focuses on Liberia, Sierra Leone, Zimbabwe, and the DRC. For example, research in Liberia seeks to overcome informational asymmetries and collective action problems in rural areas by providing groups of women with access to a safe space to listen to unbiased political radio broadcasts by United Nations peacebuilders. Results show that overcoming these barriers can increase female political participation in many forms both at a national and a local level, with communities that received the intervention also exhibiting smaller gender gaps across most outcome indicators.



Going Forward

Key findings and accomplishments from the DIME FCV program portfolio, as highlighted above, demonstrate the significant contributions made in generating valuable knowledge on what works, and why and how they do in order to resolve challenges in key FCV areas. These include the production of four white papers which summarize existing knowledge in the four priority areas. These white papers have been used to engage academic and policy makers and practitioners at the WB and external partners in this field. Also, several IEs have generated results and produced working papers, with some already having important policy influence. Going forward, our work will focus on aggregating all the produced knowledge and making it more operationally relevant through an aggressive research uptake strategy. Specifically, this will be accomplished through four distinct activity areas:

1. Sharing produced knowledge at a global stage through 2nd generation workshops and targeted policy engagement action.

The objective of these workshops is to contribute to institutionalizing the generation and use of rigorous evidence for addressing poverty, employment, and violence in FCV contexts. The idea is to present and discuss existing evidence relevant to current FCV policy and operational priorities through small group brainstorming sessions and panel discussions. Project teams with completed IEs or advanced IEs will share knowledge and lessons in an effort to better design more appropriate interventions and programs. Another objective is to match government or Bank program staff with academics and IE experts to develop initial impact evaluation concepts, which address current FCV policy

and operational priorities. Cross-country evidence from specific interventions will be shared through Policy Research Talks (PRT), which target WB policy makers in an effort to influence their thinking and decision-making. A PRT in January 2018 presented the evidence and knowledge gaps on the interactions between employment interventions and crime and violence in FCV contexts to WB operational teams.

2. Aggregating produced knowledge through systematic reviews and strengthening research uptake.

Systematic reviews, meta-analyses, and cross-study comparisons can provide valuable evidence for evidence-informed policymaking. Since the literature on what works and what does not work in FCV countries is starting to expand, and there has been accumulation of IEs in specific strategic themes/topics/clusters, such results provide opportunity to conduct rigorous systematic reviews that will aggregate the knowledge and draw important policy implications for policy makers. Within the DIME portfolio, two thematic areas where IEs have been conducted across several FCV countries include employment interventions, particularly, impact of labor-intensive public works programs and promoting political participation in post-war countries. In addition to the systematic reviews, one possible output is production of topic-specific policy briefs that will make the knowledge readily accessible to practitioners and policy makers.

3. Leveraging existing IEs to deepen the impact of current programs.

One of the objectives of the DIME FCV portfolio is to understand how and why interventions work, in addition to what works to

improve development outcomes in FCV contexts and what does not work. In order to explore the mechanisms or channels by which interventions deliver impact, DIME seeks to incorporate one or more novel treatment arms to existing IEs or deliver new treatments to a subset of participants from the previous IE/program, with the view to generating more and better evidence that can enhance operational learning and policy. Four approaches being used by DIME include:

- **Mainstreaming the effects of gender-based violence (GBV) in employment programs.**

Anecdotal evidence seems to indicate that women's economic empowerment leads to reductions in GBV due to factors such as reduced financial pressure, increase in consultative decision-making between couples and, for some, empowerment within the household. The evidence however to substantiate this relationship is not only mixed and inconsistent, but also severely limited, especially in places affected by violence and conflict. To address this gap, DIME is in the process of designing several cutting-edge initiatives aimed at mainstreaming GBV prevention and response into existing employment programs targeting both women and men. DIME plans to add treatment arms to existing IEs with an aim of measuring behavior change towards intimate partner violence (IPV). In doing so, DIME plans to initiate an education-through-entertainment (Edutainment) campaign that could include: community-based recordings, TV spots, or both. The aim of this intervention would be to raise

awareness on IPV prevention and response, while educating people and promoting positive behavior change. An intervention is being designed in Egypt with the potential of being replicated in other contexts.

- **Adding soft-skills to employment programs.**

Education contributes to developing technical skills of potential employees, however, many youth lack the soft skills needed for success in the workplace, such as how to interact with customers, work in teams, act professionally, and even how to properly represent themselves in job interviews in the first place. Building on studies such as Groh et al (2016), DIME plans to continue experimenting with the soft skills add-ons to the jobs programs in order to examine labor market effects, socio-psychological impacts, and outcomes related to violence and crime

- **Promoting female entrepreneurship through capital injection.**

Given that the income earned through employment interventions such as labor-intensive public works programs is merely sufficient to fund daily consumption and expenditure, the likelihood that some of it is set aside for business investment is very low, especially among women who almost always spend their earnings in consumption, particularly in health and education for the child. To address this gap, DIME seeks to explore interventions that aim to foster female entrepreneurship through small business grants for sub-samples of female graduates of the public-works programs.



IEs that explore whether capital injections strengthen women's economic empowerment and sustain livelihoods in the long term are being tested and/or designed in the DRC, Indonesia, Tunisia (and possibly Egypt).

- **Enhancing the economic and social reintegration of refugees and IDPs in labor markets of host communities.**

Recurrence of conflicts has led to millions of forcibly displaced populations with limited human capital and who lack access to basic services and employment opportunities. DIME's FCV portfolio spans across countries covering interventions that aim to enhance the economic and social reintegration of these populations including, refugees, IDPs, and returnees in labor markets of host communities. This includes the provision of capital grants to refugees and IDPs or providing them with access to online employment opportunities through interventions being

implemented/designed in the DRC, Niger, Azerbaijan, among others. This program is in close coordination with the Trust Fund managed by the WB FCV group, as well as potential collaboration with GIZ.

4. Strengthen and expand partnerships with DFID country programs in strategic areas of overlap.

Last but not least, we will continue to build on existing collaborations with DFID country offices to carry out joint IEs on selected programs that fit with E4P thematic priorities. Currently, DIME is working with DFID on employment of marginalized youth in Northern Nigeria. In the DRC, following a long collaboration on the Tuungane project, DIME is pursuing its engagement with DFID by engaging in discussions on a new evaluation of the UNICEF implemented WASH program. We plan to leverage these existing partnerships and expand to other country programs in DFID priority countries.

6.2 Agriculture

Urgent Need for Evidence in Agriculture

The Sustainable Development Goals (SDGs) note that agriculture is the single largest employer in the world. Globally, 40 percent of the population earns its income from agriculture, and it is even higher in Sub-Saharan Africa (61%) and South Asia (55%).⁴ The SDGs urge the international community to make the investments needed to double agricultural incomes of small-scale food producers.⁵ However, current yield trends suggest a need for path-breaking innovations to come to the rain-fed areas of the world to meet this target (Figure 6.3). Astonishingly little evidence exists to rigorously inform the investments needed to meet this

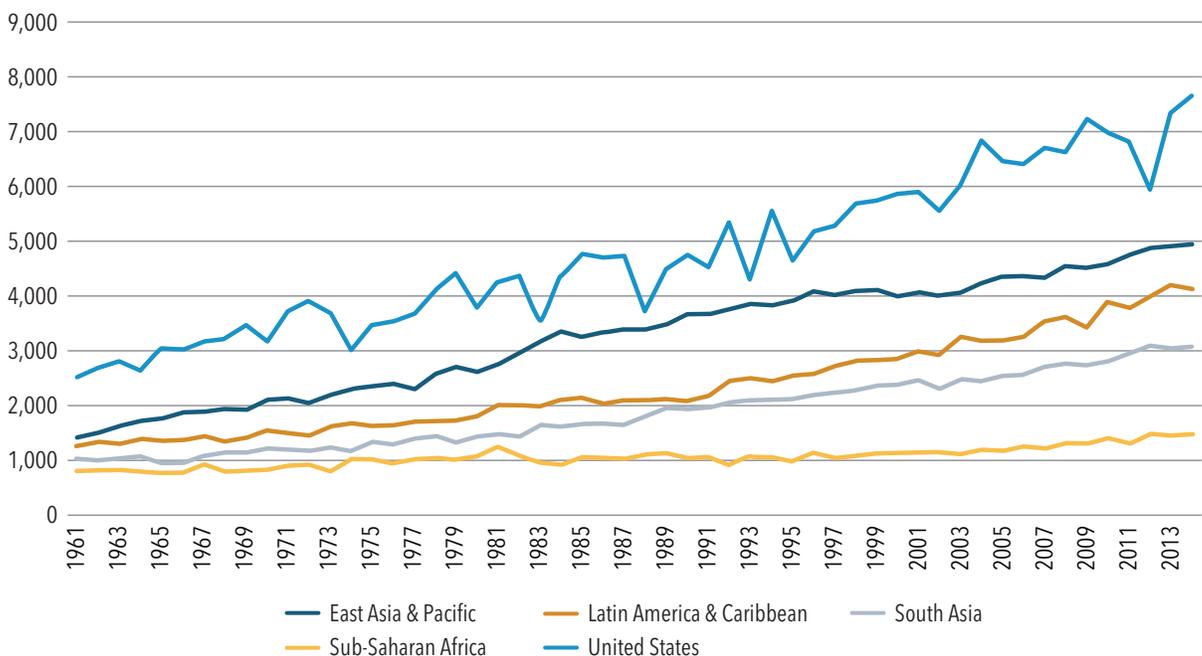
urgent goal. For example, a 2015 systematic review on the effects of training, innovation, and technology on smallholder productivity in Africa identified only 19 studies that met the scientific standard to be included in the review, making it impossible to assess which interventions yield the highest returns (Stewart et al. 2015). The gap between the urgent need for action and the evidence available to inform such action is therefore greater in agriculture than in many other sectors in development.

Agricultural development is crucial, not only for poverty reduction, but for many other SDGs as well. Ending hunger and improving nutrition for the hungry 13 percent of the developing world requires restructuring the agricultural value-chain: from farmers who grow food all the way, to retailers who sell it to consumers. As a sector contributing both carbon emissions and capture, and uniquely susceptible to climate and extreme weather, agricultural innovations will need to address climate change through both mitigation

⁴ World Bank World Development Indicators. <http://data-bank.worldbank.org/data/reports.aspx?source=world-development-indicators>

⁵ <http://www.un.org/sustainabledevelopment/hunger/>.

■ ■ **Figure 6.3** Agricultural productivity (cereal yield, WM WDI)



and adaptation. DIME's agriculture portfolio builds evidence on the innovations that best address these complex and overlapping challenges.

Policy-Driven Evaluation Design

Many of DIME's i2i-supported impact evaluations in agriculture were launched following a workshop on Agriculture Innovations held in June 2014 in Kigali, Rwanda. Ahead of this event, the Africa Region of the World Bank organized a high-level meeting of decision makers from ministries of finance and agriculture, researchers, and other policy makers to set priorities for research. This task force leveraged the evidence produced by the first-generation program of impact evaluations in agriculture, launched by DIME in 2009/10, and produced a series of learning priorities, including the urgent need for evidence on the impact and sustainability of rural infrastructure and the role of markets towards a rural transformation. The June event then took the resulting recommendations to a gathered set of policy makers, project staff, and researchers to embed research questions and designs into the project. This model of involving policy makers from the earliest stages of designing evaluations and building the evaluation directly into projects ensures buy-in from projects and immediate policy relevance of research findings.

In November 2016, DIME convened a conference on Evidence for Agriculture to share findings from ongoing and completed evaluation in the agriculture portfolio and to identify emerging priorities for evaluations in the sector. This event engaged participants from fifteen institutions, including university researchers, policy makers from governments, multinationals, and donor agencies.

In January 2017, a Policy Research Talk at the World Bank on the topic of Retargeting Investments in

“Now we have more information. Hence, before making a decision we look back into the impact evaluation results . . . with the live maps and market surveys we can track in which markets there are tomato shortages and intervene on time if needed.”

IE CLIENT, RWANDA RURAL FEEDER ROADS

Agriculture distilled the learning from the first eight years of DIME's work demonstrating strategies for enhancing agricultural productivity. This talk provided the basis for engaging policy makers in the World Bank and beyond on frontier issues in agricultural research and innovation.⁶

In November 2017, DIME convened a meeting of the Global Agriculture and Food Security Program (GAFSP), a US\$1.2 billion multi-donor trust fund that has received over US\$100 million in contributions from the UK. The workshop brought together GAFSP projects that have worked with DIME for the past five years to do impact evaluations along with six new projects receiving grants from the GAFSP to implement projects for the next five years. This event set a course for strategic learning within the GAFSP, and led to the new projects, starting out with rigorous impact evaluations to test targeting methods, evaluate the constraints

⁶ <http://www.worldbank.org/en/news/video/2017/01/24/retargeting-investments-in-agriculture>

to commercialization, and to measure the returns to infrastructure investments, like roads and irrigation. The event was followed by an invitation by WFP for a full-day DIME session on impact evaluation results relevant to WFP programming.

In addition to large global events, the team organizes local in-country events on a regular basis. These one-day events bring together relevant stakeholders with the objective of exchanging experiences, presenting results, and providing training.

Some examples of such events are:

Nepal. On January 18 2017, members of DIME led an IE Methods Workshop in Kathmandu. Participants for the workshop included senior officials from Ministry of Agriculture, Ministry of Livestock Development, and three of the agriculture projects supported by the Bank in Nepal: the Rani Jamara Kulariya Irrigation Project (RJKIP), Project for Agriculture Commercialization and Trade (PACT), and the Agriculture Food and Security Project (AFSP). The workshop aimed to provide policy makers within the ministries an introduction into IE methods and to serve as a platform for participants to learn from one another's experiences. Given DIME's growing portfolio of evaluations in Nepal, the event showcased for parts of the Ministry of Agriculture, not yet undertaking impact evaluations, the increasing capacity already existing within the Ministry to conduct and learn from this kind of work.

Mozambique. In November 2017, the team brought together several key actors in the irrigation sector of Mozambique. Representatives from the National Irrigation Institute, the Ministry of Agriculture and Foods Security, AfDB, WB, and local universities participated in the one-day event, which was meant to spark the exchange of lessons learned

across the different IEs. The team shared experiences from the innovative data collection systems deployed in each project, as well as early impact evaluation results. The event was also the kick-off of the design of a new impact evaluation seeking to explore how small emerging farmers can play a multiplying role in providing extension services, inputs, and access to markets to smallholder farmers. These discussions led to plans with other donors in the sector to incorporate impact evaluation methods into their investments.

Rwanda. Over the course of two high-level workshops held on September 28 and October 2, 2017, the team launched the DIME-EU partnership in Rwanda. The workshops brought together key decision makers and stakeholders in Rwanda's agriculture-policy and research space to present a vision for evidence-based action in the sector. The partnership aims to build on DIME's existing portfolio of IEs in Rwanda and to increase the role of evidence in policymaking-at-large. As the government looks towards its agenda of Rural Transformation and puts in place a policy framework for the next phase in its development path, the countrywide program of evaluation will bring key evidence on a range of investments in the rural space. More details on the DIME-EU partnership can be found in the partnership spotlight section at the end of this chapter.

Active Impact Evaluations

The Agricultural Adaptations (AADAPT) portfolio includes 17 impact evaluations in 16 countries across Africa, South Asia, Latin America, and the Caribbean. The evaluations are distributed across five knowledge gaps identified as constraints to the design of effective agriculture policy. These topical areas of focus are highlighted in Table 6.1. Over half of these evaluations (nine) have

■ **Table 6.1** Knowledge gaps and priority research themes.

Knowledge Gap	Example IE results	Active IEs in this Area
<p>Commercialization What are the public investments needed to ensure that farmers have access to markets and receive fair prices for their products?</p>	<p>Brazil: Providing rural organizations with matching grants to invest in machinery and marketing increased farmers' likelihood of providing commercialization, increasing overall sales value by 86%.</p>	Haiti, Brazil, Liberia
<p>Financial Constraints How do financial barriers and institutional constraints prevent farmers from making profitable investments? What are the simple interventions that can overcome these constraints?</p>	<p>Haiti: A subsidy for fertilizer caused farmers' rice yields to <i>decrease</i> by 30%. This counterintuitive result occurred because most farmers were already purchasing fertilizer. Because subsidized fertilizer was delivered late, farmers eligible for the subsidy applied fertilizer at the wrong time and experienced lower yields than farmers who paid full price but were able to use fertilizer at the right time. In the future, the government will shift away from fertilizer subsidies and toward promotion of agroforestry.</p>	Rwanda, Benin, Haiti, Uganda
<p>Rural Infrastructure Are large infrastructure investments always profitable? Beyond construction, how can we ensure sustainability of investments by building effective users groups to manage infrastructure?</p>	<p>Rwanda roads: Households in remote villages are typically the poorest. These households see the largest benefits from road rehabilitation. When roads are completed, remote households' income increases more than 20%, enough to catch them up to the initially more-connected villages.</p>	Rwanda, Mozambique, Kenya, Nepal
<p>Information Are farmers aware of the productivity gains to be realized from adopting new technologies and methods? If not, what are the most efficient ways to help them learn about these opportunities?</p>	<p>Bangladesh: Allowing farmers to experiment with new technologies on their own farms increases adoption more than traditional technology demonstrations. The adoption gains are driven both by "learning-by-doing" and learning from others. The important role for "learning by doing" implies that this mechanism should be incorporated into the design of extension programs.</p>	Bangladesh, Mozambique, Malawi, Rwanda, Nepal
<p>Natural Resource Management How can we encourage rural communities to manage and protect natural resources such as forests, clean water, and soil, while supporting livelihoods that rely on these resources?</p>	<p>Ghana: Small payments to farmers can incentivize farmers to adopt tree crops that are costly in the short run but profitable in the long run and have environmental benefits. Payments of less than \$100 can increase participation in tree-crop cultivation from 38% to 98%. Behavioral nudges are being tested to ensure that participating farmers keep their trees alive.</p>	Ghana, Burkina Faso

conducted a baseline survey, and six of those are predicted to be completed in the next year.

Improving the State of Knowledge and Implementation

DIME produces rigorous evidence on understudied issues relevant to agricultural policy and which advances knowledge that can be

used to design policies to improve productivity in the sector that provides the largest source of income and jobs for the world's rural poor. DIME has already produced rigorous research on how to adjust extension programs to optimize knowledge diffusion, the relationship between land rights and technology adoption, and the role of gender in learning about technology.

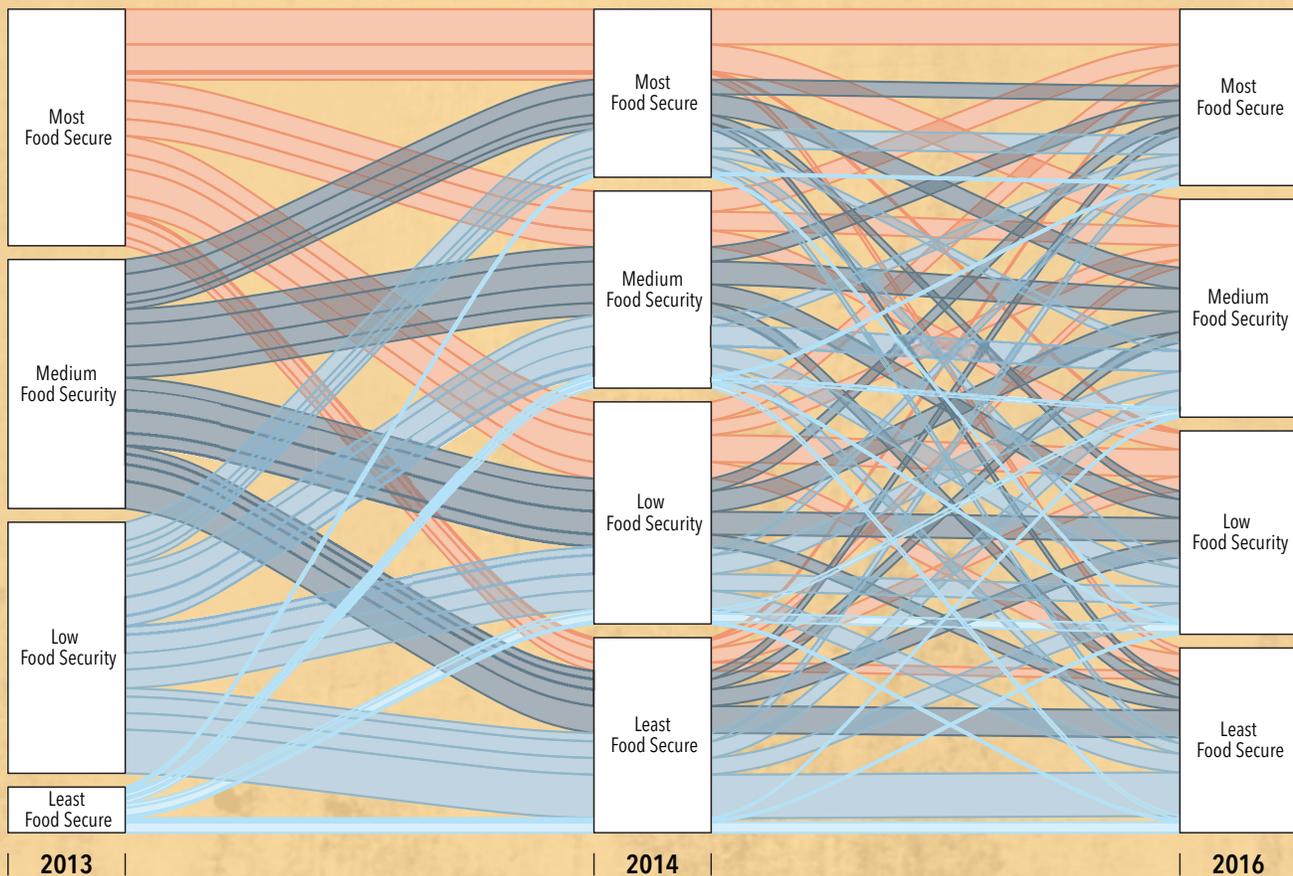
ANYONE CAN BE FOOD INSECURE

DIME's research shows major changes in household food security status over time



The graph shows the change in household food security status over 3 survey rounds in a client country in Africa. While a small fraction of the households remain in the same relative category that they started off at in 2013, the great majority of households see drastic changes across the 2013 to 2016 period. The underlying distribution and the apparent volatility in the outcome is guiding further analysis into the drivers of food security for this population and the mechanisms that determine these changes.

Major Changes in Household Food Security Status Over Time



DIME's research influences policy directly through intensive interaction with partners from governments and multinationals and changes the way that agriculture programs operate throughout every stage of the impact evaluation. For example, an impact evaluation in Rwanda empowered monitors from within farmer groups to report on

maintenance and water-sharing issues and the weekly data they report has helped the project identify additional training and support needed for optimal use of the infrastructure. In Malawi, impact evaluation results showed that demonstrations from average farmers were a more effective way to increase adoption of new technologies than

traditional extension services, which prompted further tests of decentralized demonstration and “learning-by-doing” in Bangladesh and Rwanda. For an irrigation project in Mozambique, DIME developed a simple transparent system to identify farmers cultivating below an area threshold set by the project. This allowed the project to roll out a structured selection process, improving the inclusion of the projects’ priority farmers, compared to a traditional community selection process. The system is easily replicable in other projects.

Developing data systems allows for uniquely comprehensive assessment of project impacts, and always lays the ground for rapid testing of innovations at scale. For example, the Rwanda roads project will make use of the established data system to test and learn from different monitoring systems, to ensure that the newly-constructed roads are well-maintained.

As the program is maturing, many projects are reaching final analysis stage. Examples of results across the topical areas can be found in Table 6.1. Many of these results are used to motivate further research, even in other countries working on similar issues. Further, research showcasing IE results from the portfolio often appears in working papers and top journals in the field of development economics.

Going Forward

Following multiple international consultations with DFID, GAFSP, WFP, FAO, EU, and others, DIME is moving from a phase of consolidating results from the first wave of i2i supported evaluations to bringing that evidence and practice of evaluation to new partners. In the coming few years, the focus of the agriculture portfolio will be committed to three approaches: sector strategies, country partnerships, and data systems.



Taking a sector approach means identifying key intervention types and building complementary evaluations across country contexts. One example of this approach is pursuing coordinated evaluations in Mozambique and Haiti that are both focused on strategies to leverage private sector investment to stimulate commercialization of farmers who are emerging from small to medium or large scale. Country partnerships are intense engagements with single countries across the agriculture sector. Engaging with multiple initiatives related to the agriculture sector ranging from extension to transport to commercialization allows the team to identify complementarities in the constraints to transformation of the rural and agricultural economies. The Rwanda partnership is currently the most mature example of these types of engagements, and the model is relatively advanced in Mozambique, where four evaluations are completed or near completion and at least two others are planned or ongoing. Finally, a core principle in the upcoming agricultural work is integrated data systems in which increasing large scale data collection links the household surveys traditionally employed in impact

evaluation to remote sensing and administrative data on roads, markets, land transactions, use of infrastructure, and other key types of data allowing for holistic understanding of the environment in which rural transformation happens.

Consolidating evidence into policy-relevant messaging. As it is maturing, the program will work to consolidate its policy lessons and ensure dissemination of those findings within the local government, across the portfolio, and the wider development community, as well as publication in peer-reviewed journals.

In Mozambique, an evaluation of the Mozambique Sustainable Irrigation Development Project (PROIRRI), a smallholder irrigation project, demonstrated the feasibility and usefulness of a high-frequency, user-managed irrigation monitoring system. One year of monitoring irrigation systems revealed that water scarcity arises primarily from misallocation. Simple, low-cost reminders of how much water crops need reduced overwatering and eliminated farmers' reports of conflicts over water. These findings led to the government counterparts scaling up the measurement in two ways. In the first half of 2018, a text message system of reminders about crop watering requirements will test whether the insights from this measurement system can be cost-effectively scaled into a full extension system. Second, the World Bank's support for the strengthening of the National Institute of Irrigation will include a component to build capacity for the Institute to take over the systems to collect and manage this type of high frequency data.

Another example of demonstrated learning and adaptation from projects within the i2i agriculture portfolio comes from Haiti. For five years, the government of Haiti conducted a large pilot project of subsidized fertilizer distribution using

vouchers. A DIME impact evaluation found that this strategy was not an effective approach to improving farmers' income and productivity. Because baseline fertilizer use was higher than the government expected, the subsidy vouchers simply replaced fertilizer transactions that would have happened anyway, meaning that the subsidies did not add value. Before starting a new project, a representative of the Ministry of Agriculture from Haiti attended an event in the summer of 2017 at the World Bank to discuss the experience of these findings. The event included representatives from an evaluation in Mozambique where fertilizer subsidies had been more successful and one in Tanzania where results had been mixed. This event consolidated a consensus that input subsidies need to carefully address the issue of baseline adoption before being rolled out. The next iteration of the Haiti investments will move away from fertilizer subsidies for staple crops and toward the promotion of agroforestry practices.

Focusing on emerging priority areas. In consultation with the Agriculture Global Practice, the AADAPT team have identified areas where practitioners within the Bank feel that additional focus is needed. One example of a new approach is understanding complementary investments and goals in agricultural programs, particularly those related to nutrition, social protection, and climate change. Ongoing evaluations in Nepal and Rwanda seek to uncover how nutrition interventions and social protection interventions can enhance the impact of agricultural productivity-enhancement efforts. A second area is a push toward expanding the commercialization portfolio. Newly-launching projects in Senegal and the DRC will focus on value-chain interventions, warehousing, and marketing as channels to translate agricultural productivity into income.

DIME PARTNERSHIP SPOTLIGHT

MINAGRI-EU-DIME PARTNERSHIP FOR EVIDENCE TOWARDS RURAL TRANSFORMATION IN RWANDA

Scaling up the use of evidence. As part of its new strategy, DIME is piloting country programs of impact evaluations that take a sector-wide approach to evidence-based policy. The first country program was launched in Rwanda in September 2018 under the auspices of DIME's agriculture program, under the leadership and with financial support from the EU delegation in Kigali. The collaboration spans a broad range of issues in rural development, from rural roads, land management, to irrigation, commercialization and technology adoption.

Building deep knowledge of a sector. DIME's work with the Ministry of Agriculture and Animal Resources (MINAGRI) of the Government of Rwanda (GoR) began in 2011 in collaboration with the Global Agriculture & Food Security Program (GAFSP). Quickly, MINAGRI's enthusiasm for evidence led the program to expand to other investments, including investments financed by the EU, NL, and USAID. Presently, the program spans multiple experimental evaluations as well as natural experiments. The evidence generated contributes to building the science of delivery for a range of investments: terracing, irrigation, feeder roads, land management, technology adoption, rural finance, accountability in extension service delivery, as well as understanding mechanisms for operation and maintenance of rural roads and of irrigation projects. The process through which one policy research engagement matured into a sector-wide approach to impact evaluation is worth describing in further details.

A process of policy engagement. Working across this program of IEs, the research team has engaged closely with program implementers and policy makers in the agricultural sector. Through support from the European Union delegation in Rwanda, DIME's work in Rwanda is set to continue and deepen, with a focus on enriching the policy dialogue in a country that is moving from a target of improving the agricultural sector to transforming the rural economy. The partnership will continue to work on existing themes, and build on this through a deeper engagement with the policy apparatus, focusing on capacity-building within relevant line ministries and institutions. Signed in August 2017, the partnership signals a long-term commitment by the EU to support and generate a culture of decision-making within policy circles that relies heavily on data and evidence.

Building and leveraging data systems. Under the partnership, DIME aims to move from evaluating individual programs and projects, towards developing a data-driven evidence ecosystem. One example is an ongoing initiative to set up a large-scale data system that collects and collates information from across the country and in a multitude of sectors. This has involved bringing on board a number of stakeholders and decision makers in Rwanda from various government and independent institutions. A combination of existing administrative data from several line ministries and extensive primary data collection across the country's districts, this data system aims to track indicators including commodity availability and prices, migration patterns, household income and more. This wide-ranging data dashboard will enable policy-makers in Rwanda to track the movement of key outcomes over time. Further, the system aims to leverage data within and across sectors in order to increase policy-makers' capacity to draw on lessons from the data trends as they design, roll-out and improve policies.



Policy impact. For the past four years, DIME has worked with MINAGRI to design rigorous evaluations and innovative implementation modalities, roll out Randomized Controlled Trials within government programs, and collect agricultural data on an annual basis. The Land Husbandry, Water Harvesting and Hillside Irrigation (LWH), and Rural Feeder Roads (RFR) projects are examples of large infrastructure investments with goals to profoundly transform the agricultural sector. DIME has taken a lead in designing IEs aimed at answering questions related to the delivery, sustainable maintenance and use, and overall impact of these flagship programs.

In the case of the LWH Project, the evaluation aims to formally document the impact of the intervention in project sites, using as a comparison group similar pre-identified watersheds that will not receive LWH project activities. The main identifying assumption is that the only difference between pre-identified sites that receive LWH and those that do not is the project, with a pairwise matching strategy to tease out the program's impact. Data has been periodically collected across treatment and control sites beginning in 2013 and initial results are encouraging. After two years of project interventions, there is an increase in the proportion of households with access to public extension services (18 percentage point difference between treatment and control), higher adoption of radical terracing (59% of HHs in treatment vs. 11% of HHs in control sites), as well as improved investments in a variety of ag-inputs (a 280% increase in inputs in Season B) which subsequently lead to gains in crop production (an increase of RWF 19,000 in 2013 Season A) and sales (an annualized increase in sales of RWF 15,000 in 2014). With an endline survey in the pipeline, these and many other outcomes will continue to be tracked, providing rigorous evidence of the short and medium-term results of the project in the tracked sites.

In the RFR program, the impact evaluation will use an *event study* design, using high frequency *pre-post* comparisons of outcomes at the road segment level.⁷ Identification requires that the exact timing of the intervention is as good as random. In this context, idiosyncratic factors such as donor disbursement calendars, construction delays, permits and weather all suggest that is indeed the case for construction completion date for each road segment. Causal changes in outcomes of interest (i.e. prices, quantities traded) can therefore be estimated around the timing of the event (rehabilitation). Towards this end, DIME has begun the process of collecting data from a variety of sources to set up a robust information system that will feed into the IE. In addition to a standard household survey (conducted annually), the IE will use administrative data already collected by the GoR, and high-frequency market data collected through mobile technology. Road segment-level data (international roughness index and traffic count) will be collected by the Rwanda Transportation Development Authority (RTDA) annually. Land and migration records collected by local government were digitized for this IE, and will be compiled for analysis. By working with data already collected by the GoR, the research team is gradually and sustainably building the capacity of existing data collection systems. Market data will be collected to create a comprehensive list of local markets and the prices of goods. The high frequency (bi-monthly) data will allow us to clearly track trends over time, and identify changes around the specific rehabilitation event. The data system being set up to identify the impacts of this program will – as discussed earlier – aim to affect policy priorities for the government across sectors.

While the active collaboration between DIME and MINAGRI across these projects reflects a long-term relationship, the effort has been to strategically test delivery mechanisms on a number of interventions—directly affecting the results of the LWH and RFR projects, and contributing to MINAGRI's objectives at-large.

⁷ The roads are currently under construction, with the WB project scheduled to end in 2018

One example of this is a collaboration between DIME and LWH running a Randomized Control Trial (RCT) with 80 Self-Help Groups in Karongi district in 2012-13 aimed to introduce and rigorously test two new types of savings products: a targeted savings account and a commitment savings account. An important component of LWH was transitioning from subsistence to commercial agriculture. DIME worked with the rural finance specialist to create and rigorously test innovative financial products designed to help farmers manage their finances to provide for their families and have money available at the beginning of each season to purchase agricultural inputs. The new savings products had fairly typical take-up rates: 25-30% of households registered for one of the new types of accounts.⁸ The offer of the commitment savings account increases savings activity, and increases the maximum balance held in the accounts in the lead up to Season A. Male-headed households offered commitment savings accounts increase their investments in fertilizer for Season A. Female-headed households offered commitment savings accounts also invest more in fertilizer for Season A, but that does not translate to gains in agricultural productivity.

Based on these positive results of the pilot in Karongi, the LWH team decided to test similar products at scale and worked with DIME to implement an RCT with 5 agricultural cooperatives in Rwamagana district in 2013-14. First-order impacts of the introduction of the program were optimistic - 80% of farmers who attended a training registered for one of the new accounts. However, with the scale-up that relied heavily on existing institutions and with a far greater degree of hand-holding from DIME and the project team, the positive impacts of the pilot did not follow through. This is consistent with the fact that even during the pilot, beneficiaries did not actually use the accounts – mental accounting was the crucial piece in the impact chain, and the scale-up did not affect this margin. The promising results of the pilot provided the government team with enough evidence to scale up the activity. However, when it was expanded and extended to a different district, outside of the careful management of the project team, the impacts of the program did not hold – in fact they were statistically significant and indistinguishable from zero across a range of indicators. Based on the evidence of the lack of impact in the larger program in Rwamagana, the LWH team decided not to extend the program to other districts.

In addition to rural finance, testing the modalities to improve extension services are central to LWH's aims of capacity building and technology diffusion. In LWH project areas, farmers purchase agricultural services (inputs and extension) from One Acre Fund (OAF). DIME worked with OAF and the MINAGRI to design, introduce, and test innovative farmer feedback tools. The government was interested in monitoring the extension services provided by OAF, and farmers' satisfaction with the service. In addition, the research team was interested in whether feedback tools could actually increase demand and improve low participation rates. Together, the team set up a large field experiment, in which two types of feedback tools were randomly assigned to groups of OAF clients. The team also tested the cost effectiveness of different feedback modalities. The evaluation resulted in a number of important lessons. First, feedback tools help sustain demand for the service among current clients. Farmer groups offered the opportunity to provide feedback were half as likely to have members leave the service the following year as control groups. Second, and more surprisingly, this demand effect spills over to non-users in the vicinity of the treated groups, who are more likely to sign up in the following season. Farmers groups who have access to feedback tools are 28 percentage points more likely to attract new members, relative to control farmers groups that have an 8 percent chance to attract new members. The most cost-effective

⁸ Not all of those households actually went to the SACCO to formally open the new account. Further, analysis of the SACCO's transaction data shows that very few households (<1%) actually made deposits or withdrawals into the new accounts. This is possibly a consequence of administrative errors or poor record-keeping by the SACCO.

feedback mechanism piloted (a hotline) was adopted and scaled up by OAF throughout Rwanda the following season. In addition, the satisfaction data drawn from the feedback tools themselves helped convince MINAGRI to continue its partnership with OAF and scale it up to new LWH sites.

The extension experiment is an example of a happy marriage between research and operations: an impact evaluation motivated by a direct policy question, but going beyond simple program evaluation. DIME had an opportunity to improve the effectiveness of a policy that touches large numbers of farmers. In addition, useful data were collected in the process, highlighting gender difference in access to services as well as information on quality of service delivery.

The final element of DIME's ongoing work in the country is a number of experiments around irrigation systems that are in early stages of implementation. Given the extent of the expense involved with setting up the irrigation system, the government is thinking deeply about the cost-effectiveness and the sustainability of these investments. Towards this end, the research team is working with the government to produce rigorous evidence that will guide the scale up decisions, as well as strategies to improve the performance of irrigation schemes. Construction of the large-scale irrigation schemes is complete in three sites; all farmers in the irrigated area have access to irrigation, with fees charged to all households that irrigate starting in 2016. This evaluation has two components. First, to evaluate the impact of access to irrigation on farmers, outcomes of farmers whose plots are just inside the irrigation scheme will be compared with farmers whose plots are just outside the irrigation scheme – a spatial regression discontinuity design. Second, we test mechanisms designed to address the two primary concerns about the sustainability of irrigation investment, using Randomized Controlled Trials (RCTs). We test the impact of escalating irrigation fees on farmers' adoption of a high-value cropping system, through fee subsidies; the impact of experiential learning, using demonstration mini-kits; and the impact of empowering water user groups to develop, monitor, and control their own operation schedule and maintenance.

DIME's portfolio of IEs in Rwanda has evolved over time, with several short- and medium-term outputs that have already begun to shape the government's approach to a rural transformation agenda. Through support from the EU, DIME will continue to closely align with MINAGRI's sector-wide approach. It will support program implementation by testing alternative delivery mechanisms, and building knowledge feedback loops under key pillar of MINAGRI's poverty reduction and agricultural strategies. In practice, this will be done through three main channels: (1) an impact evaluation team that merges research, operations and project management, implements the work, with the day-to-day technical support of a field coordination team based in Kigali; (2) the impact evaluations will focus on key policy areas agreed on with MINAGRI, and will be specifically designed to yield actionable recommendations that back operational decisions with hard evidence; and (3) wider capacity building for evidence-based policymaking in MINAGRI and the Ministry of Finance (MINECOFIN), as well as the National Institute of Statistics (NISR) and the local research community.

The MINAGRI-EU-DIME partnership is a flexible instrument to accumulate learning to increase program impact, and to adapt to emerging needs on the ground. As results come out of ongoing experiments and descriptive analyses, they will inform new testing for subsequent agricultural seasons. The partnership will build a community of practice for evidence-based policy making including partnerships for impact evaluation and long-term engagements linking policy makers, researchers, development agencies, and other stakeholders. Leveraging existing and ongoing work, the partnership will allow DIME to take a multi-sectoral approach to policy making. Through the partnership, the research team aims to be a proactive partner in framing upcoming national strategies and generating real-time evidence of what works, focusing on approaches to improve the effectiveness and sustainability of the government's wide range of development strategies.

6.3 Governance

Government organizations provide essential public services in key areas such as health, education and infrastructure, and the size of the public sector is especially large in developing countries. However, key determinants of effective governments are still largely unknown. For instance, *which factor is most important to ensure effective delivery of public goods, or which mechanisms can ensure a more transparent and accountable public procurement process* are still unanswered policy research questions. Governance reforms are often long term, complex, and difficult to measure. Rigorous evidence on what works in the sector is, therefore, in short supply with the governance field representing less than 3 percent of registered impact evaluations.⁹

The ieGovern program

DIME and the Governance Global Practice launched the ieGovern program in 2013 to produce rigorous evidence to improve governance project results and to push the frontier of available evidence on what works in governance reform. To date, the program has a portfolio of 38 IEs in all regions of the world that study four main themes: (i) civil service reform, (ii) public financial management (tax and procurement), (iii) justice, and (iv) decentralization/subnational public-sector management.

The program has reached maturity as most IEs in the portfolio are currently being implemented. Over the last year, important advances were achieved in the justice theme as the portfolio of IEs in justice has evolved into a full-scale research program, the Data and Evidence for Justice Reform (DE JURE). On the civil service

theme, the Bureaucracy Lab research program has convened a remarkable conference, *Innovating Bureaucracy*¹⁰ (see more details below).

Theme 1: Civil Service Reform - The Bureaucracy Lab

IE research has mainly focused on studying performance incentives for frontline staff—such as teachers, nurses, and doctors—that address, for instance, problems of absenteeism or underperformance. To go beyond this, the IE work in the civil service reform pillar focuses on research questions related to those civil servants who work in government’s core ministries, such as the ministries of finance and education, and who bear the responsibility for designing a country’s policies, collecting its taxes, and so on. Key policy questions being studied include how to motivate public sector workers to perform better with different (monetary and mission-based) incentives (Liberia and Pakistan); how to improve the governance of maintenance public infrastructure (Tanzania); how streamlined information flows within the public sector can improve project performance (Pakistan); and how a public-private partnership can facilitate access to public services for marginalized groups (India).

The IE work on civil service reform fits into a broader research program, called The Bureaucracy Lab, which is an initiative co-led by DIME and the Governance Global Practice. The Lab is creating improved administrative data on the characteristics of public officials and their organizations to inform the operational design of public sector organizations. In addition, the Lab is undertaking experimental work within large-scale surveys of civil servants to generate an evidence

⁹ 3ie Impact Evaluation Repository (<http://www.3ieimpact.org/evidence/impact-evaluations/>), accessed 1/14/2016.

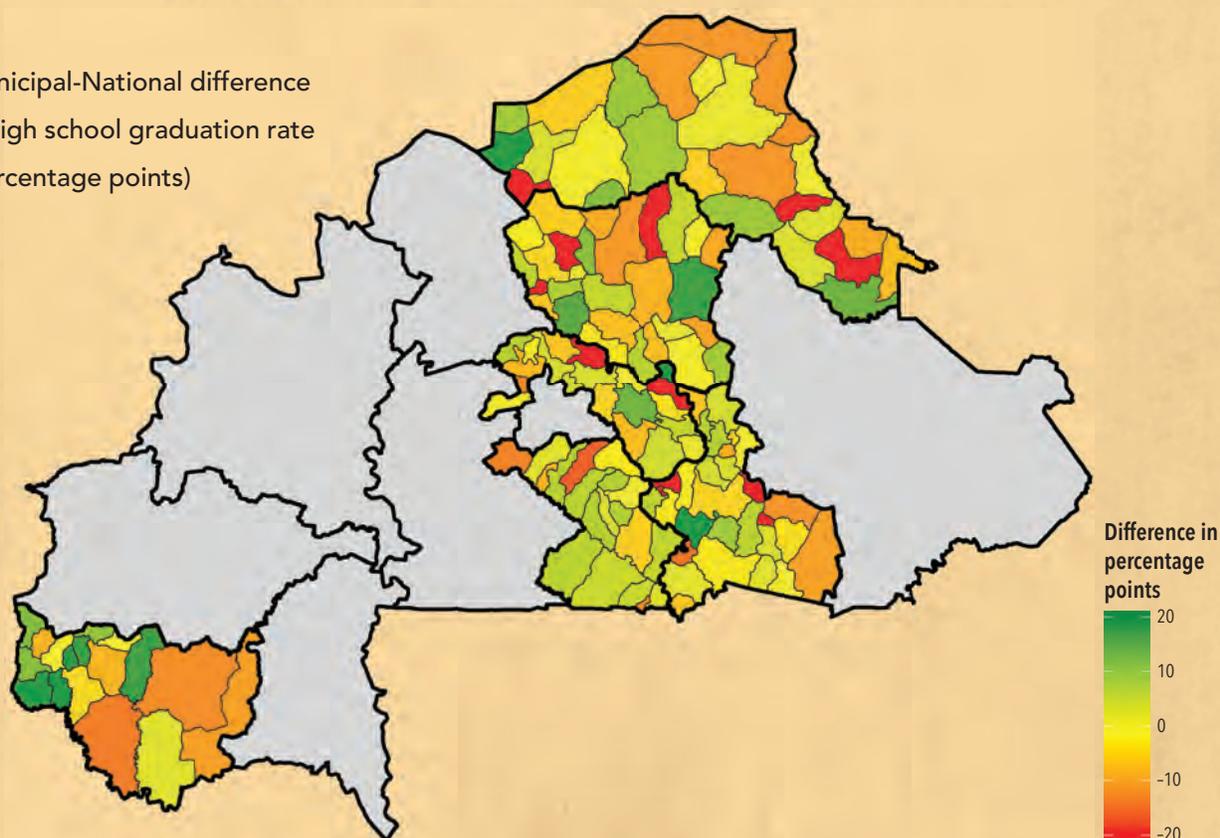
¹⁰ See <http://www.worldbank.org/en/events/2017/10/16/innovating-bureaucracy>.

IN BURKINA FASO, DIME'S ANNUAL MUNICIPAL PERFORMANCE SURVEY HELPS GOVERNMENTS TARGET THEIR POLICY EFFORTS



In Burkina Faso, DIME's work led to the creation of the SUPERMUN annual municipal performance survey which tracks key indicators of local government performance and capacity. This image shows the difference in municipal high school graduation rates relative to the national average, highlighting considerable variation across municipalities and within regions. Even in data-poor environments, such key performance metrics can be used for decision-making at national and sub-national levels and for citizens to hold their governments accountable. SUPERMUN was initially implemented in six of thirteen regions and, as of this year, will be scaled nationwide.

Municipal-National difference
in high school graduation rate
(percentage points)



base on how to survey civil servants effectively. The Lab is also working with academic anthropologists and sociologists to create a detailed picture of civil services across the world. Each of these elements uses the ieGovern program as a platform for research, whilst providing inputs that feed back into the design of the evaluations.

The *Innovating Bureaucracy* conference has allowed the Lab to present its approach to an audience of leading scholars and practitioners in the field. The approach combines embedded micro-level diagnostics based on primary data on public officials and their activities; experimentation and adaptation in reform implementation;

and an appreciation of the political economy in which the public sector is situated. The conference has also allowed for a review of existing evidence on which these reforms will be developed.

Theme 2: Tax and Procurement

Public Financial Management (PFM) reforms have been a core of support in client countries by the World Bank and other donors for a long time, yet only few IEs exist on the effectiveness of different PFM systems. The IE research work under ieGovern has tried to fill this gap with several IEs in the PFM subsectors of tax and procurement. Research questions being addressed include the impact of the adoption of e-procurement systems on competition and market entry of new firms, prices and value for money of government purchases (Bangladesh and Brazil); how centrally coordinated framework agreements affect the procurement process and quality of services procured (Colombia); and how behavioral nudges and facilitation measures can affect willingness-to-pay taxes and tax compliance (Tanzania).

The ieGovern portfolio of tax research has helped spark the creation of a broader research program embedded in the Governance Global Practice, which is the *Innovations in Tax Compliance* program. The objective of the research program is to influence the design of World Bank tax operations through the development of a multifaceted approach to improving tax compliance that explores strategies that are both i) technically appropriate and ii) lever the Bank's broader governance operations to engage citizens, and progressively build trust, reciprocity, and support for tax compliance.

To do this, the project will develop a framework that holistically looks at enforcement, facilitation, and trust as key mechanisms to improving tax compliance. The work recognizes that technocratic

“Elements from the IEs, including the annual municipal performance survey and the municipal scorecard intervention, are being scaled up under the Bank-assisted Burkina eGovernment project and PACT Additional Financing. This is part of the project restructuring.”

IE OF BURKINA FASO LOCAL GOVERNMENT SUPPORT PROJECT

reform focused on enforcement and facilitation remains essential, but more substantial and long-term improvements are ultimately likely to depend on building relationship of mutual trust between government and taxpayers. The project will serve as a convening force for research partnerships both inside the Bank and with outside academic/research institutions including International Centre for Tax and Development (ICTD), Institute for Fiscal Studies (IFS), and Chr. Michelsen Institute (CMI).

Theme 3: Data and Evidence for Justice Reform (DE JURE)

An efficient, fair, and accessible justice system safeguards peace and security, encourages investment and growth, and is fundamental to notions of citizenship and trust in government. Yet, there is little empirical research on justice system reform, in large part because data is not easily available. In recent years, however, governments around the world have embraced electronic case-management systems and have used innovative technologies to expand access

to justice. Leveraging the Bank's relationship with governments, the ieGovern work on justice is uniquely positioned to take the lead in justice research. The justice portfolio has evolved into a self-contained research program, Data and Evidence for Justice Reform (DE JURE), that aims to establish a global data infrastructure for the justice sector, and, through rigorous analysis and experimentation, to expand the evidence base on the economics of justice reform.

The DE JURE program has three pillars:

Data. Work with client governments to strengthen case management and administrative data systems, and pilot a public data depository that brings together key elements from these systems in the form of a series of *Doing Justice* indicators;

Measurement. Use administrative and survey data, in conjunction with economic theory and literature, to develop an empirically-validated measurement framework that lays the foundation for research on the economics of justice reform.

Learning. Work with client governments to tackle priority policy research questions to understand the impacts of changes in laws and regulations, information and monitoring systems, and incentives and enforcement mechanisms by embedding experimental research into the rollout and scale-up of justice sector interventions.

The DE JURE program was launched in 2017 at an impact evaluation workshop that brought together government officials, Bank staff, and academics from Azerbaijan, Brazil, Colombia, Croatia, India, Kenya, Mexico, Peru, the Philippines, Senegal, and the Solomon Islands, to identify immediate opportunities for expanding the evidence base in justice reform. Current research engagements include analysis of high-frequency administrative data in Croatia and Senegal,

experimental and quasi-experimental work on access to justice in Colombia and the Solomon Islands, and randomized controlled trials of performance management and behavioral interventions in Kenya and the Philippines. Going forward, the program aims to pilot a global depository of administrative and survey data on justice, and to identify additional priority countries in which to experimentally and iteratively test the impacts of new justice reforms.

Theme 4: Subnational Public Sector Management/Decentralization

Transferring power and responsibilities to local entities has been a very popular reform for many countries, including in OECD, middle-income, and poor countries. However, the evidence base of how decentralization reforms fare in practice has not kept up with the number of reforms. Unanswered research questions include how to measure and incentivize the performance of local governments; how to deal with potential elite capture at the local level; and how to make sure local governments have sufficient capacity to handle increasing responsibilities and collect their own revenues.

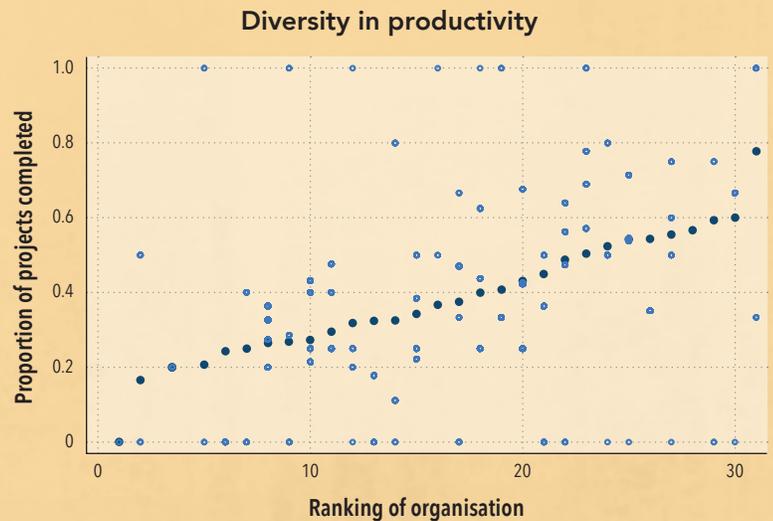
The research program is exploring several dimensions of the decentralization puzzle. In Cambodia and Dominican Republic, IEs are testing how to harness social accountability interventions to improve service delivery of local governments. The impact of demand side actors such as community officers and community-based organizations in making local governments more accountable is being studied in Burkina Faso and Solomon Islands. Transfers of resources from central government to local governments based on their institutional performance are being tested in Tanzania to assess



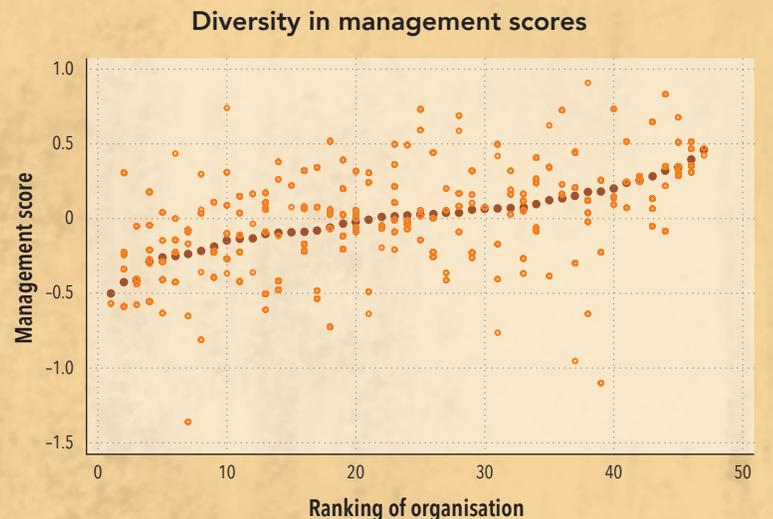
WALK ACROSS THE CORRIDOR TO INCREASE PERFORMANCE

DIME's research in Ghana shows huge variation within organizations

DIME supported research in Ghana collected data on management and productivity from across the civil service. Surveying nearly 3,000 civil servants in the central government and assessing completion of 3,628 projects across government has provided us with a detailed view of how effectively the government is functioning, and indicates why. This graph shows the diversity in projects completed by organization (solid dots) and by divisions within those organizations (hollow dots). There is huge diversity across organizations in the proportion of projects that get completed, but within organizations we see even more variation. Looking solely at organizational averages would hide the immense variability in performance across divisions – some divisions complete 100%, some complete 0%, even within the best performing organizations. The implication here is that walking across the corridor in the Ghanaian public sector could mean transforming productivity.



Why are there such large gaps from one office to the other within the same building? DIME research focused on the management practices within and across organizations. The graph shows huge variation in the quality of management associated with these organizations (solid dots) and divisions (hollow dots), mirroring the diversity in productivity. Do management practices explain any of the variation in productivity? The impact of a one standard deviation improvement in management practices leads to an increase in project completion of 39 percentage points. This is against the backdrop of 21% of projects never starting (see Rasul, Rogger and Williams, 2017). This finding has important policy implications - since Ghana's government expenditure is roughly 20% of the economy, a standard deviation improvement in management practices would lead to an 8% increase in GDP in direct productivity gains. Of course, there would also be indirect productivity gains through the improved interactions between the public and private sectors, so this is a lower bound.



whether this program-for-results type of incentive scheme is effective in improving local service delivery.

A Systems of Accountability Approach

All Governance impact evaluations in DIME programs are gradually adopting a systems of accountability approach. Critical service delivery questions will be addressed by diagnosing the governance environment in which they are embedded using frontier measures of government activity. DIME's programmatic approach to impact evaluation allows to create experimentally-validated measurement tools. Additionally, the World Bank's long-term engagement with governments has allowed DIME to initiate global databases of micro-level data on government that anchor learning within and across contexts. These diagnostic tools provide a platform to understand the critical relationships that underlie the corresponding systems of accountability. Experimentation will target systemic change in

these relationships towards strengthening service delivery.

For example, a current impact evaluation addresses the question of whether e-procurement is an effective device to improve transparency and efficiency of the public procurement process. The next generation of evaluations would look at the elements of the systems of accountability within the public sector that should be studied to ensure that e-procurement can lead to a more transparent and efficient public procurement process. By intervening in both the e-procurement system and in corresponding civil service incentives, our work would enable a broader understanding of the impact of the reform.

6.4 Climate Change

What is the right balance between meeting the World Bank's objective of eradicating poverty, while limiting environmental consequences? The Sustainable Energy for All initiative aims to achieve



universal access to sustainable energy; however, this would mean providing electricity access to over one billion new people. Globally, approximately two-thirds of greenhouse gas emissions already come from energy extraction and use (IPCC 2014). Agriculture is the largest sector in many developing country economies, but deforestation contributes between 10 and 17 percent of annual carbon emissions (Samii et al 2014a).

While we are on target to eradicate extreme poverty by 2030, we are also on target to increase the Earth's temperature to irreversible levels that are anticipated to have far-reaching, long-term consequences on economic growth, vulnerability, and the environment. While the problem of climate change is fundamentally a global collective action challenge, there are important program-level activities and insights that can help us mitigate its effects and strengthen resilience.

The initiation of an Energy and Environment (E&E) program was motivated by the dearth of rigorous impact-evaluation evidence in these sectors and the influential role they play in poverty alleviation and climate change. In 2011, the World Bank had 12 ongoing or completed impact evaluations in both energy and environment (compared to over 100 in education), despite the fact that these programs consist of almost 20 percent of the World Bank's lending portfolio.

Energy and Environment Impact Evaluation Program

The E&E program was launched in Lisbon in October 2014. It brought together 19 project teams (financed through DFID, the GEF, CIF, IDA and IBRD) and 28 researchers from 11 academic institutions to refine research opportunities based on project interest and operational

“CONAFOR [The National Forestry Commission for its acronym in Spanish] is now working on the design of the Emission Reduction Initiative in which we want to test an integrated territorial management model (several programs contribute to solving a problem within a polygon). We designed this program with an impact evaluation in mind (treatment, control, baseline, annual data collection, end line 6 years after the baseline). A pilot is being implemented in 5 states.”

IE CLIENT, MEXICO PAYMENT FOR ENVIRONMENTAL SERVICES SCHEME: A RETROSPECTIVE EVALUATION

feasibility. This was complemented by a parallel set of workshops focused on measurement opportunities in the sector.

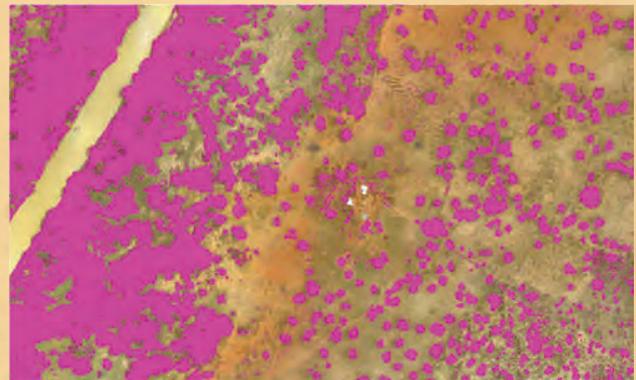
The first workshop, held jointly with CEGA in August 2014 in Berkeley, brought together engineers, economists, and World Bank counterparts to explore leveraging new technologies to improve measurement in energy and environment projects and research. A follow-up Berkeley



ARTIFICIAL INTELLIGENCE AT WORK

DIME's ongoing work shows that high precision mapping of trees in dry forests is possible

Understanding the real impacts of forest conservation policies requires accurate measurement of forest cover stocks and trends. On the left is a drone image taken inside Bontioli forest in Burkina Faso (centroid located at -3.077711, 10.805546). The 0.1m resolution image shows sparsely distributed trees, built-up areas, water bodies, and crop fields. In the image on the right, a machine learning technique called Random Forest was overlaid on the drone image to predict tree cover with an accuracy rate close to 100%. This high precision mapping technique enables policy makers to make data informed decisions.



measurement workshop, focused on innovative measures for climate resilience, was held in June 2015. To strengthen the economic theory underpinning each IE, a research workshop was then held in collaboration with the University of Chicago economics department. A set of project teams were given a chance to present their current design and receive critical feedback from leading academics within the E&E research team to ensure the work is able to maximize its contribution to the global knowledge agenda.

The original set of 12 impact evaluations was selected from a funding window in January 2015

that laid out the research program focusing on two pillars: (i) environmentally sustainable electricity supply, access, and efficiency; and (ii) natural resource and sustainable land-management issues, with a focus on incentive schemes, governance, and vulnerability. Transport, industrial pollution, and urban development relate directly to energy and environment and present major development challenges with important economic and environmental implications. However, these topics are addressed by other programs.

The research agenda has benefited from direct engagements with the World Bank's Climate

Change Cross-Cutting Solutions Area, Energy and Extractives Global Practice, Environment and Natural Resources Global Practice, Water Global Practice within the World Bank, the Global Environment Facility (GEF), the Climate Investment Funds (CIF), the Energy Sector Management Assistance Program (ESMAP), and DFID (including the evaluation department and climate-change teams). Work under the main themes is summarized below.

Theme 1: Energy Access, Reliability, and Efficiency

There is a growing body of evidence on the impacts of energy access on health, education, and productivity. But, important questions remain about the cost-efficiency of investments to balance the coverage (extensive margin) and per-connection availability (intensive margin) of providing electricity. The average annual energy consumption of electrified households in Kenya is 20 times less than the average American household. Therefore, understanding the demand and impact of different tiers of access from solar lanterns (Tier 1) through to full grid access (Tier 5) becomes an important concern to help governments efficiently allocate resources. There is an important trade off: while lower tiers of access may provide less opportunities for economic growth (for example, being unable to power large appliances and machines), the benefit of lower investment costs and easier expansion may outweigh this concern.

We explore this question by looking at the demand for and impacts of different levels of energy access. In rural Kenya, for instance, preliminary evidence suggests that one and a half years after the provision of grid connections to households, energy consumption increased to

levels that could be fully served via stand-alone pico-PV systems. While the economic and educational effects of grid electricity were modest, there is suggestive evidence of gains in employment and own businesses for women, especially for wealthier households.

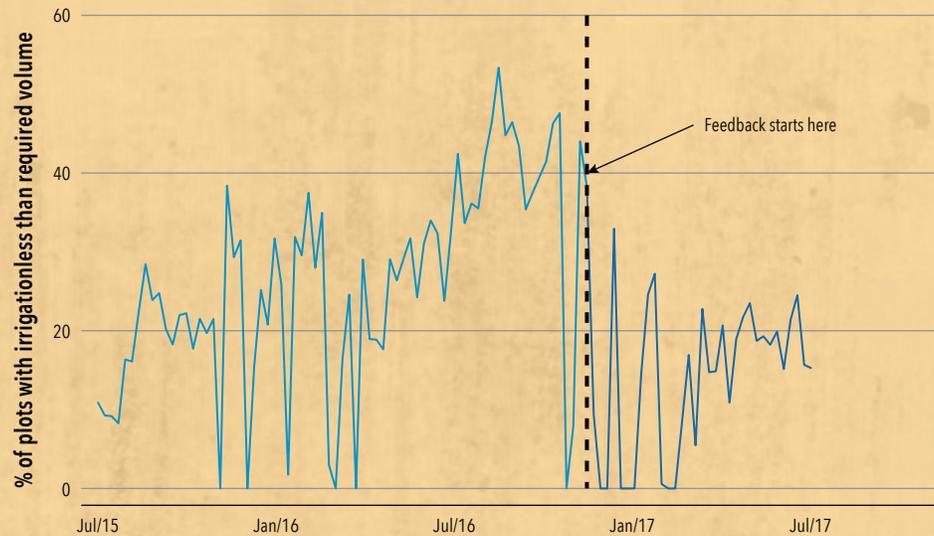
The results further indicate that the connection costs are multiple times higher than the willingness of households to pay, even after taking into account possible economies of scale (i.e., higher community connection rates reduce per unit costs). For example, at a price of US\$150, slightly more than 20 percent of the population in a community is willing to connect to the grid. The unsubsidized cost of connection per household is approximately US\$1,150, implying that a very large subsidy is required to increase take-up among rural households, which would require large social gains to be economically efficient. Yet, results show minimal effects on neighboring households within the same community.

An increasing number of national electrification strategies have relied on off-grid technologies and market-driven approaches. Many developing countries such as India are well on track to reach universal energy access by 2030. However, market failures limit faster growth and access to higher tiers of energy access. In Senegal, for instance, DIME found that imperfect information (mostly about product quality) faced by consumers when choosing among different lighting options causes an inflated demand for low-quality products at the expense of high quality pico-PVs (i.e., a “market for lemons”). Direct consumer outreach campaigns based on leaflets and posters aimed to improve knowledge about the good quality products increased their demand by 6.5 percentage points (this corresponds to a 50 percent increase considering the low status quo

SIMPLE FEEDBACK TOOL REDUCES IRRIGATION INEFFICIENCIES AND CONFLICT OVER WATER



DIME created a high frequency water use monitoring system to assess how effectively farmers are managing their irrigation water. We find that most farmers greatly overwater crops early on in the growth cycle. As a result, others are not allocated sufficient water to meet recommendations for their crops. Simple feedback tools greatly reduced these inefficiencies and reduced conflict among farmers over water.



level). Building on this work, the program will analyze additional interventions aimed to overcome this problem of imperfect information including third-party quality disclosures and guarantees.

Theme 2: Incentivizing Sustainable Land Use and Natural-Resource Management

The poor management of natural resources can be the result of multiple factors such as externalities, unclear property, or high discounting of the future. However, the gap between need for action and empirical evidence to guide those actions, especially in the environmental sector, means that efforts to increase the evidence base on the effectiveness of various environmental policy tools is crucial for the achievement of the Sustainable Development Goals.

For instance, a common and increasingly popular intervention to address the externalities associated with sustainable forest and land management is Payments for Ecosystem Services (PES). PES, in the larger sense, is an instrument to incentivize individuals or communities to engage in conservation activities such as reduced deforestation, afforestation, adoption of climate-smart land use practices, or watershed protection, for instance. The rationale behind such a policy instrument is that, without compensation, the targeted groups would have to incur the costs of providing the ecosystem services on their own while they typically reap only a small share of the conservation benefits, leading to under-provision of those services despite their benefits for the community. By offering financial compensation, conditional on environmental service delivery, PES is thus a means for changing the resource owner's cost-benefit

evaluation outcome in favor of conservation. While this makes economic sense, and while several countries are adopting PES as part of their REDD+ strategy, rigorous evidence on their effectiveness remain limited (Samii et al. 2014b). There are also several issues that may arise in the design and implementation of PES schemes. Understanding how the details of implementation of PES can affect their effectiveness and efficiency is important, both in terms of resource conservation and poverty reduction.

DIME's E&E program is contributing to filling this knowledge gap through a growing portfolio of IEs on PES in Mexico, Burkina Faso, Ghana, and Uganda, touching on key unanswered questions related to the tool.

Are PES effective at reducing deforestation?

Our first completed IE on PES studies the Mexico's Federal Payments for Environmental Services (PES) program, with a focus on the 2011-2014 cohorts. The evaluation seeks to understand how the program has affected land-cover management activities and land cover change, communal social capital, and socioeconomic indicators of participant communities and households, comparing outcomes for beneficiaries and similar rejected applicants close to program scoring cutoffs for each state, year, and sub-program. The main findings include significant increase in land-cover management activities suggesting the program effectively generated behavioral changes at the community and household levels that support the provision of ecosystem services, thereby providing supportive evidence for its positive impact on reducing deforestation.

Do PES effects last beyond the end of the program?

What happens when the payments are withdrawn matters for the decision to adopt PES as

policy tool for natural resources management. Therefore, DIME's research is also exploring the dynamics associated with incentivizing long-term behavior change towards conservation. For instance, in Uganda, we are exploring the impact of PES after incentives are removed. Does deforestation remain low, return to pre-intervention rates, or increase to catch up with total deforestation in control areas? Each scenario has plausible justifications, but results in very different interpretations around the overall role that PES schemes can play to mitigate our impact on the climate.

How to choose the right compensation scheme?

Finding the appropriate compensation level remains one of the main implementation challenges in PES schemes. This question is at the heart of an ongoing DIME IE in the context of a program with private landowners in Ghana. Potential participants in PES schemes differ in the opportunity costs of providing the services, and such information is private to them. The challenge of the PES implementer is how to set the payment so that it is high enough to attract and enroll as many participants as possible on a limited budget. The findings indicate that, when directly proposed a price, 100% of farmers sign up to participate in the program and be paid conditional on planting trees on their land. Sign-ups are around 10% lower when the farmers are first asked to think about how much they would need to be paid before being told what the offer is. These results suggest that the way offers are framed can influence whether the most enthusiastic participants are the ones included from a fixed budget for payments.

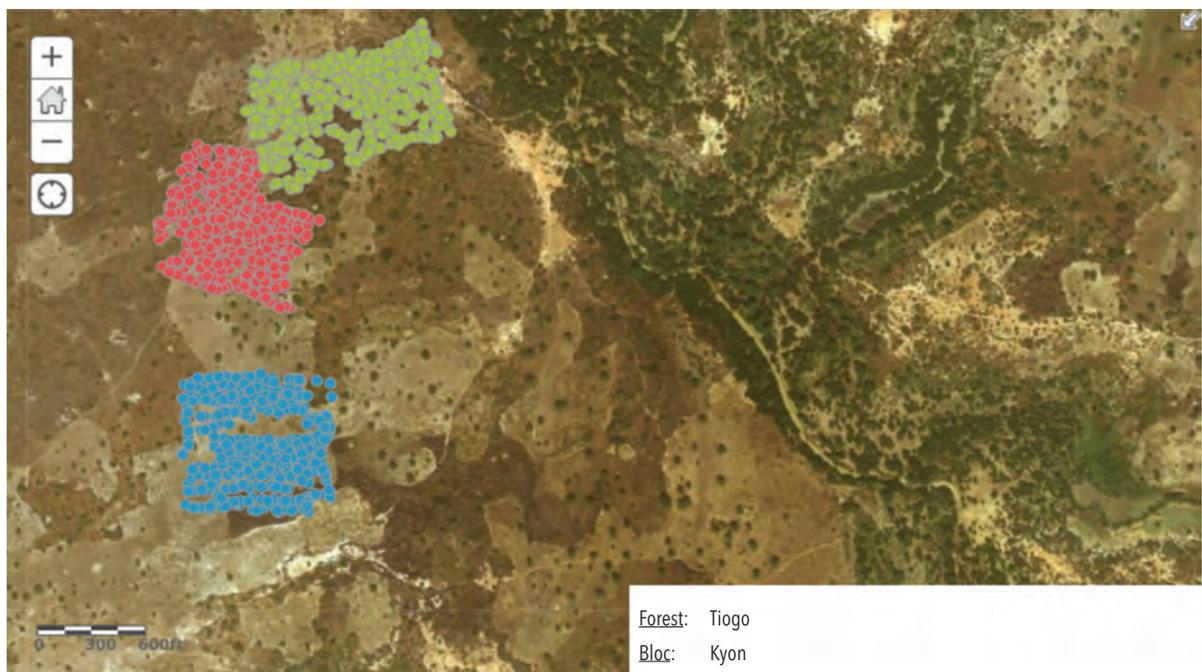
How does contract design affect PES outcomes?

The structure of the PES contract can also greatly influence its performance and overall efficiency. When group dynamics are involved

as in community-based PES, coordination failure or social dilemma can negatively affect the overall outcome. To unravel this question, DIME is implementing an IE embedded in the Burkina Faso Forest Investment Program, which is part of the national REDD+ strategy. Every year the Burkina Faso government organizes afforestation campaigns targeting selected gazetted forests in the country. As part of this, communities living around the forests are invited to participate in tree planting activities in well-defined areas of the forests in exchange for immediate cash. After tree planting, groups of individuals are enrolled into PES contracts, whereby they will receive additional payments based on tree survival rates. The DIME IE is testing the impact of alternative contract design options on

group performance measured by tree survival rates. More specifically, we are testing threshold-based payments versus piece rate payments. For this, we have georeferenced every seedling that was planted during the August 2017 afforestation campaign and have collected data on the specific groups of people selected to take care of those trees (see figure 6.4). Verification of survival rates will take place in May 2018 and people will receive the corresponding monetary compensation based on the type of contract they were given and the survival rate observed. The results of this evaluation will provide crucial information to the Burkina Faso government about how to maximize the outcome of their PES schemes through better contract design.

Figure 6.4 Reforestation Campaign 2017, where 1,330 trees were planted across three plots in Tiogo forests, Burkina-Faso. Tiogo in Burkina-Faso is one of 11 forests that were subjects to reforestation activities during the 2017 campaign. We counted 33, 547 trees planted overall across the 11 forests in August 2017.



Going Forward

Since the program began three years ago, the focus on using rigorous evidence in the energy and environment sectors has only modestly increased. It still requires a more concerted effort to catch up to other evidence-led sectors, like education and health. The mapping of evidence to development projects is currently skewed in favor of subtopics that are more amenable to impact evaluation. While the impacts of energy access have been a preoccupation in current economic literature, the reality is that the vast majority of development funds are directed towards generation and supply.

Tackling questions on the drivers of energy availability to connected customers will be the primary focus of the energy agenda moving forward. This is more aligned with the major development challenges in the sector. The program aims to work with utilities and other service providers to explore the interplay between pricing, service delivery guarantees, billing and payment schemes, and enforcement. This will help identify the bottlenecks and associated solutions to optimally utilize electricity infrastructure and provide reliable energy to households and industry.

For environment topics, we aim to expand the focus area beyond financial incentives to also include co-management practices and regulatory influences to better represent the major development tools available to practitioners and policy makers.

6.5 Financial and Private Sector Development

The literature on firms has a lot to say on the key determinants to productivity, but much less so on how to increase firms' productivity

(Syverson 2011). In fact, all the efforts made to increase firms' capabilities through skill training and matching grants reveal that increasing firm's productivity and performance is not an easy task. In reality, this challenge has to do with identifying ways to increase efficiency of factors of production that are somehow being misallocated; and finding innovative ways of combining factors of production to increase growth potential. Identifying effective mechanisms to increase technology adoption among firms at scale is paramount given the role played by micro, small and middle-sized enterprises (MSMEs) in job creation and growth.

The Framework

The Trade and Competitiveness (T&C) agenda in DIME is structured to test and identify effective ways of increasing firms' productivity through both efficiency gains and shifts in the production frontier. Efficiency gains are understood as changes in the production process to help firms move closer to the efficient production frontier. Many factors can prevent firms from using resources efficiently, such as: (1) market imperfections, e.g., when firms underinvest in training because they cannot fully internalize the benefits; firms cannot access subsidized public credit lines because are informal; or firms cannot access credit because do not have collateral or credit history; (2) behavioral biases, such as misperception of returns associated with a given business practice or lack of motivation and incentive to adopt better production process (Gibbons and Henderson 2012; Nguyen and Nguyen 2016); and (3) organizational barriers that prevent firms from adopting new technologies (Atkin et al. 2017) and using inputs more efficiently. In this light, finding ways to overcoming barriers to adoption of new (though proven) technologies is key to firms' (and economic) growth.

“ [The IE] helped assess that firms with 5 to 20 employees should not be targeted [by] the program. ”

IE CLIENT, IMPACT OF MANAGEMENT UPGRADING ON SMALL- AND MEDIUM-SIZED COMPANIES IN BRAZIL (BANCO DO NORDESTE)

In this context, for a *technology change* to materialize, firms need to be first and foremost willing to adopt the new technology. Once the barrier to adoption is surpassed, firms need to be ready to take full advantage of the new technology, otherwise, the shock in the production process won't necessarily be conducive to higher productivity. Technology change and productivity growth go hand in hand with technology adoption.

Interventions aimed at shifting the production frontier (e.g., those caused by improvements in business regulations, innovation, and infrastructure) may show limited effects if firms do not

figure ways to use inputs efficiently. Even so, interventions intended to shift the production frontier tend to be more difficult to evaluate with RCTs, either because changes in regulations are not necessarily theoretically grounded or because some policies are highly likely to have general equilibrium effects (spillovers). IEs can still be used, for instance, to help quantify spillover effects and increase compliance with a new law/rule when lack of information or problems with the enforcement mechanism are at play, as shown in the work done by the ieConnect team.

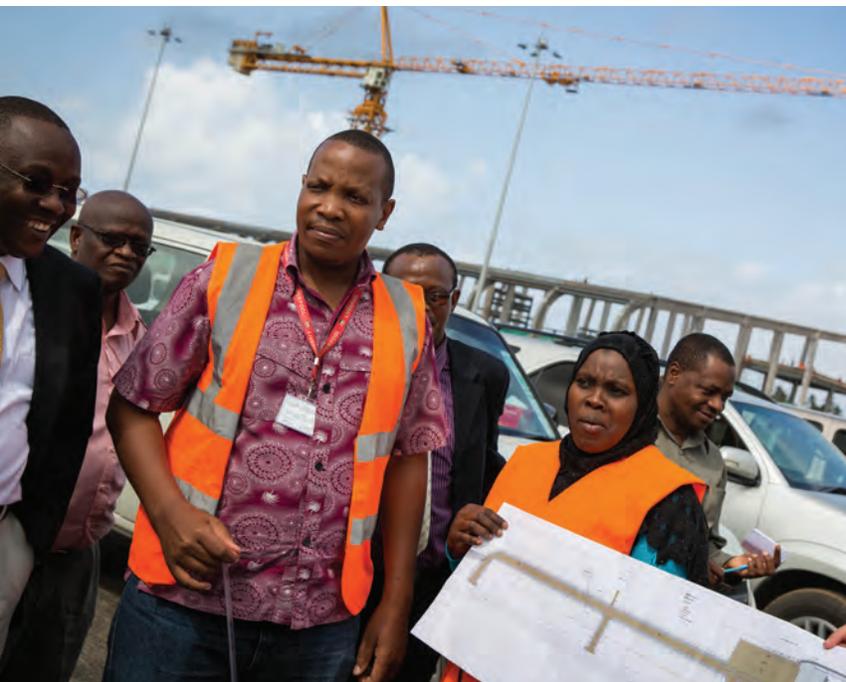
DIME Trade & Competitiveness (T&C) Program

Background

Even though the T&C agenda in DIME accommodates both efficiency gains and growth to increase firms' productivity, the projects in the pipeline tend to be more concentrated on testing policies aimed at raising firms' productivity through the former. That has not always been so.

In 2010, DIME and the Finance and Private Sector Development (FPD) and Gender groups at the World Bank co-organized the first IE workshop in the city of Dakar. The event focused on interventions to improve firm capabilities through matching grants and other services to MSEMs, such as job training.

A year later, DIME and the FPD group co-organized the second IE workshop in the city of Rio de Janeiro, Brazil. The event focused on supply-side interventions and financial literacy/inclusions. World Bank practitioners held a strong view that the main barriers preventing MSMEs to grow was related to lack of capital or skilled labor. The initiative of supporting the financial literacy agenda was based on the idea that individuals,



youth and households, need financial education to deal with the increasingly complex financial system, and that better financially-educated agents could potentially affect macroeconomic effects by increasing savings rates, thereby reducing household overborrowing and indebtedness.

Several training programs have been evaluated since 2010, but almost all impact evaluations found null effects on jobs creation and firm productivity (McKenzie and Woodruff 2012). Several matching-grants programs, on the other hand, could not even be evaluated because of low take up rates in the programs (Campos et al 2013). Despite the disappointing results documented by the first wave of rigorous evaluations in this sector, the failure was not in vain. In fact, the impact evaluations revealed that matching grants are not “free money” (as paperwork can be cumbersome); local market for services (e.g., local consultants) might be underdeveloped; firms do not necessarily know what they don’t know; quality of implementation of training programs matter; and, more importantly, the evaluations identified what should be avoided in supply-side interventions. In a nutshell, these failed attempts showed that supply-side interventions are unlikely to succeed if constraints on the demand-side that hinder participation and adoption of new technologies are ignored.

In 2012, the Bank decided to attack the constraints that are outside firms’ control. DIME and the Business Climate Unit at IFC co-organized an IE workshop with the objective geared towards generating rigorous evidence on how incentivize small informal firms to formalize their businesses. At that time, there was a consensus that excess of taxation and red-tape costs were important explanatory factors of firms’ decision to remain informal, so reducing business registration costs

and simplifying taxes should lead to higher formalization rates and business growth. This initiative can be seen as a first attempt to tackle other constraints firms may face, other than inadequate inputs. The experience with easing firms’ formalization through tax simplification and/or reduced registration costs brought our underlying assumptions into question. The overwhelming evidence produced since then shows that most small firms do not want to formalize their business, even when registration costs are fully subsidized. In fact, DIME learned that if informal firms have access to credit and other types of services and do not face major penalties for being informal, then cutting taxes and registration costs shouldn’t produce the expected effects on formalization rates. Interestingly, the research produced in this area also shows that even among firms that do formalize, the impact on performance indicators (e.g., sales and revenues) are either small or null (Bruhn and McKenzie 2014). The knowledge generated by the rigorous evaluations led to a change in IFC’s approach towards small informal firms. It no longer sees formalization as a stepping stone for growth of informal firms. The challenge now is to find ways of helping informal firms to become more productive. DIME’s evaluation of a reform in the Senegalese justice system exemplifies this view well. The IE shows that giving judges the duty and power to meet a deadline led to a more efficient justice system enhanced economy’s welfare. This new approach well reflects what has been the holistic model embraced by the T&C since 2015, as is discussed below.

Current Portfolio

Improvements in the investment climate are still central to the Bank’s agenda, but the projects in T&C are now looking towards better understanding issues that are both under and beyond firms’

control in order to better attack the constraints firms face in efforts to grow.

An example of how to better tackle participation rates in combination with supply-side intervention is that of our IE in Brazil, where medium-sized firms will be randomly split to receive information, information plus training, and a placebo intervention. As evidence shows that take-up of training programs is low, this IE aims at measuring participation incentives. One group will be offered 40 percent subsidy on training costs in addition to information, while the second will be offered 80 percent subsidy plus information. The training will be provided at firm location and consultants will pay weekly visits to monitor the adoption of good practices. In addition to testing the impact of different subsidy rates on training take-up, this design will help the implementing institution (a regional development bank) identify how much firms are willing to pay for a training program. Another example, still in Brazil and similar in spirit, is the evaluation of the SEBRAE's consulting service program for micro and small firms. The program reaches thousands of firms in the metropolitan area of Rio de Janeiro (RJ) and consists of a free-of-charge one-hour visit of a SEBRAE's agent made to a firm. The agent runs an in loco diagnostic of the business practices adopted by the firm, thus ensuring a high participation rate into the program. DIME is working with SEBRAE's regional office in Rio de Janeiro to test different ways of encouraging firms to follow up on the problems as identified in the diagnostic stage. We are shocking firms with an informational package and a more handholding approach in order to understand whether firms do not adopt best managerial practices due to lack of information or lack of skills.

In Georgia, DIME is testing whether a training program focused on helping firms use

e-commerce to widen their consumers base. Since new entrants in e-commerce may need time to build reputation, this IE will also try to relax some constraints on the demand side firms may face by generating the first ten orders to a randomly selected group of firms.

Overall, the current T&C portfolio accounts for 28 IEs, distributed across 23 countries (three in the preparation phase, sixteen ongoing, and nine completed). The total estimated budget is US\$18 million, of which 30 percent is funded through i2i. Over half the program (16 IEs) evaluates World Bank projects, representing a total of US\$630 million in loans. In terms of outputs, the T&C teams have produced eight reports, ten working papers, and three publications.

Going Forward

The lessons generated so far have substantially shaped the current T&C Global Practice (GP) agenda. A more structural approach was envisioned in the IE workshop held in Istanbul in 2015 and it has matured since then. The Competitiveness Policy Evaluation Lab in T&C GP (ComPEL) and DIME teams pushed projects selected for funding support to pay special attention to some of the cross-cutting themes that were identified as major issues in previous IEs such as: (i) potential spillovers occurring in the market; (ii) low take-up rate of supply-side interventions; and (iii) using the IE to improve program's targeting.

Recent IEs that managed to measure spillovers found that some interventions may benefit the treatment group of firms at expenses of the control group (McKenzie and Puerto 2017). This is something that previous studies completely ignored. A recent IE of a matching grant program also showed that with simplified application



process, take-up rates can be high enough to permit the evaluation to be carried out (McKenzie, Nabila and Cusolito 2015).

Those cross-cutting issues gained momentum in the T&C program and played a critical role in the selection of 18 projects that attended the latest IE workshop in Mexico City between February 27 and March 2, 2017. Three priority areas were recently identified by the T&C GP as strategic for knowledge generation through rigorous impact evaluations:

- i. Firms' access to markets and spillovers.
- ii. Identification and support to high-growth small and medium-sized enterprises (SMEs).
- iii. Regulatory efficiency.

There is an increasing focus on how to improve firms' linkages to both consumers and larger firms in the global value-chain. Interventions will then

have to look carefully at both supply (for example, training and matching grants) and demand-side/ or institutional constraints (for example, technology adoption and diffusion, access to markets, i.e., business-to-business and business-to-consumers, and regulatory environment), in order to (i) maximize the chances of intervention success at least on implementation grounds, and (ii) increase the odds of policy effectiveness. Selected projects (four in total) were encouraged to think creatively about how interventions could be designed to deal with cross-cutting issues, and integrate complementary tools (e.g., machine learning, psychometrics or business plan competition) to help pre-identify high growth potential firms. Projects aimed at exploiting changes in regulations were also encouraged to think of complementary interventions (ideally through experiments) to increase compliance and strengthen enforcement mechanisms.

Moving forward, we intend to explore the complementarities of research agendas in DIME to widen our knowledge of how large-scale infrastructure interventions affect firms and workers, who and why they benefit most, and what types of skill complementarity policies should be put forward by government and international organizations to help firms and workers take advantage of skill-biased technological changes.

6.6 Transport

Transport and associated infrastructure are critical investment sectors for economic development. A large percentage of lending by the World Bank and of development finance and grants from other donors is aimed at this sector. However, there is limited evidence of their impact through rigorous experimental or quasi-experimental evaluation. In fact, impact evaluations on transport accounted for fewer than 1% of impact evaluations globally between 1981 and 2012 (Cameron et al. 2016).

Motivated by the dearth of impact evaluations in transport, the goal of the ieConnect for Impact program is to generate evidence on the impact of transport policy and investments, including indirect benefits, at sufficient scale to substantially improve the evidence base for policymaking. DIME's ieConnect for Impact program, which was developed in collaboration with the World Bank's Transport & Digital Development Global Practice and DFID, aims to fill this gap by linking projects with research teams and enabling them to develop innovative and rigorous impact evaluation designs in selected evidence gap areas. The focus is on identifying and estimating the impact of transport investments themselves, as well as developing and testing ancillary interventions that can maximize their impact.

The ieConnect for Impact program began in 2015 as part of the i2i program. During the first phase, two major workshops were hosted in Rio de Janeiro to launch the program and in East Africa to build capacity and strengthen designs on new transport impact evaluations. A number of impact evaluations were developed in the first phase and the program has now grown into its own stand-alone program of transport impact evaluations.

Expansion of the ieConnect Program

DIME's original impact evaluation work in transport has been supported by funding from DIME's main i2i program, while 2017 marked an important new phase in this program with the creation of a portfolio dedicated exclusively to transport. Transport IEs typically are on a larger scale and have more intensive data requirements than IEs in many other sectors. This new program, referred to as Phase 2 of ieConnect, can accommodate these scale and data requirements thanks to a new grant from DFID.

Phase 2 was officially launched at the ieConnect for Impact workshop held by DIME and the Transport & Digital Development Global Practice in Lisbon, Portugal in July 2017. The workshop brought together 126 participants on 22 different project teams comprised of government officials and clients, IE specialists, transport specialists, and academics. The 22 projects were strategically selected in cooperation with the Transport & Digital Development Global Practice for their IE feasibility and relevance to the knowledge priorities and were financed or co-financed by a variety of institutions including the World Bank, DFID, European Union, African Development Bank, Asian Development Bank, and Islamic Development Bank.



In addition to launching Phase 2, the workshop objectives were to build capacity for clients to generate and use rigorous evidence in the transport sector. The agenda consisted of training sessions, targeted project “clinic” sessions, and panel discussions on development corridors, gender, urban mobility, road safety, technology and data, and the environment. The workshop also provided an opportunity for IE teams to report back on early data collection and to share preliminary results from three Phase 1 ieConnect impact evaluations in Tanzania, Brazil, and Rwanda.

Following the workshop, a call for expressions of interest was made in August 2017. Proposals were received for 29 new impact evaluations. These proposals then went through rigorous technical and policy review and the i2i/ ieConnect Technical Committee selected 22 projects on a competitive basis to begin concept note preparation. Project teams that were selected to receive seed grant funding were notified and, in

“The [IE] results are expected to inform adjustments and potential operational/fare interventions that can help optimize the impacts of the BRT on travel by potential public transport riders (and particularly poorer riders and women) as well as on urban development changes. Very good, innovative, and collaborative work by the coordinated DIME/GP team.”

IE CLIENT, IMPACT EVALUATION OF THE DAR ES SALAAM BRT SYSTEM

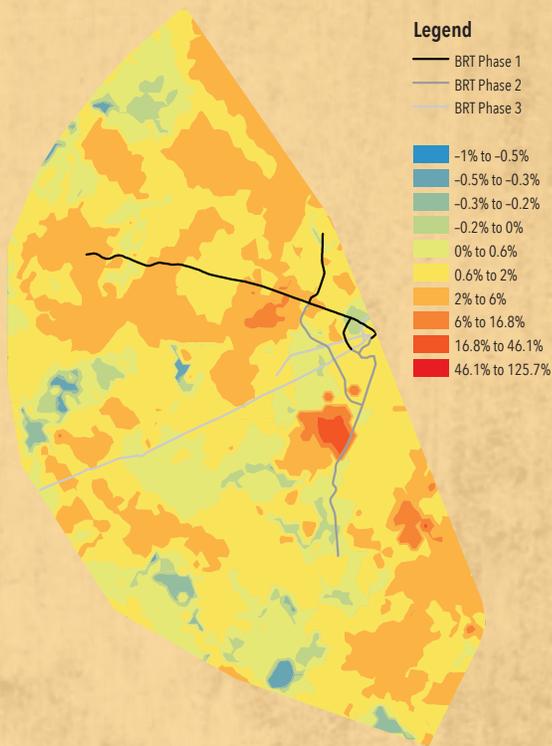
UNDERSTANDING COSTS OF GENTRIFICATION IN TANZANIA

DIME research builds evidence for transport policy

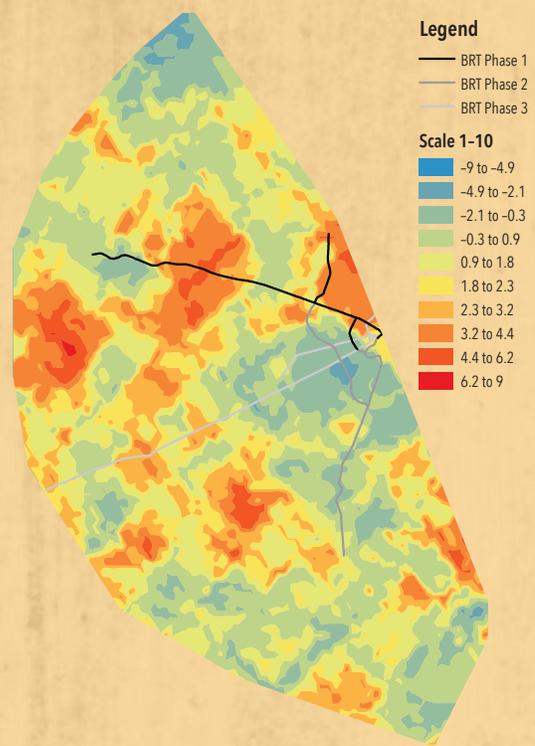


The maps below show current and planned Bus Rapid Transit (BRT) lines in Dar es Salaam. Using an innovative spatial triple-differences approach to evaluate transport investments, DIME found that residents who live close to the BRT line were 4 percent more satisfied with commute time, but also experienced a 6 percent increase in their monthly rent payments.

Estimated Monthly Rent Per Person — % Change Baseline To Midline



Change In Satisfaction With Time To Main Job Baseline To Midline



December 2017, work began in earnest on concept note development and IE implementation for the Phase 2 IEs.

The ieConnect Agenda

The ieConnect for Impact program seeks to generate rigorous evidence on the effects of large-scale development investments in the transport

sector. The program targets the subsectors of urban mobility, transport corridors, and road safety, with a thematic emphasis on gender, female economic empowerment, data systems, and measurement frameworks. The geographic priorities include, but are not limited to, Africa and South and Southeast Asia, with a goal of having 50% of the IEs in the category of fragile and conflict-affected states.

Evaluations not only aim to improve the theory and evidence base of project design, but also to structure experimentation during project implementation so that it can inform mid-course decisions and improve delivery and effectiveness. The ieConnect program also aims to influence the design and implementation of transport projects outside of the IE itself, and to partner with other multilateral development banks and bilateral development partners at the operational level of the impact evaluation. This transfer of knowledge and tools to organizations outside of the World Bank through partnerships on IE projects is a core mission of the ieConnect program. At this time, there are ongoing or planned IEs of investments from other Multilateral Development Banks (MDBs) within the portfolio including the Interamerican Development Bank, the African Development Bank, and the Islamic Development Bank.

The ieConnect program is a joint initiative with the World Bank's Transport & Digital Development Global Practice, whose initiative, Sustainable Mobility for All (SuM4All) is framed around the goals of access for all, efficiency, safety, and green mobility. The ieConnect program recognizes that along with filling knowledge gaps, there is also an emerging need for taking a systematic approach to the achievement of sustainable mobility and so the impact evaluation program has been aligned to be consistent with these goals.

Active Impact Evaluations

Phase 1 of ieConnect first launched in 2015 with seed funding from i2i. This set of eight evaluations range from exploring the economic impact of road rehabilitation and its effect on women's access, to education and health services in

Peru, to the impact of road access on household income and price dispersion across national markets in Rwanda, to harnessing big data to better understand safe driving behavior on semi-formal bus transit systems in Nairobi. While still in progress, a number of projects have results from initial data collection or from midline surveys. For example, an impact evaluation on the Bus Rapid Transit (BRT) system in Dar es Salaam has early results that housing prices have changed as a result of the BRT expansion. This study aims to help policy makers understand the dynamics of displacement, earnings and jobs in the context of BRT investments. Another IE is testing interventions to reduce the high economic costs to women associated with violence and harassment in public transport in Rio de Janeiro. Initial findings have quantified the willingness-to-pay to avoid harassment.

Following the Phase 2 expansion of the program, 22 projects have concept notes under development. These proposals cover a wide range of transport topics with a significant number of projects added in the road safety and transport corridors thematic subsectors, some of which are discussed below. New IEs in the urban mobility and gender subsectors are discussed in the chapter on Shared Prosperity.

Road safety is a key area that has been largely unexplored by IE research work in developing countries. Worldwide, road traffic crashes kill more than 1.3 million people every year and severely injure another 50 million or more, with the vast majority of fatalities (92%) occurring in developing countries. Road traffic injuries are estimated to cost developing countries 5 percent of GDP, are a key obstacle to economic development, and a burden on public health. Phase 2 of the ieConnect program has greatly expanded



DIME's road safety research agenda. Proposed IEs evaluate a range of interventions, from physical road safety infrastructure, to information campaigns and driver training programs. One project that is currently preparing the concept note will evaluate the impact of post-crash interventions in Malawi on disability and mortality, an understudied theme that is one of the pillars of the UN Decade of Action for Road Safety. Another concept note under development will evaluate the effect of changes to physical road design, driving training, and passenger empowerment on speeding and traffic fatalities and injuries in Liberia.

Phase 2 also significantly expanded the number of IEs on transport corridors. This program examines the impact of corridor investments on trade-related outcomes, agglomeration, and access to basic services in less developed and

remote regions. For example, a new IE in this sub-sector is evaluating the Eastern Africa Regional Transport, Trade and Development Facilitation Project in Kenya. This IE will provide evidence on the impact of both hard and soft infrastructure programs and impact of complementary trade facilitation policies on the welfare of communities along the road and on trade volumes between Kenya and South Sudan. Another IE in the concept note development phase examines the effect across markets of a highway upgrading project in Tunisia. Improving roadways lowers transport costs which can, in turn, increase market access and economic opportunities in remote markets. However, reduced commuting costs may also encourage out-migration from lagging areas. This can benefit those who leave, but also requires that destination markets can support newcomers. A better understanding of the effects of improved market connectedness

and its differential impact across communities will provide evidence that can be used to design and coordinate complementary interventions to ensure gains are shared across the region.

Developing Data Systems

Following the success of the data collection system developed for the rural roads impact evaluation in Rwanda through critical support from the European Union, the ieConnect program has prioritized the development and expansion of data systems using new technologies to harvest large amounts of data at higher frequency or with greater geographic coverage and spatial resolution than traditional survey methods allow. A key methodological issue within transport impact evaluation has been identifying the impacts of investments when selection of project sites are non-random and often integral to project design. The data systems which we seek to develop allow research teams to mitigate these methodological challenges and to use innovative IE methods.

In Nairobi, ieConnect research teams are developing a detailed real-time geo-referenced crash map. Currently under development, the first stage in this project will generate high-frequency data on crashes and crash density around urban hotspots and use a real-time verification process to record different characteristics related to the crash. Once this process has been successfully tested for Nairobi, the code will be open and available so that this type of data collection can be expanded to other countries. The second part of the project will use the information on the characteristics of the different crash sites to conduct a series of experiments on the behavioral and infrastructural interventions that can be used to improve road safety and mitigate

the health impacts of increasing motorization in Nairobi.

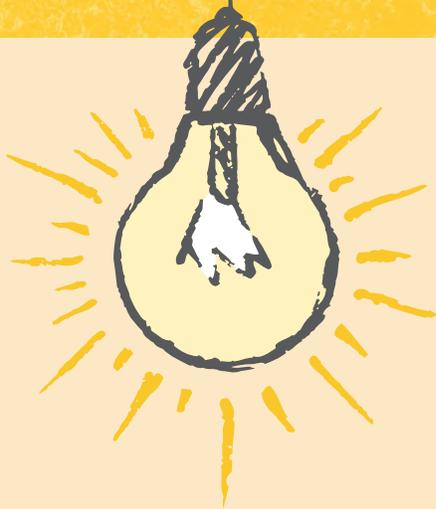
Going Forward

Coming off an exciting year of scaling up the ieConnect program, in 2018 the effort will be focused on (a) developing full concept note proposals for the new IEs that recently received seed funding, (b) supporting ongoing IEs through implementation, (c) designing and advancing data systems, and (d) developing measurement frameworks in road safety and transport corridors.

In the first half of 2018, research teams for the 22 IEs, selected from the 2017 call for proposals, will be developing and submitting IE concept notes. After technical and policy review, a subset will be approved for full implementation grant funding. In 2017, ieConnect hired three new team members whose expertise in economic geography, data science, and geospatial analysis, has allowed us to increase our support to projects across the portfolio. A major focus for the coming year is to continue to promote and support the development of data systems that can increase the rigor and success of IEs in the transport sector and the use of innovative IE methods. The ieConnect team will also develop framework papers that will inform the theory of change across all IEs within the corridor and road safety thematic subsectors in the ieConnect portfolio. Finally, we will continue to expand gender-specific interventions and gender- and age-disaggregated data collection and to increase the work on fragile and conflict (FCV) or conflict-affected settings. We will also continue to build partnerships and collaborations with other MDBs and bilateral development partners on developing and implementing impact evaluations in transport and will ensure that knowledge-sharing is occurring across organizations.

PARTNERSHIP SPOTLIGHT

IECONNECT FOR IMPACT: DEEPENING ENGAGEMENT THROUGH A SECTOR-BASED FUNDING MODEL



Innovations in data and technology have created new opportunities to work on the evaluation of large infrastructure investments. In partnership with the UKAID's Department of International Development (DFID), the ieConnect for Impact program is generating a significant body of evidence to improve our understanding of the transformational potential of transport on global economic development. This program will generate data and experimental and quasi-experimental evidence of the impact of transport policies and investments to transform the way we think of mobility as an economic force and at a scale that will substantially improve the evidence-base for policy making. The strategic advantages of this sector-based funding model include deepening knowledge in important evidence gap areas, leveraging new data opportunities, and building partnerships across organizations, thereby inducing global shifts in transport policy and financing.

Taking a systematic approach, ieConnect is building evidence across a large portfolio of 30 impact evaluations in 19 different countries. This program will increase the depth of knowledge in the thematic subsectors of urban mobility, transport corridors, and road safety. The ieConnect team is developing framework papers to take stock of existing evidence and create the conceptual framing to strategically align activities across the portfolio. As results are generated across IEs, the lessons learned in each project will be synthesized in thematic knowledge reviews for policymaking, and in cross-cutting reviews on transport in fragile and conflict affected states and on transport and female economic empowerment. It is with this dedicated, sector-specific funding that the program can take advantage of these strategic opportunities to work across projects to generate evidence and lessons learned at the portfolio level.

Transport IEs are typically on a larger scale, have more intensive data requirements than IEs in many other sectors, and use a mix of existing and newly collected high frequency data and surveys. Following the success of a data collection system developed to monitor the impact of rural roads on market prices in Rwanda, the ieConnect program has prioritized the development and expansion of integrated data systems. Using new remote sensing and crowdsourcing technologies to harvest large amounts of data will provide information at higher frequency, greater geographic coverage, and higher spatial resolution than traditional survey methods allow. This, in turn, can build local capacity and strengthen government agencies' basis for evidence-based decision-making for transport investments in the long-term.

To evaluate the impact of transport infrastructure requires a region- and country-wide approach and coordination across donors investing in this sector. The program has reached out to large financiers to build a coalition of partners who want to explore the range of socio-economic impacts of transport investments. The approach has been successful in soliciting interest and engagement from the African Development Bank, Asian Development Bank, Islamic Development Bank, Inter-American Development Bank, Development Bank of Latin America, Export-Import Bank of China, Export-Import Bank of Korea, the European Union, and DFID. The ieConnect program aims to influence the design and implementation of transport projects by partnering with donor institutions at the operational level of the impact evaluation in order to transfer knowledge and tools across organizations. For example, in Senegal, researchers are working with the World Bank, African Development Bank, and Islamic Development Bank to evaluate the impact of the express train and bus rapid transit systems. These projects span donors, but jointly affect mobility and quality of life. Because individual transport projects are typically part of a broader transport program, partnerships across donor projects to structure experimentation during project implementation can inform mid-course and scale-up decisions, and lead to improved delivery and effectiveness.

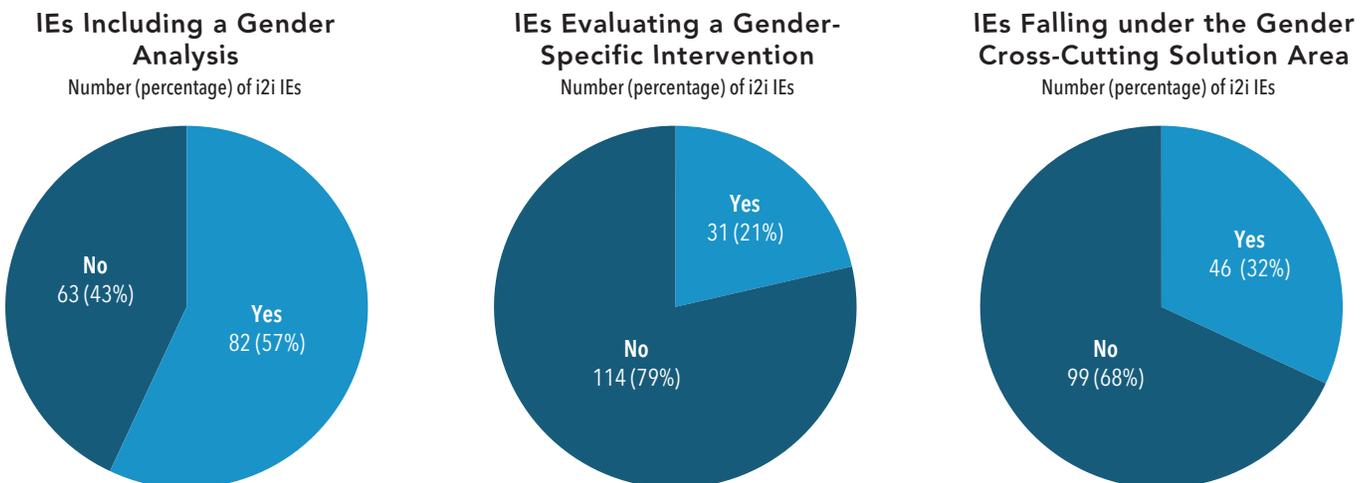
6.7 Gender

While gender equality is a fundamental human right, it is also pivotal to achieving prosperous societies and sustainable development. It is a necessary foundation for achieving other development goals such as health, education, and growth. In short, it is smart economics. But while progress has been made, much more remains to be done to ensure that women are equally included in the political, social, and economic life of their societies. In this context, gender equality for development is a core theme of DIME's work under the Gender Program. Closely aligned with the Sustainable Development Goals (SDGs), specifically Goal 5 on Gender Equality, it stretches across almost all indicators (gender-based violence, social protection, and political participation among others). Alongside SDGs, the portfolio addresses UK-DFID's priority areas under the gender strategy through evaluations ranging from those aimed at preventing risky sexual behavior to those that facilitate access to financial and human capital, address GBV in public and private spheres, and those that promote education among women and girls.

i2i Gender Program

DIME's gender-related analytical work is organized across two streams: those that integrate gender disaggregated analysis across the IE, and those that are specifically designed to address constraints that women face. Moreover, these two categories include rigorous evaluations of policy actions that aim to relax supply-side constraints (e.g., improving service delivery for clean water, sanitation and maternal care) as well as market and institutional constraints (e.g., reducing systematic differences in earnings). As it stands, over half of DIME's current portfolio includes gender-disaggregated analysis, while 22% is testing interventions tailored to address gender issues. More specifically, the thematic coverage of the i2i gender program seeks to fill the gaps in four areas identified in the 2012 World Development Report (WDR), Gender Equality and Development: (i) human capital, (ii) economic productivity, (iii) access to finance, and (iv) empowerment. In DIME's gender portfolio there are 31 IEs that address these constraints, some of which cross-cut through the four areas, hence providing for a comprehensive outlook of

■ ■ Figure 6.5





the impacts. The complete list of ongoing and forthcoming IEs that specifically target these constraints can be found in appendix 2.

Theme 1: Addressing Human Capital Gender Gaps through the Lifecycle

Despite being the majority, and a crucial part of economies, women remain disempowered and excluded. Extensive unpaid care work leaves no room for women to seek employment opportunities outside of their homes, while low school completion rates leave them unqualified and unprepared for paid jobs. Financial and business literacy is even less pronounced among women, preventing them from capitalizing on their entrepreneurial potential.

While the gender gaps in human capital are well-documented, little is known on how best to close them. The i2i research agenda focuses on how to design interventions that reduce women's vulnerability to shocks that disrupt human capital acquisition. An IE of a vocational training program in Malawi (Cho et al. 2015) found that family obligations limited participation and resulting skills development for young women. Another IE testing the impact of a business literacy course for female micro-entrepreneurs with relatively low education in five different states in Mexico, finds significant improvements of managerial skills. As a result, nine new states in Mexico have submitted proposals to expand the program to their states. Another study in Nigeria, the MAFITA program which targets youth at risk, evaluates the impact of a package of entrepreneurship interventions (including apprenticeship and vocational skills training, entrepreneurship training, and access to finance interventions) on human capital, labor market,

economic welfare outcomes, and socio-psychological welfare, among others. This evaluation includes an innovative gender component which comprises of an additional home-based apprenticeship element. This female-only intervention that targets women in their homes has been specifically designed keeping in mind the local cultural norms that restrict women from working outside their homes. This is an attempt to ensure that women receive the same opportunities for human capacity building as their male counterparts in an effort to close the economic employment and wage gaps between men and women.

Theme 2: Economic Opportunities

Women's economic empowerment is pivotal to sustainable development, and pro-poor growth. Greater gender equality enhances productivity, makes institutions more representative, and leads to better development outcomes. Women's access to economic opportunities however remains low due to a wide range of constraints, such as lower access to employment production inputs, business linkages, information and essential social services among others. Against this backdrop, DIME has carried out a wide range of evaluations, and already has substantive results on women's access to economic opportunities.

For example, an IE in Kenya tested a recent policy innovation, known as "micro franchising," which provides unemployed participants with a proven business model and the specific capital and business linkages based on the hypothesis that many unemployed youths, would like to be generating income, but lack both experience to be competitive and the financial and human capital. Early results from the study found that for

“The IE led to the development of a new IE on female entrepreneurship, targeting women beneficiaries from the same project. We are taking gender more and more into account in all our approaches to providing services. The experience of some beneficiaries, especially women in THIMO, highlighted the importance of women's entrepreneurship.”

IE CLIENT, LABOR-INTENSIVE PUBLIC WORKS, DRC

young women the program increased self-employment. This is an important finding considering around 55 percent of urban women in Kenya aged 15 to 25 are unemployed. Similarly, in Kenya, DIME conducted a lab experiment prior to launching a Business Plan Competition and found a significant gender difference in competition entry when individuals faced a mixed-gender competitive environment: only 31 percent of women chose to compete as compared to 69 percent of men, and they were more likely to do so when faced with female only competitors. In Malawi (BenYishay et al. 2016) and Mozambique (Kondylis et al 2014), i2i IEs showed that women can make effective extension partners:

they are at least as good as men at encouraging adoption of improved technologies. Throughout Sub-Saharan Africa, women are disproportionately limited in their land ownership and transfer rights. An IE in Benin showed that land demarcation increases soil fertility investment in female-managed landholdings, shifts household decision-making and reduces spousal conflict (Goldstein et al. 2015).

Regarding ongoing experiments, in Afghanistan DIME is measuring the impact of a program aimed at lifting the poorest out of extreme poverty by providing a way to transition into sustainable and profitable economic activities and linking them with microfinance programs. DIME applies the program in a setting where female labor force participation is among the lowest in the world (15%) and with a strong focus on supporting female-headed households, tackling multiple constraints simultaneously to provide households with a big push out of extreme poverty.

In Tunisia, Egypt, and Côte d'Ivoire, DIME has evaluated large-scale Labor Intensive Public Work (LIPW) programs as part of its continuous engagement in the social protection and labor related work, that the Bank has been implementing globally. The IE of the LIPW program in Tunisia's Jendouba Province suggests that an increase in the probability of female employment is significantly correlated with the intervention; however, the program also positively correlated with slight increase in experience of emotional violence. In Egypt, the evaluation found that the intervention was successful in reaching the targeted female participants suggesting a significant uptake. The evidence also suggests that the improvement in the economic situation of the households are concentrated

within the female participant subsample with them reporting increase in savings and improvement in psychological wellbeing. Lastly, in Côte d'Ivoire, the evaluation of the LIPW program suggests that the intervention has had positive impact in vulnerable population, including women, when it comes to earnings. The IE found that a program that targets the most vulnerable and reduces costs outside transfers would become much more cost-effective, i.e., break even, in about 3 years versus 23 years for running the intervention under the status quo. This has policy implications for governments on limited budgets aiming to maximize their social protection programs.

Reducing transaction costs by improving transport infrastructure has the potential to change the way women access markets. In Ethiopia, a large expressway construction is combined with the development of a large industrial zone. Since most of the employment in the industrial zone will be of young women, this will be an opportunity to study the effect of a large labor market shock (60,000 plus jobs over a period of several years) on young women's economic and social outcomes in the vicinity of the zone. A complementary intervention will be set up to experimentally study the role of skills, information, and access to employment opportunities. In Peru, an intervention to promote women's access to health services and education is being evaluated in the context of a rural road rehabilitation project.

The gender and urban mobility component has seen IEs study strategies to address gender-based violence in public transportation within the Dar Es Salam BRT project and the Dakar BRT project and to evaluate the impact of sidewalk rehabilitation on the most vulnerable users' mobility. This is following on existing IE work

done in Brazil, where sexual harassment on the metro system was quantified, both in terms of the number of incidences occurring and the willingness of women to pay to avoid harassment.

Theme 3: Access to Finance and Capital

In general, women are less likely than men to access information, productive assets, as well as credit, and capital needed to start a business, however small it may be. This gap can be associated with low experience and literacy rates, as well as operating in low-growth sectors. Women are also prone to gender stereotyping and biased lending when it comes to borrowing from banks or the informal sector.

Whether in the case of women farmers or micro-entrepreneurs, access to productive assets remains a severe constraint. An i2i IE in Rwanda tested targeted and pre-commitment savings accounts. Initial findings show that women are more likely to earmark their savings to buy durable goods, relative to men who invest in agricultural inputs, suggesting that intra-household bargaining over resources plays an important role in women's investment decisions. In India, an IE showed that women who participated in a women's empowerment and rural livelihoods program had improved access to loans, accumulated assets, and invested in education, which further made them feel more empowered. Early results from an ongoing IE in Benin examining several incentive mechanisms to attempt to get business owners to formalize found that male business owners formalized much more than female business owners. Ongoing analysis is exploring the reasons for lower formalization amongst female-owned firms. In the Dominican Republic, an ongoing IE is studying the impact of

financial literacy and job skills especially benefiting women on household finances management, savings, credit, usage of formal sector financial products, ability to search for, obtain, and retain formal employment, management of small businesses, new businesses opened, and income levels. An ongoing IE in Malawi on identification and fingerprinting, a topic that is at the heart of the developing agenda, is testing whether requiring authentication for transactions by fingerprinting alleviates access to credit more for females and improves repayment more for those that borrow. This is also important because it would make it impossible for male relatives to seize control of women's assets upon the death of the husband, as is common in Malawi. In DRC, Indonesia, Tunisia, and Egypt, i2i is supporting new cross-country evaluations of unconditional "cash injections" to women on the creation of sustainable livelihoods and on long-term poverty alleviation. The relative impact and complementarities across these interventions will be captured by the experimental design. Focusing on vulnerable women, the impact of these additional grants on long-run consumption and labor market outcomes will be compared to the outcomes of those who merely participate in short-term, labor-intensive works.

Theme 4: Promoting Women's Empowerment and Agency for Economic Development

A growing body of evidence shows that placing women in the center of the development agenda can increase efficiency in the management of institutions and resources, and that female leaders can have beneficial impacts on social norms. Women empowerment goes beyond economic constraints and includes their role in social and political life. The i2i research agenda thus focuses on using gender empowerment to combat

domestic violence, testing interventions such as cash transfers and active labor market policies to economically empower women and role of law and justice in achieving gender equality, among other areas.

Addressing gender norms, one specific study in Nigeria, measuring the impact of entertainment education through soap operas on attitudes and behaviors about safe sex and HIV testing, finds positive impact on both outcomes (Banerjee and others, forthcoming). In Pakistan, an IE is evaluating the impact of women inclusion mandates and ratification in village-level grant management, which imposes an inclusion mandate where 50% of individuals organized in a village have to be women. As results come in, it will show whether having more women in these village-level bodies changes the composition of projects that are funded and to an overall better allocation of resources. Limited access to basic social and legal services presents another barrier that prevents women from social and economic empowerment. This is even more pronounced in post-conflict settings, or places affected by fragility and violence. To address this shortcoming, in Colombia, DIME is testing the effects of a national government program that serves as a front door for victims to access justice services under the Mobile Victims Unit (MVU). Similarly, women are severely underrepresented in the political life of their societies. As such, they remain deprived from being in a position of power to address the barriers that they face in an institutional context. But the supply-side constraints are not the only ones that they face. Women's political participation, first and foremost through voting in elections, is hindered by their low participation rates. This could be associated with lack of information pertaining to their political rights and limited access to political information in general. In this

context, a growing number of DIME IEs investigate the effects of information-provision interventions and a variety of delivery mechanisms designed to remove or circumvent these constraints. In Liberia, DIME investigates the effects of providing rural women with access to a United Nations Radio's elections-related programs on their political attitudes and voting behaviors. The results point to positive significant effects of the intervention on women's political participation both on a national and a local level. In Zimbabwe, DIME looked at the effects of an intervention designed to reform village-level governance via horizontal pressure on gender inclusion and empowerment.

6.8 Edutainment

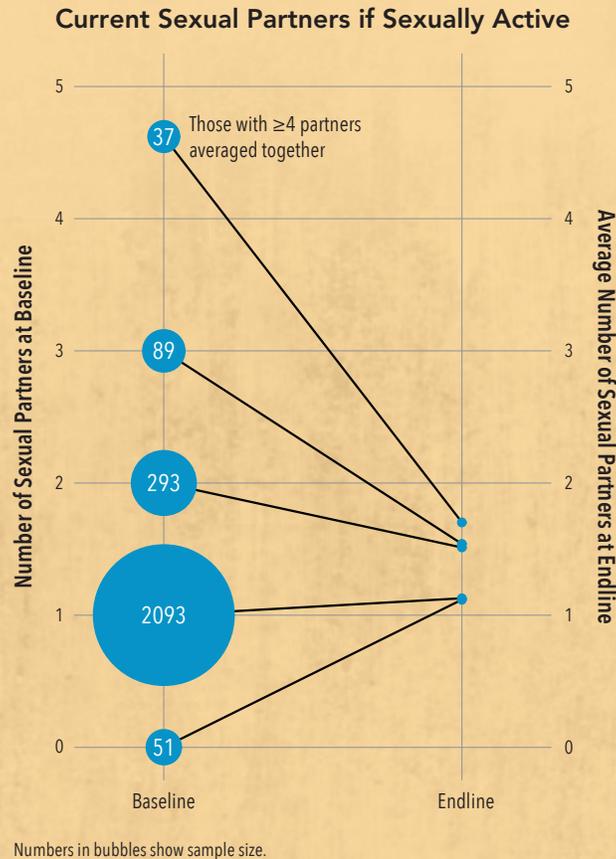
Every year, the World Bank and client governments invest millions in behavior-change campaigns across almost all development sectors. However, many of these campaigns are unconvincing, lack inspiring narratives, and are communicated through outmoded and uninteresting outlets such as billboards and leaflets. Systematic reviews of these campaigns from risky sexual behavior to handwashing consistently show little or no effect on behavior, especially in the long term.

There is an unprecedented opportunity to use entertainment media to change the lives of billions of people, especially in urban areas. Entertainment education or edutainment can be a game-changer for development. Unlike traditional behavior-change campaigns that convey abstract concepts and can become repetitive quickly, educational narratives are easier to follow and remember than abstract information. Characters in mass media have the power to be

SEX IN THE CITY

DIME research finds that an MTV program reduces number of simultaneous sexual partners in Lagos

MTV Shuga reduced the number of current sexual partners when evaluated at the mean, though the effect size depended on the number of partners reported at baseline. For people who only had one partner, the impact of Shuga is zero. For people who had two and three sexual partners at baseline, the average number decreased by 0.18 and 0.35 partners.



MTV SHUGA: A DRAMATIC EVALUATION

Produced by the MTV Staying Alive Foundation, Shuga is a television drama targeting African youth. Having starred Oscar-winner Lupita Nyong'o in the first two seasons, the show is now broadcast in over 70 countries, with potential audiences over 500 million people worldwide according to MTV.¹¹ The DIME study found that eight months after seeing Shuga, viewers were twice as likely to get tested, reported fewer concurrent sexual partnerships

(as shown above), and reduced gender-based violence. Among female viewers, chlamydia infections were halved. These are substantial impacts, especially in light of the limited effects found in other HIV behavior-change trials. The study design and preliminary results have been discussed in TEDx talks, Bloomberg TV, The Hollywood Reporter, an interview during the IMF-World Bank 2017 Spring meetings and WB blogs, among other media outlets.¹²

The impact evaluation results aided the MTV Staying Alive Foundation in obtaining an additional grant of US\$5 million from the Bill & Melinda Gates Foundation to scale up Shuga in Nigeria. The new seasons will focus on Lagos, Kano, and Kaduna with the theme of family planning and creating demand for contraception. With support of new donors, new seasons are also being launched in India and Egypt, where they will address gender-based violence issues.

¹¹ <https://www.mtvstayingalive.org/blog/2017/07/mtv-shuga-goes-global/>

¹² <http://blogs.worldbank.org/voices/newest-weapon-against-hiv-aids-africa-mtv> — <https://www.youtube.com/watch?v=0mCv5U5LRG4>



role models, inspire audiences to engage in new thinking about “what is possible,” and change the perception of what is “normal” and socially acceptable behavior.

The 2015 and 2016 World Development Reports respectively highlighted the untapped potential of entertainment education and mass media in development practice. However, the evidence base regarding the effectiveness of entertainment media remains thin, especially to advise the scale up of entertainment media as a development tool across different sectors. There is a lot to learn about the best way to maximize the impact and minimize unintended consequences of entertainment media, a powerful tool that is largely untapped for development. DIME is starting to expand this evidence base with ongoing experimental evaluations that explore the relative effectiveness of radio spots versus printed narratives to promote adoption of solar lanterns in rural Senegal; the use of a Nollywood (the Nigerian film industry) movie to promote financial savings, and of the MTV Shuga drama to reduce risky sex and gender-based violence in Nigeria.

A Multi-Sectoral Program

The Entertainment-Education program was launched in May 2016 to explore the use of entertainment-education and, more generally, how mass media behavior-change campaigns can be designed to change perceptions of social norms, achieve adoption, and sustain healthier behaviors.

The multi-sectorial program aims to contribute to a series of Sustainable Development Goals (SDG) s: Ongoing impact evaluation studies explore topics like: the use of a Nollywood movie to promote financial savings among entrepreneurs (SDG 1); the impacts of the MTV Shuga drama on risky sexual behavior and gender-based violence (SDG 3, SDG 5, and SDG 16); the use of social-norms campaigns to encourage families to enroll girls in primary school (SDG 4 and SDG 5); the relative effectiveness of radio spots versus printed narratives to promote adoption of solar lanterns in rural areas (SDG 7); the impacts of including entertainment education in in-school life-skills programs to reduce bullying and to prevent drug and alcohol consumption among young people (SDG 3 and SDG 16).

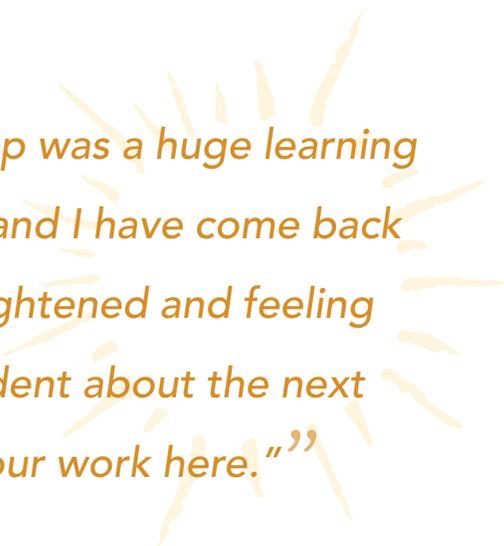
So far, the DIME Narrating Behavior Change program has conducted research workshops and impact evaluations in the entertainment hubs of Brazil, India, Mexico and Nigeria. Its innovative research is supported by different World Bank units, development partners and leading media houses from the “Hollywoods of the world”, including the Asian Center for Entertainment Education, Cinapolis Foundation, Discovery Learning Alliance, ITVS, Life Changing Experiences-Cinemapark, MTV Staying Alive Foundation, Population Foundation of India, and the University of Southern California-Hollywood Health & Society.

The impact evaluations being conducted in the first phase also address important questions related to the indirect or spillover effects of mass media on community members who may have heard program messages from their friends. Another important element being studied is the role of social networks in disseminating and magnifying potential impacts. Finally, the impact evaluations also study how best to reinforce edutainment messages through new interactive technologies, from mobile messaging to social media outlets to videogames.

The program has conducted DIME workshops in the entertainments hubs of Mexico, Nigeria and India.

DIME “Narrating Behavior Change” Workshop (Mexico City, Mexico, May 16-20, 2016)

The official launch of the “Narrating Behavior Change” program took place during a DIME impact-evaluation workshop, jointly conducted with the Inter-American Development Bank. The event brought together 22 project teams from Latin America, Sub-Saharan Africa, and South Asia, and producers and researchers from leading media organizations and universities to



“The workshop was a huge learning curve for me and I have come back to India enlightened and feeling super-confident about the next stage of our work here.””

VINTA NANDA (INDIA), MANAGING DIRECTOR, ASIAN CENTER FOR ENTERTAINMENT EDUCATION AND CEO, THE THIRD EYE

design the next generation of impact evaluations of entertainment media and behavior-change campaigns. The workshop outlined the evidence base and knowledge priorities and through clinics, allowed project teams to work with researchers to develop interventions and evaluation proposals relevant to their projects. Over 90 percent of participants reported being satisfied with the technical content and to have learned what works and what doesn’t to measure the impact of a program.

“DIME Using Entertainment to Change Behavior Forum” (Lagos, Nigeria, May 4, 2017)

As part of the DIME “Beyond the Status Quo: Using Impact Evaluation Research to Drive Innovation and Improve Outcomes in Health” workshop, forum panelists discussed the potential of entertainment education in development and the required public-private partnerships for scaling it up both in the public and private sectors. Nollywood was well represented. Desmond Elliot, a Nollywood celebrity and now legislator from Lagos state, opened the event. Representatives from BBC Media Action, MTV Staying Alive Foundation, and the Nollywood production company

Ultima Limited, discussed the challenges and the opportunities in Nigeria. Otomola Ekehinde, Nollywood super star, philanthropist, and one of Time Magazine's most influential people in 2013, encouraged policy makers to work with the industry to reach large audiences. The Lagos forum helped launch new impact evaluations on education and gender empowerment.

DIME Mass Media Entertainment to Improve Development Outcomes" workshop (New Delhi, India, February 19–20, 2018)

The latest DIME workshop took place in New Delhi, where DIME is launching new evaluations of social media interventions aimed at reducing gender-based violence. Dr. Shравan Kumar, the Joint-Secretary for Culture, opened the event; and Professor Abhijit Banerjee (MIT) delivered the keynote address. The impact evaluation workshop brought together counterparts from government, development partners and edu-tainment producers working in India and in other entertainment hubs. Presentations offered evaluation evidence for different sectors and mediums, as well as innovations on how best to reach vulnerable populations. Bollywood is betting on mass media for positive behavior change. The week before the New Delhi DIME workshop, the Asian Center for Entertainment Education launched the first edition of the Raj Kapoor Awards for Excellence in Entertainment in Mumbai. The brothers Rajiv, Randhir and Rishi Kapoor were among the awardees of the 2018 Raj Kapoor Awards. These and similar events aim to increase awareness about the potential of mass media as a tool for development.

Impact evaluation workshops are key to DIME's efforts to generate demand for evidence-based policy and to create new impact evaluation

studies that meet an expressed need from policy makers. The success of the launch event is reflected in the positive feedback we received from participants:

"Let the magic that we all witnessed last week transform the world with all these projects, ideas, connections, and impact evaluations."

Lorena Guillé-Laris (Mexico),
Director, Cinepolis Foundation

"Thank you for a wonderful gathering of the most interesting leaders in EE and for provoking stimulating discussions in plenary and small teams throughout the week. I hope many new EE collaborations will take place as a result, and I look forward to working with you all!"

Sandra de Castro Buffington (USA), *Director, UCLA Global Media Center.*

Scaling up in Main Entertainment Hubs

Despite being launched in 2016, the Edutainment Program has generated important knowledge in the field of mass-media entertainment. DIME has or will soon have three published papers of edutainment interventions to: i) promote financial literacy and savings among entrepreneurs in Lagos; ii) reduce risky sexual behavior and gender-based violence among youth in Nigeria; and iii) to adopt solar panels in rural Senegal. These rigorous evaluations study not only the effectiveness of edutainment in different sectors, but also the effectiveness of different mass-media outlets (that is, movies, TV series, radio spots, and printed material). Study results have been presented in academic, policy maker, and producer

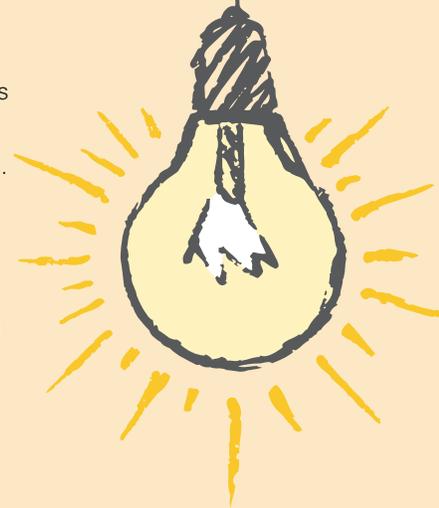
DIME PARTNERSHIP SPOTLIGHT

NEW NORAD-DIME PARTNERSHIP: IMPROVING LITERACY THROUGH MOBILE EDUTAINMENT IN FRAGILE AND CONFLICT STATES

Many children that attend school are not learning basic skills. Many of these students lack proficiency in the language of instruction or do not have access to local language learning resources. Broad multi-stakeholder initiatives like the Global Book Alliance or the EduApp4Syria innovation competition are radically increasing the availability of free, engaging, quality local language learning resources for download by anyone, anywhere. Given the increasing availability of smartphones also among the poorest and the declining prices of these devices, mobile learning has a strong potential to provide supplementary learning opportunities at the community and household levels.

Under the sub-theme of Playful Learning, DIME is partnering with the Norwegian Agency for Development Cooperation (Norad) to carry out a research portfolio of mobile-edutainment innovations, specifically, the use of mobile video-games and digital books in local languages. This new partnership aims to launch two impact evaluations and a series of qualitative evaluations and policy maker workshops over three years, 2018–2020. The research will focus on marginalized groups, including girls, children, and youth affected by crises and conflict, and children and youth with disabilities.

The partnership with Norad is a critical step to continue expanding DIME's research and policy influence in the field of entertainment education. DIME has active discussions with a series of partners, including producers and investors of Hollywood and Bollywood, to carry out similar systematic partnerships for the use of documentaries and movies for international development.



circles. As mentioned above, the results have received media coverage beyond development outlets.

The program has extended to the major entertainment hubs of Nigeria, India, Mexico, and Brazil. These countries have a combined population of 1.87 billion and their entertainment industries produce for their respective regions, thus working there should facilitate translating research evidence into development and industry strategies

for global impact. Going forward, the program will continue to fundraise to expand its portfolio and generate the evidence needed to introduce edutainment into development mainstream. Our focus is on innovations that can potentially promote and sustain behavior change among the largest number of individuals. Thematically, a new window would support research projects in the following sub-themes: Sex in the City, Stopping Violence, Empowering Men and Women, Keeping Clean, and Playful Learning.





INNOVATION IN DATA QUALITY AND MONITORING OF POLICY INFLUENCE

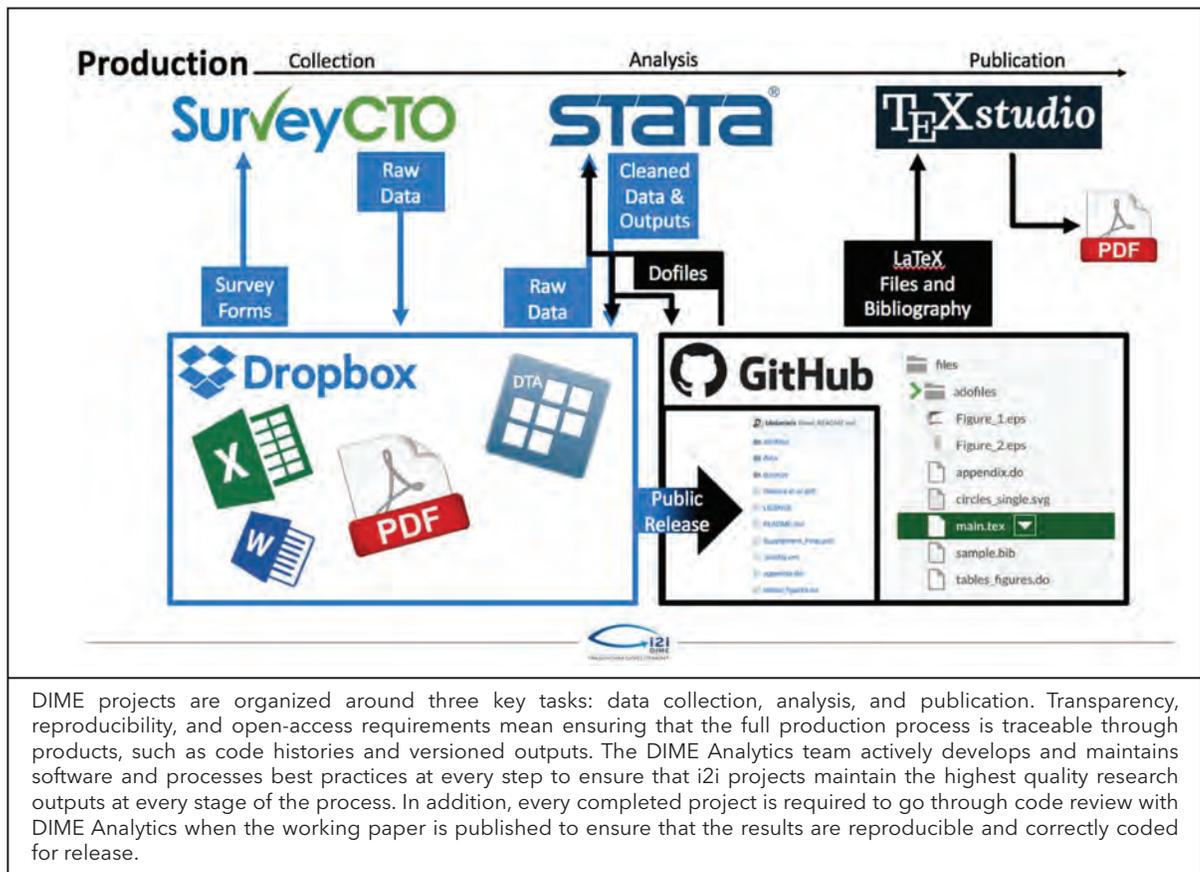
7

7.1 DIME Analytics

The “reproducibility crisis” in social science has finally arrived in economics. In the October 2017 issue of *The Economic Journal*, a special feature, “The Confidence Crisis in Science,” included four papers exploring systemic issues in the replicability of social science experiments like Impact Evaluations. WIRED wrote, “The Dismal Science Remains Dismal, Say Scientists.” Around the same time, a researcher from the London School of Economics released a paper based on reproductions from 32 papers published by the AEA, showing most were underpowered for the statistical methods they had used. Taken together with the seriousness of this issue in other disciplines (e.g., the recent New York Times piece, “When the Revolution Came for Amy Cuddy”), it is critical for DIME to be at the forefront of the field in best practices of data quality and reproducibility.

The DIME Analytics group is leading that effort. The DIME Analytics group works directly with the full team of DIME economists, research assistants, and field coordinators to hold all DIME work to the highest standard of transparency and reproducibility across the complete research production workflow shown in figure 7.1 and, second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities. and, second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities. and, second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities. and, second, to make public training and tools available to the larger community

■ ■ **Figure 7.1** DIME Analytics research production workflow



of development researchers who might not have the same capabilities. and, second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities. and,

second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities. and, second, to make public training and tools available to the larger community of development researchers who might not have the same capabilities.

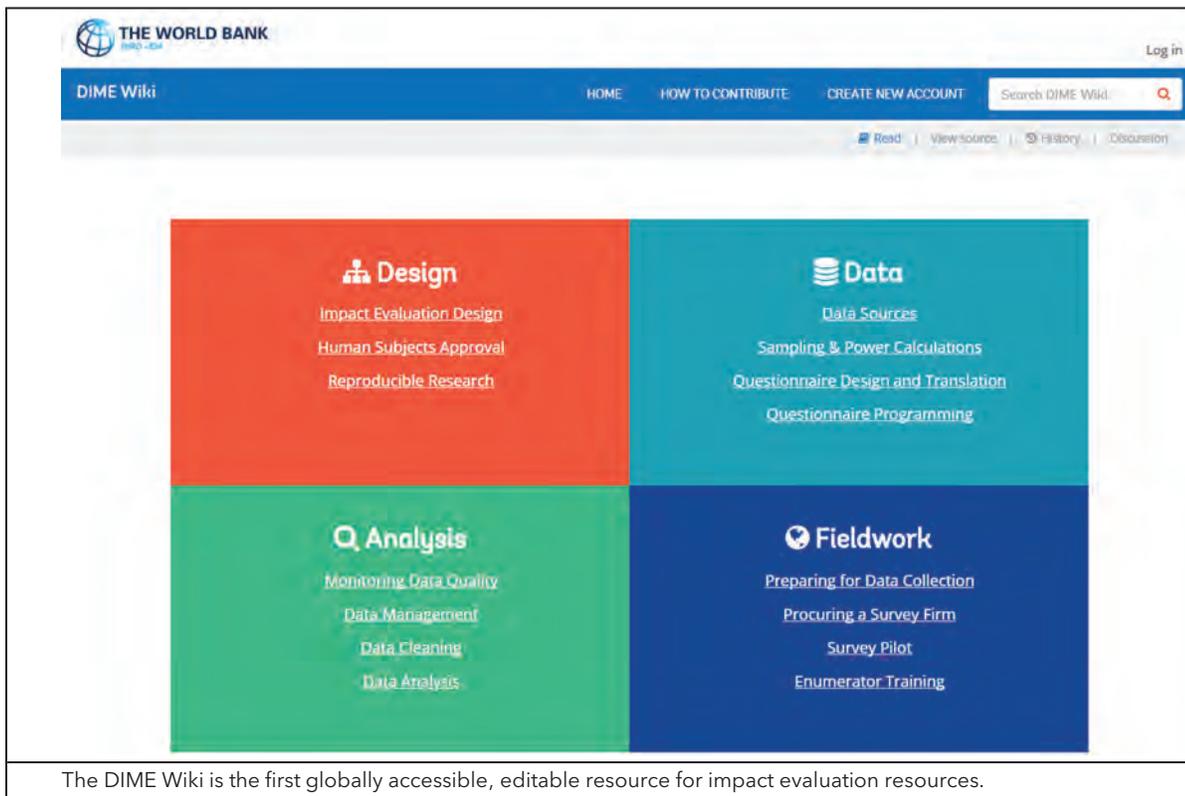
“Thanks so much for this very useful toolkit. It will save time for every single impact analysis. Thanks once again to DIME!”

RESEARCH ANALYST, YAOUNDE, CAMEROON

DIME Analytics’ primary contributions are:

- DIME Wiki: a one-stop shop for research solutions.
- Open-source programming to increase quality of data analysis.
- Structured trainings to enhance global research quality.
- Software and capacity building to improve DIME research reproducibility.

■ ■ **Figure 7.2** Main page of the DIME Wiki



DIME Wiki: A One-stop Shop for Research Solutions

The DIME Wiki¹³ (Figure 7.2) is a one-stop shop for resources on all phases of an impact evaluation: design, fieldwork, data, and analysis. Each article contains a summary of best practices and key resources for successful execution of a particular IE task. All DIME Wiki content is publicly accessible and editable, and the DIME Analytics group has been active in recruiting contributions from other leaders in the field.

While there are many existing impact evaluation resources, none meet the specific need the DIME Wiki aims to fulfill, such as being accessible to the public (non-proprietary); easily

searchable; suitable for users of varying levels of expertise; up-to-date with the latest technological advances in electronic data collection; and curated by a vibrant network of editors who are experts in this field.

As a result of this open structure, the DIME Wiki is continually expanding to house all the tools and best practice guidelines developed by DIME with corresponding links to external resources. Currently, the Wiki hosts 141 articles. There are 58 registered contributors, who have made a total of 4,422 edits. Articles range from introductory material on impact evaluation design, data sources, and reproducible research, to technical discussions of survey protocols; best practices in questionnaire programming; spatial analysis; and use of randomization inference. The Wiki will be broadly promoted to a public audience in early

¹³https://dimewiki.worldbank.org/wiki/Main_Page



2018, at which point we expect the content and user base to grow substantially.

DIME Analytics will continue publicize the DIME Wiki in order to increase the quality of development research globally. It is targeted to all researchers and M&E specialists at the World Bank, clients who are managing data collection efforts in the field, donor institutions, universities, NGOs, and governments.

Open Source Programming to Increase the Quality of Data Analysis

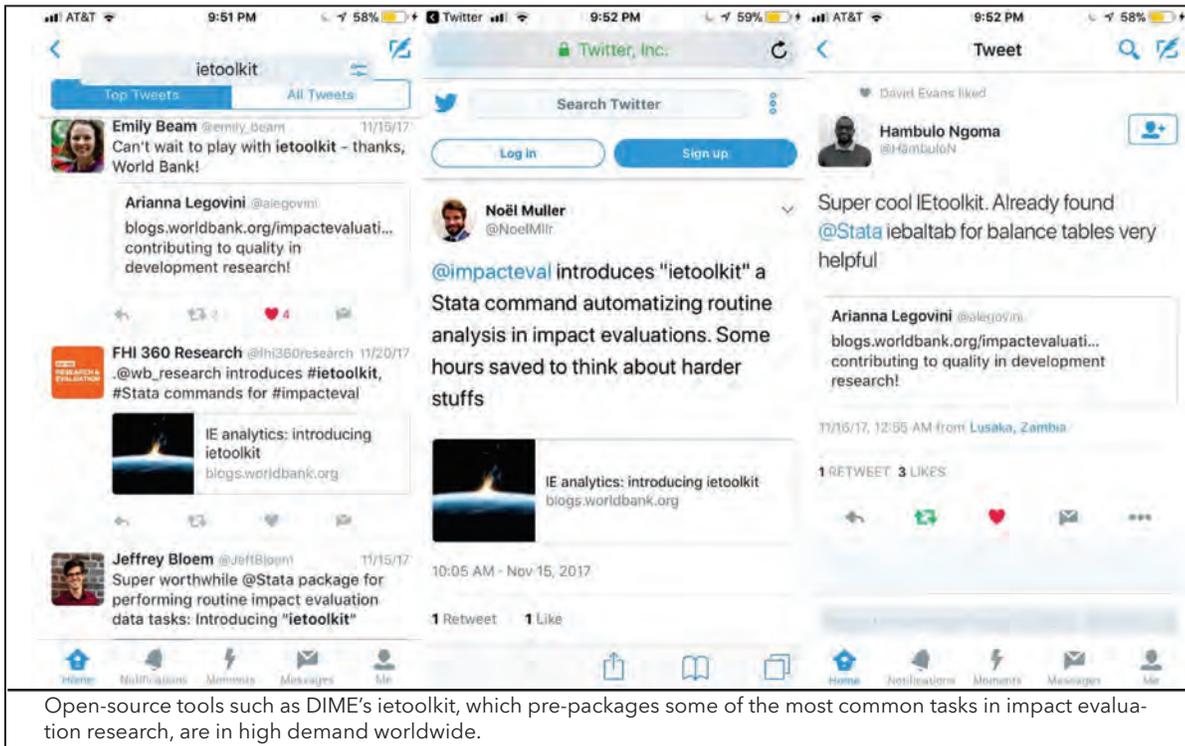
DIME Analytics' first major software release, *ietoolkit*¹⁴, is a Stata package containing commands to routinize common analytical tasks in impact evaluations. This statistical software is a direct result of the Analytics team's effort to gather, document, contribute to and disseminate best practices for data work. It can be installed through SSC, the most popular repository for user-written Stata extensions, and the source code is available for public review and contribution via GitHub.

¹⁴<https://worldbank.github.io/ietoolkit/>

The development of *ietoolkit* produced large economies of scale within DIME's own project portfolio, and its release has been broadly endorsed by the global research community. After the release of *ietoolkit* on the Development Impact blog this fall, many economists took to Twitter to note its merits (figure 7.3). Demand for *ietoolkit* is clear: the software package has been downloaded an average of 615 times per month over the three months since its launch on the Development Impact blog, which is itself one of the ten most-read World Bank blogposts of 2017. It is now ranked 117th in terms of total number of downloads on the SSC archive and is currently the fifth-ranked Stata repository globally on GitHub.

The *ietoolkit* package is an excellent example of the potential for modular programming as a tool to continue to standardize common impact evaluation data tasks. By reusing and building upon this type of prepackaged command, users avoid repeating mistakes others have made before and are common for new research assistants globally. Since the code for these tasks is now standardized in *ietoolkit*, a single easily readable command now accomplishes the work of many lines that later may be difficult to understand; and it is easier for someone else to work with and build on standardized code, facilitating team work.

■ **Figure 7.3** Twitter reception of ietoolkit after public release



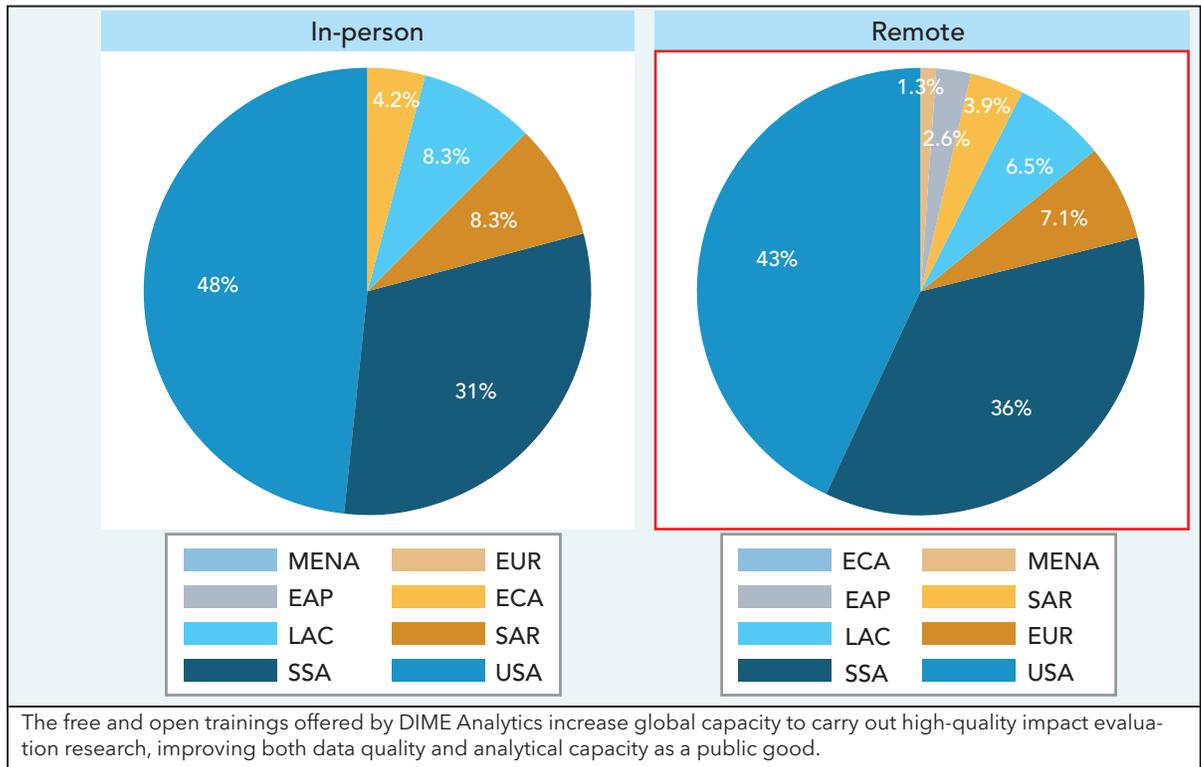
Structured, Open Trainings to Enhance Global Research Quality

The flagship DIME Analytics training is our annual event, “Manage Successful Impact Evaluations.” This week-long, hands-on training is designed to improve the skills and knowledge of impact evaluation implementers, familiarizing them with critical issues in IE implementation, recurring challenges, and cutting-edge technologies. The training is in high demand. For the past two years, we have reached the enrollment cap on in-person participants (50 persons) weeks in advance and have had a few hundred participants join remotely. In 2017, 82% of in-person participants worked on i2i impact evaluations. Of the remote participants, 23% worked at the World Bank, 30% at international or local NGOs (e.g., Plan International, Technoserve, Oxfam), 27% for research institutions (e.g., Oxford Policy

Management), and 13% were from universities (e.g., Universidad Privada Boliviana). As figure 7.4 shows, participants come from diverse backgrounds, ensuring global impact. The training is also highly effective: scores on a before-and-after knowledge test improved by 64.5%.

The course covers impact evaluation tools and concepts, but the primary focus is on how to successfully manage impact evaluations in the field. Participants learn to plan and budget for data collection; design and program electronic survey instruments; monitor data quality and provide real-time feedback to field teams; manage complex survey data and produce descriptive analysis for policy makers; work with dynamic documents to increase transparency and reproducibility of research outputs; leverage qualitative research techniques for quantitative IEs; and build client capacity for impact evaluation. All presentations

■ ■ **Figure 7.4** Work location of training participants, by type of participant



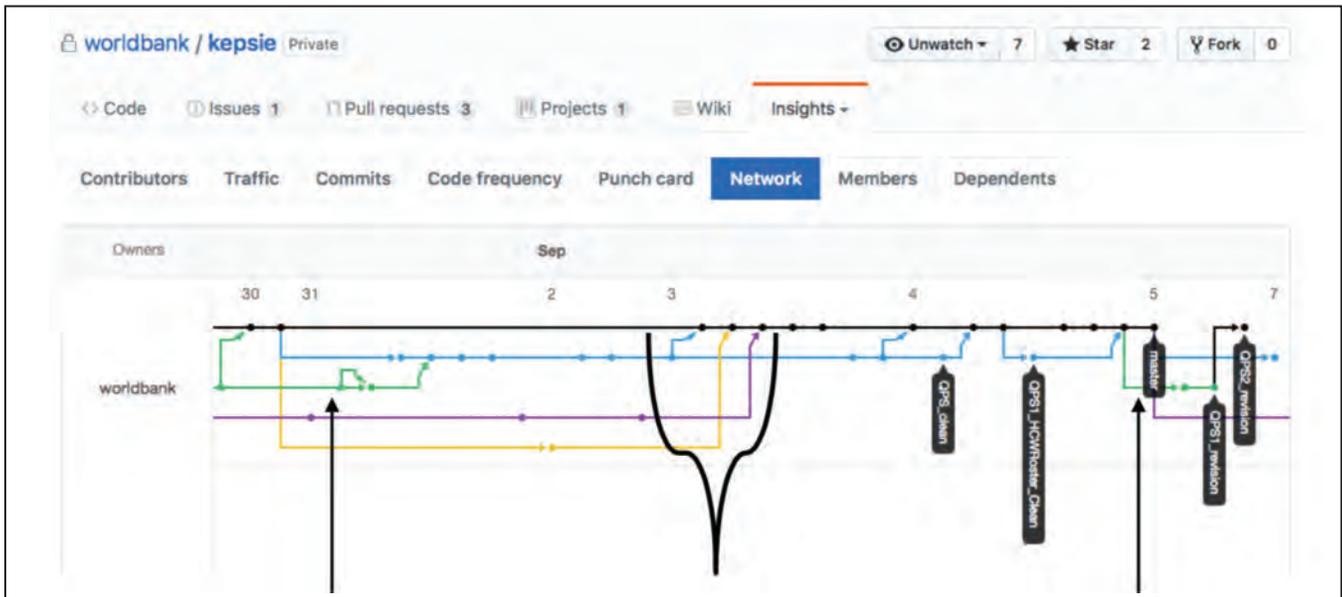
and training materials from the course (including lab exercises and solutions) are publicly available.

Software and Capacity Building to Improve DIME Research Reproducibility

The DIME Analytics team provides direct technical support to all i2i impact evaluations. The team provides routine and miscellaneous econometric and software support, through which the team identifies and shares code and process innovations across the larger DIME group to improve the rate of knowledge sharing across the portfolio. The team also regularly responds to requests to develop and offer personal trainings and broader reading groups within the i2i community for new data tools and econometric methods.

After surveying the DIME economists and RAs to discover their challenges and needs, DIME Analytics created hands-on trainings for using LaTeX and GitHub for impact evaluation research (figure 7.5), which all new-hires attend. DIME Analytics also serves as the primary liaison for feature requests, support, and feedback with the developers of popular tools including SurveyCTO and Overleaf. DIME Analytics also serves as the DIME point of contact for internal inquiries into issues, such as data security and best practices for services such as GitHub, Dropbox, and encryption of personally-identifying information (PII) datasets. In addition, we collaborated with the Berkeley Institute for Transparency in the Social Sciences (BITSS) and the Big Data team at the World Bank to offer a reproducible research workshop open to all Bank staff.

■ ■ **Figure 7.5** Collaboration for reproducibility on WB GitHub



In GitHub, the full version history of code is viewable in a graphical format, as above. Each line traces one of many development changes that happen over the lifetime of the project so that it is easy to track changes and revert to previous versions. GitHub also enables multiple research assistants to work simultaneously on the same or different parts of the code without creating “conflicting copies,” as in some other cloud collaboration software, and GitHub merges these changes together seamlessly. Migrating all i2i projects to the World Bank GitHub creates a useful collaboration environment—preserving these easily-read code histories and unpublished results for complete research transparency, reproducibility, and accountability.

7.2 MyIE Monitoring System

MyIE is a state-of-the-art monitoring system to track the evolution and progress of the portfolio of i2i/DIME supported impact evaluations (IE). It is a user-friendly cost-effective and low-maintenance database-management system and M&E web-based software. It covers two hundred indicators on: profile and status, evaluation design, data collection aspects, monitoring and quality indicators, counterpart details, influence on programs and policies, and produced documentation.

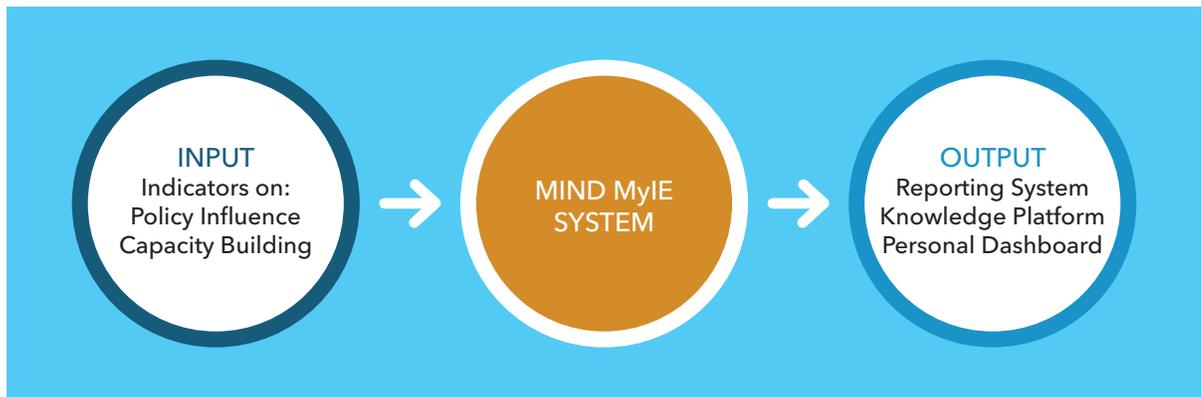
An Innovative Platform

The MyIE Monitoring System facilitates IE portfolio management and the production of reports tailored to different users and clients (DIME, GP

or country management, and donor or client reporting). These reports are automatically generated on any issue of interest (cost, timeline, analytical design, data collection, geography, etc.) across the portfolio or for specific programs or sectors.

The system serves to monitor progress, identify issues, and to highlight how IE products can best contribute to development practice. By collecting data on how IEs feed into project design, support capacity building, influence outside projects and/or motivate scale-up or down of an intervention, the MyIE system helps us identify opportunities to better inform policy decisions and increase our focus on research with impact. The summary statistics were created to help understand the challenges and lessons learned from the DIME/i2i portfolio.

■ ■ **Figure 7.6** Mind MyIE framework



Features and Content

The system asks IE Task Team Leaders (TTLs), or other authorized respondent, to report on their activities during the entire life cycle of their IEs. The data is collected annually for all ongoing IEs and currently translates into 191 i2i and/or DIME IEs, of which 76% are DFID/i2i funded. The latter thus englobes the 146 IEs presented in this annual report. A complete list of all collected indicators can be found in appendix 3. These are divided among 8 sections: IE profile, Evaluation Design, Data Collection, IE Monitoring and Quality Indicators, DIME Involvement, Counterpart Details, IE Influence on Program/Policy, Documentation and Research Outputs.

On top of the IE questionnaire tab (“Manage my Projects” in the system), users have access to a Reports Section. It consists in aggregated graphs and tables for each of the following:

- Map of IE distribution across the world
- IEs across GPs and sub-themes
- IEs across regions
- IEs across IDA countries
- IEs by life cycle
- IEs by duration
- IEs involving gender components or analysis

- IEs in fragile and conflict affected settings
- IEs and main counterparts
- IE budget distribution
- IE secured funding
- IEs across evaluation method
- IEs across number of treatment arms
- IE data collection rounds and response rates
- IEs that have ethical clearance and/or study registry
- IEs that influence project design or implementation
- IEs that generate evidence used to support project adoption, scale-up, scale-down, continuation, or cancelation decisions.
- IEs that have contributed to improving the coverage, quality, delivery, output/outcomes, or cost of the program or other intervention outside the IE
- Number of previously-ongoing impact evaluations completed and reported (for example, as working papers or policy briefs published online)

In 2017, the system was further developed to include additional functionalities serving not as a data collection tool, but as a real resource platform. MyIE users now have access to their own personal dashboard, which will reflect the same

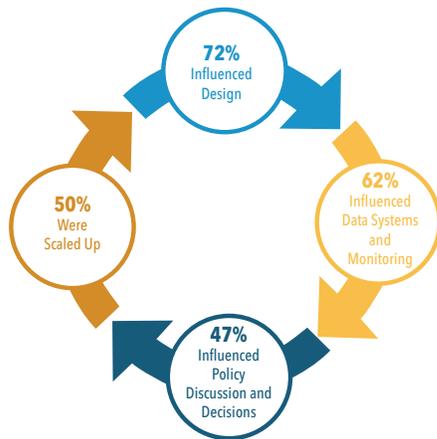
■ ■ **Figure 7.7** MyIE home dashboard

Last Updated: January 2018



IE Policy and Program Influence

Percentage of IEs in the relevant phase

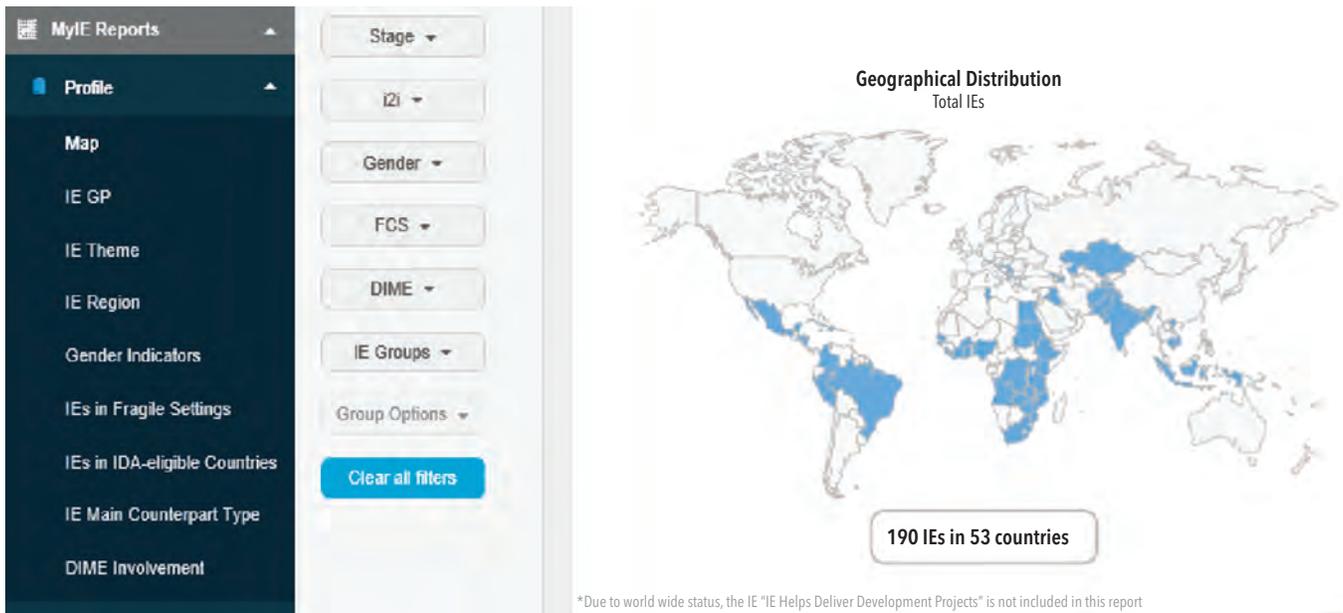


IEs across WB Networks

Percentage of IEs by network



■ ■ **Figure 7.8** Sub-tab of MyIE reports



indicators that are in the Reports Section, but on a portfolio level. The disaggregated data will be presented relative to the overall average statistics, hence giving users indication as to where they stand in relations to other portfolios. Additionally, a Databank and Document Library have been made available. This allows users to directly download data and IE outputs from the system, conditional on if the latter is public. Finally, in the long term, the system will have a public access feature, where unregistered users will be able to obtain aggregated descriptive statistics and access to documentation for completed evaluations. The final objective is for MyIE to serve as a recording system for all World Bank-related impact evaluations.

7.3 Policy Influence Results and Verification

Measuring IE Policy Influence Across Multiple Stakeholders: 2017 Policy Influence Survey

i2i and DIME's goal is to generate high-quality data and evidence to motivate policy change via a close collaboration with policy implementers throughout the policy cycle. Understanding when, where, and how policy is influenced is a priority for our program. Entry points for influencing decisions are during the design of the project, after baseline, while testing implementation alternatives, and after IE completion, when replication or scale decisions are taken. To develop a good understanding of the channels of IE influence on policy and their extent, however, two challenges must be overcome: (i) defining what to measure and how, and (ii) identifying the link between data/evidence and decision-making, as attribution is complicated by the absence of a proper counterfactual.

Within this context, we measure policy influence through: (i) a yearly survey of IE TTLs, who report on instances during the IE cycle when, in their own understanding, the IE activities and outputs have led to policy action; and (ii) a survey of WB and government and implementing clients to collect their understanding of when and how IEs have influenced their decisions. We measure policy influence through the indicators developed for the MyIE monitoring system. These indicators are conceptually closely linked to the chain of inputs to outputs in our theory of change to affect policy, where securing policy relevance of the IE questions and the investments on client ownership and capacity are critical as inputs (see MyIE Monitoring System presented in chapter 4 and chapter 7.2). The novelty of this exercise stems from surveying external counterparts to complement the measures of policy influence that DIME developed in recent years, methods which have been used to collect records from IE TTLs for all i2i/DIME portfolio. While the approach cannot identify causal IE influence, validating information from different stakeholders clarifies the extent of common and divergent perceptions, and brings us closer to an understanding of what the IE portfolio is likely to have achieved in this area.

In this chapter, we present the results from 44 IE teams surveyed, including 106 respondents in 22 countries and across 33 institutions implementing IEs with support of i2i/DIME. Overall, the results suggest DIME/i2i's IE work is considerably influencing the teams and projects across different stages and dimensions of the policy process. In particular, 94% of external respondents report that their IEs have added value to their programs and units; and they identify multiple channels through which the activities and outputs of the IEs have contributed, for instance, to improve the program design and delivery, and to build their capacity. Additionally, there is

SURVEY TECHNICAL DETAILS IN A NUTSHELL¹⁵

Data Instruments: DIME's Policy Influence Survey consists of two independent data collection exercises: (i) a survey of government and agency implementers and WB operations TTLs (henceforth referred to as "external"); and (ii) the latest module on policy influence collected through the MyIE Monitoring System, which is updated by IE TTLs every year on the complete DIME/i2i portfolio. Both surveys were conducted between the end of 2017 and beginning of 2018.

Sample: The DIME portfolio at the time of the survey design included 163 IEs (including 121 i2i IEs). After exclusion criteria were applied, 109 IEs were eligible for the survey; we focused on the 67 IEs that have a DIME IETTL. We excluded (i) IEs that were too early in the IE cycle to have any policy outputs; (ii) retrospective IEs or post-program evaluations, which by their nature do not involve work with counterparts during the implementation; (iii) IEs that are on hold due to security conditions

or other external factors; and (iv) older IEs that do not have a reachable counterpart.

Response rate: High response rates were secured, with IE TTLs reporting the highest (100%), followed by government counterparts (80%), and WB operation TTLs (61%). At the IE level, 44 of the 67 IEs targeted were surveyed, for a response rate of 66%.



demand for increasing activities associated to these policy outputs. The overwhelming majority of external respondents provided ideas on how to increase the benefits of the IE activities, including more dissemination, capacity building, and further efforts to improve the IE research enterprise. The results provide support to the theory of input-to-outputs for policy change that is at the core of i2i and DIME. They also provide insights on potential avenues to improve on that policy influence.

To What Extent Do IEs Influence the Programs and Policies Evaluated?

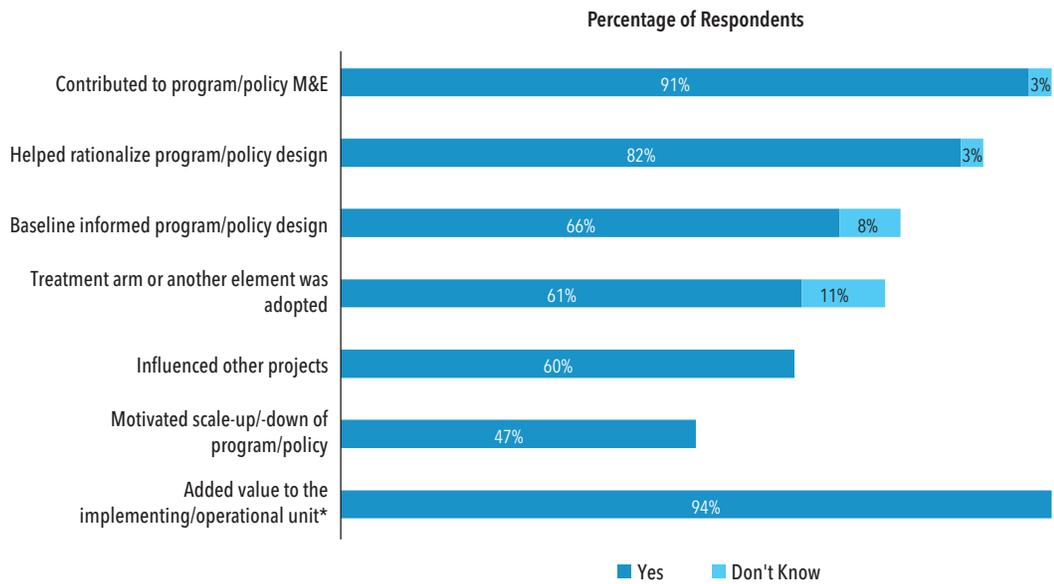
A majority of respondents report that the IE has influenced their policies and programs across the indicators assessed, and the measure of influence in these areas is relatively high. As figure 7.9 shows, the average influence by indicator ranges from 47% to 91%. The greatest

influence reported concerns the contribution of IEs to the program/policy M&E function (91%), followed by helping rationalize the design of the program/policies (82%), informing design through baseline results (66%), the adoption of a treatment arm or another IE element (61%), influencing other projects (60%), and motivating scale-up/-down of program/policy (47%). As we move from the beginning of the IE cycle, influence declines as expected as IEs are at different stages of implementation. These proportions are likely to increase as the portfolio matures (64% of the sampled IEs are after CN or baseline and 36% are completed). Finally, 94% of government/implementing agencies and WB operations TTLs report that the IE has added value to their units.

In addition, most IEs have had influence on multiple dimensions, which to some extent is an indication of high intensity. As figure 7.10 shows, practically all IEs have displayed influence in at least three indicators (out of the six assessed), 90% of IEs in at least four indicators, and 56% of IEs in

¹⁵ A detailed technical note on the survey methods is available in the Policy Survey Final Report upon request.

■ ■ **Figure 7.9** IE policy influence and added value, all respondents



Note: only asked to WB operation TTLs and government/implementing agencies.

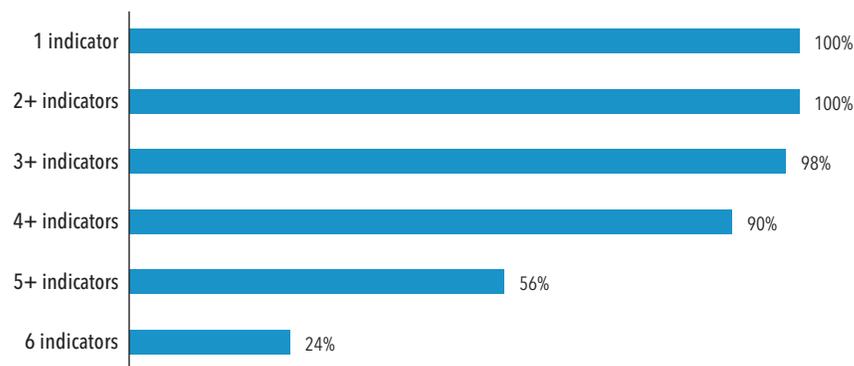
at least five indicators. Finally, almost a quarter of the IEs report influence in all six areas, including motivating scale-up, which mostly happens after the IE has been completed.

Is Policy Influence Perceived Similarly Among Key IE Actors?

Somewhat expectedly, influence is perceived heterogeneously across different types of

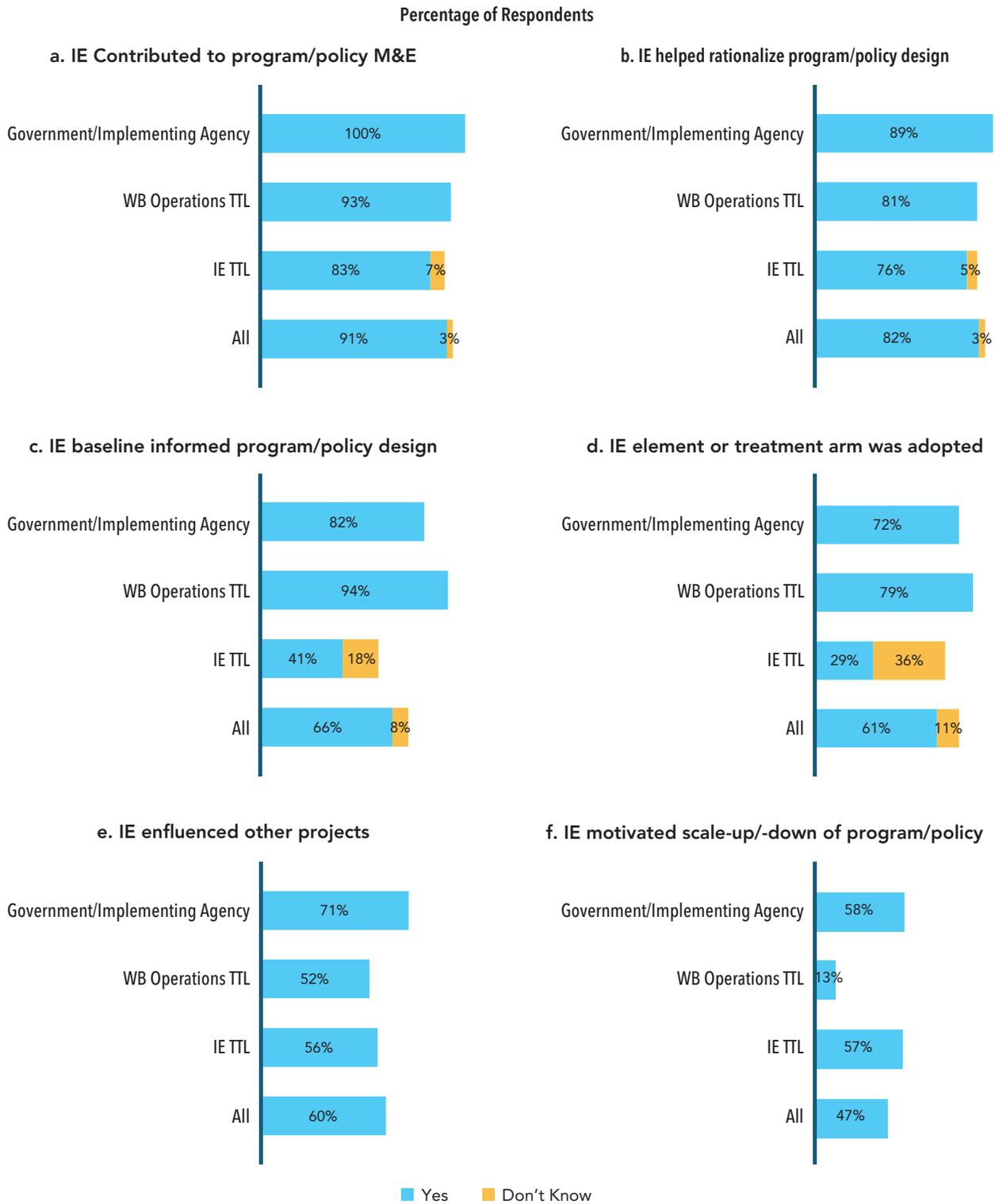
respondents, which may be partially explained by how close they are to the policy decisions. Overall, government and implementing agencies report the highest policy influence across most indicators, and IE TTLs report the lowest policy influence and the highest proportion of “don’t know” answers. The highest congruence is on the IE contribution to the policy/program M&E function (figure 7.11a) and to rationalizing program/

■ ■ **Figure 7.10** IE policy influence intensity, all respondents



Note: This figure presents the number of selected policy influence indicator at IE level. It aggregates responses from the 3 types of respondents: government/implementing agency, WB operation TTL and IE TTL.

■ ■ **Figure 7.11** Policy survey indicators by type of respondent



Note: These figures present results to the questions in appendix 3. Some questions are filtered according to their status in the IE life cycle.

policy design (figure 7.11b), which are the activities where all actors work most closely together. For instance, in terms of the M&E function, government and implementing agencies report that in 100% of the cases the IE contributed to improving this function, followed by 93% of WB Operations TTLs and 83% of IE TTLs. An important discrepancy is on whether the baseline informed policy/program design (figure 7.11c), with the WB operations TTL reporting the highest influence (94%), followed by the government/implementing agency (82%), and the IE TTL (41%), with 18% “don’t know” answers in the last group. Finally, on whether the IE motivated scale-up/-down of program/policy (figure 7.11f), the government/implementing agency group and the IE TTLs report relatively high influence (58% and 57%, respectively), while the WB operations TTLs report low influence (13%). This might be due to the fact that WB project TTLs might not be involved with a country or client after the project is brought to completion and might not be party to those discussions. WB operations TTLs for completed IEs were more difficult to reach for the survey, as well.



To What Extent Do Different Actors Agree on the Specifics of the Policy Influence?

The previous figures seem to indicate that IE TTLs tend to underreport policy influence across all indicators. However, given that they are the most likely respondents (their response rate is 100% in our sample), an important question is to what extent are they able to report policy influence details that their counterparts would interpret similarly. Therefore, an important part of the external policy survey focuses on a simple validation (*Yes, No, or I don’t know*) of the details that IE TTLs had provided in the internal survey.

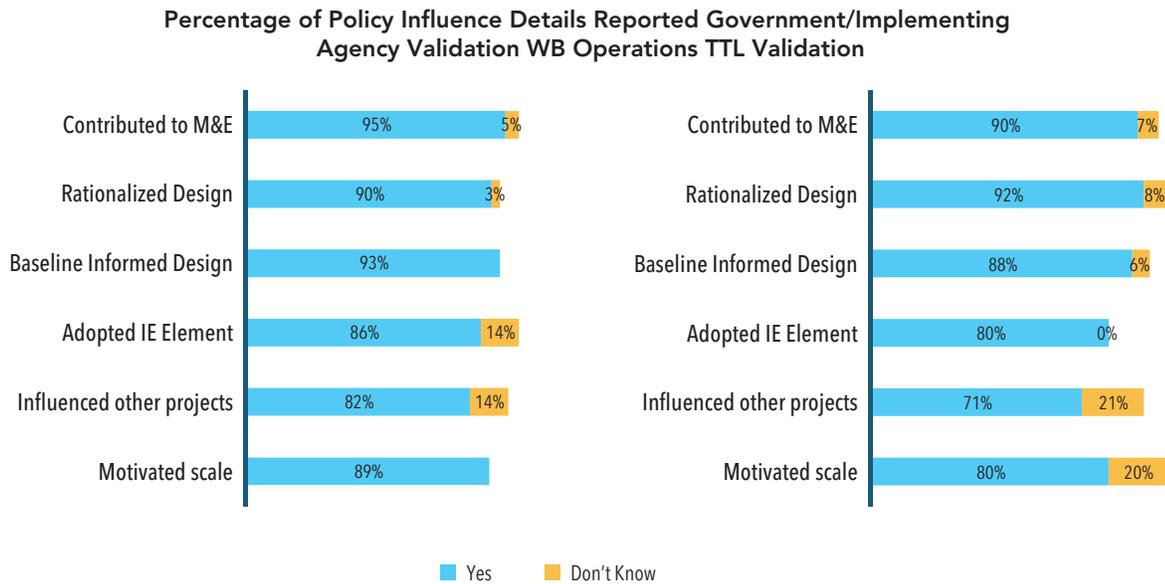
Below are a few examples to illustrate the details validated by the government/implementing agencies and the WB operations TTLs:

IE baseline informed program/policy design (IE TTL, Rwanda Irrigation Project): “The IE baseline identified that less than 1% of farmers currently sell their harvest through the cooperative. After the baseline, they [the project team] worked with the IE team to develop a new strategy for fee collection.”

IE motivated scale-up/-down of program/policy (IE TTL, Burkina Faso Local Government Support Project): “Elements from the IEs, including the annual municipal performance survey and the municipal scorecard intervention, are being scaled-up under the Bank-assisted Burkina eGovernment project and PACT [French acronym for Local Government Support Project] Additional Financing.”

IE helped rationalize program/policy design (IE TTL, Workfare to Well-being: A randomized Evaluation of the Effects of Labor-Intensive Public Works on the Urban Poor in Eastern DRC): “The IE team helped rethink the monetary

■ ■ **Figure 7.12** Validation of IE TTL reporting



Note: These figures present results to the question "DO YOU AGREE with the following reports on . . . ?" The question applies to all IEs that have IE TTL reports on the indicators presented in appendix 3.

incentive component of the intervention, which resulted in savings being extracted at the source (at payment) and commitment to not take it out until project completion. We also rationalized the PIU's [Project Implementing Unit] traditional approach to training by implementing a soft skills component."

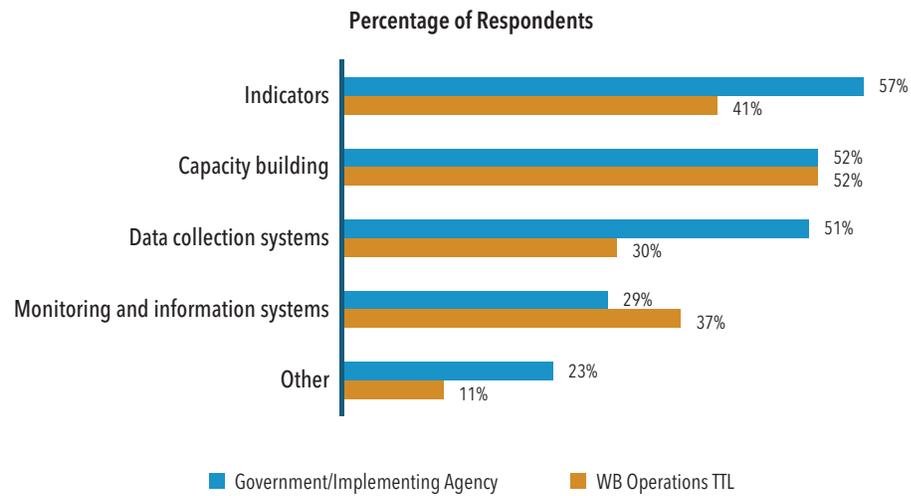
As figure 7.12 shows, across all indicators, the congruency is high, ranging between 82% and 95% in the government/implementing agency validation, and between 71% and 92% for the WB operations TTL validation. These results exclude the "don't know" answers, which for some indicators reach 20%. Overall, these numbers validate the internal data collected through the monitoring system. They also emphasize that policy influence may be perceived differently even in cases where specific details are shared, which calls for more work on dissemination and communication among the teams.

Channels Through Which IE Policy Influence Works

Given the high potential for IEs to influence policies and programs and the almost unanimous assessment from the external counterparts that IEs add value to their units, questions remain regarding (i) whether we can learn something about the channels through which this influence is taking place, and (ii) whether we can do even more to improve the extent and quality of this influence. Next, we analyze these channels for each of the six indicators assessed.

Contributing to the M&E function: 91% of the respondents reported that their IE has contributed to the monitoring and evaluation function of their project teams or units. As figure 7.13 shows, the most important channels identified by government/implementing agencies include: contributions to improve the indicators to monitor the project, including new indicators or

■ ■ **Figure 7.13** IE Contributed to program/policy M&E (Channels)



Note: These figures present results to the question “How did the impact evaluation contribute to the Monitoring and Evaluation function of your project team or unit? (Select all that apply).” The question applies to IEs in all phases.

SELECTED QUOTES FOR IE CONTRIBUTED TO PROGRAM/POLICY M&E

“Now we have more information. Hence, before making a decision we look back into the impact evaluation results . . . with the live maps and market surveys we can track in which markets there are tomatoes shortage and intervene on time if needed.”

IE Client, Rwanda Rural Feeder Roads

“The use of tablets and electronic surveys was adopted by the project as a product of the IE activities, trainings, and capacity building.”

IE Client, Moving Youth Out of the Market for Crime: Interventions in the Honduras Safer Municipalities Project

“Data previously collected on paper is now in digital format and of higher quality.”

IE Client, Do Matching Grants Create Agricultural Productivity: IE of PACT-Nepal

“The IE has shown that if routine data is collected correctly, we can do better at a lower cost. This helped to better organize the health information system as well as improve and audit data completeness and quality.”

IE Client, Stormwater Management and Climate Change IE Adaptation

“The impact evaluation research team supported MISFA [Microfinance Investment Support Facility for Afghanistan] in developing a framework to monitor the implementation of the project. MISFA is using this framework in all of its Targeting the Ultra Poor projects, also outside of the study areas.”

IE Client, Targeting the Ultra-Poor in Afghanistan

“Everything is now tracked based on the work that they [the IE team] did. When we say, we are paying subsidies, we are paying subsidies based on calculations they did.”

IE Client, DIME-WaSSIP Nairobi

“The quality of our data had a great improvement.”

IE Client, Smallholders Agricultural Productivity and Commercialization (SAPEC)

“There is daily continuous monitoring which did not exist before. Previously, there was no follow-up with [health] facilities after inspections.”

IE Client, Kenya Patient Safety Impact Evaluation

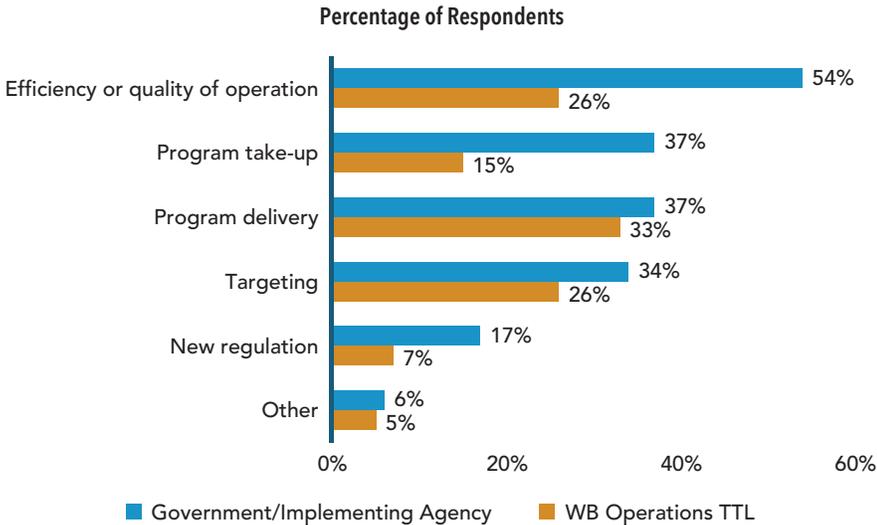
improvements on existing ones (57%); capacity building on M&E tools, for instance through training (52%); and support to improve data collection systems, for instance, through the establishment of electronic collection of project data (51%). WB operations teams reported similar importance for these channels but capacity building ranked first (52%), and support in developing monitoring and information systems ranked third (37%), for instance, by contributing to improving or developing new administrative data. Overall, these results highlight the importance of the IE work in helping improve measures and systems to monitor project progress, as well as in improving the capacity of their teams in the monitoring and evaluation function beyond the IE.

Rationalizing Program/Policy Design: 82% of respondents report that the IE helped them reorganize or rethink the program being evaluated (89% of government/implementing agencies). The channels identified include different dimensions on how the program is delivered,

its efficiency, quality, and outreach. As figure 7.14 shows, for the government/implementing agencies, the most important channels through which IEs contributed to this area relate to improving the efficiency/quality of the operation (54%), improving the take-up or the participation of the target population (37%), improving how the program is delivered (37%), and helping improve targeting or how to select or reach the target population (34%). Finally, 17% of respondents report that the IE contributed to the introduction of new regulations, which are higher-level policy influence. For the WB operations TTLs, the top three channels are improvements to program delivery (33%), efficiency or quality of the operation (26%), and targeting (26%).

Informing Program/Policy Design Through Baseline Data Results: 82% of government/implementing agencies and 94% of WB operations TTLs report that the IE baseline results informed their programs/policies. In many countries, this is often the first available data on the

■ ■ **Figure 7.14** IE helped rationalize program/policy design (channels)



Note: These figures present results to the question, "How did the impact evaluation lead to reorganizing or rethinking the design of the program that is being evaluated? (Select all that apply)." The question applies to IEs in all phases.

SELECTED QUOTES FOR IE HELPED RATIONALIZE PROGRAM/POLICY DESIGN

"[The IE] helped assess that firms with 5 to 20 employees should not be targeted [by] the program."

IE Client, Impact of Management Upgrading on Small- and Medium-Sized Companies in Brazil (Banco do Nordeste)

"Within the IE activities, a census was carried that allowed to understand better the target population characteristics, needs and interest, and improve their targeting."

IE Client, Moving Youth Out of the Market for Crime: Interventions in the Honduras Safer Municipalities Project

"It has improved how they're actually conducting the inspections and how they are introducing themselves. It is more professional, it is more approachable, but most importantly, it is more standard. Making something standard and predictable is a way of making it better. So, it is now uniform and takes away some discretion, which of course is linked with corruption and other problems, so it's just a better way of doing the inspections."

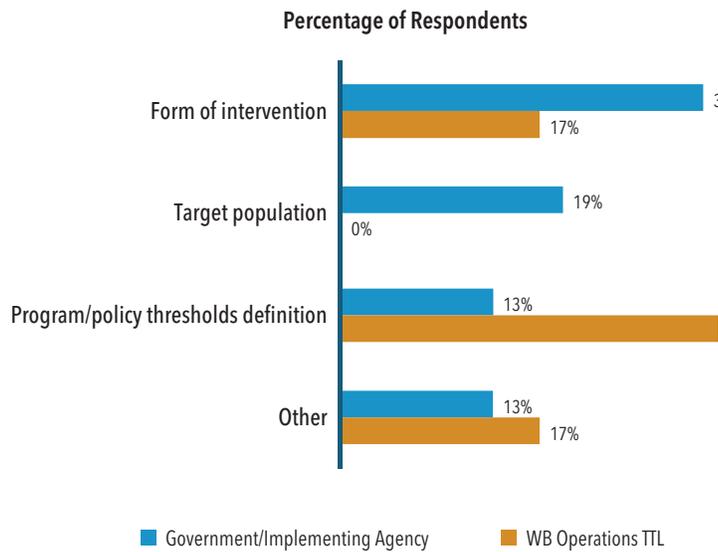
IE Client, Kenya Patient Safety Impact Evaluation



sector being evaluated, which therefore brings forth the opportunity to understand multiple key characteristics of and constraints faced by the population that the governments' aim to target. Over half of the sampled IEs are currently in the "after baseline data collection" stage and hence going through the analysis and dissemination phase of these results. As figure 7.15 presents, government/implementing agencies report that the baseline data contributed to modifying the actual form of the intervention (31%), redefining the target population (19%), and helping define intervention thresholds such as the distributed grant value (13%). For WB operations TTLs, the most important channel is through the definition of those same thresholds (33%), followed by modifying the form of the intervention (17%). It is worth highlighting that an important proportion of respondents reported that the results were not shared (19% of government/agencies and 25% of WB operations TTLs). This may be, in part, related to delays in the analysis, but it would be worth exploring further.

Adopting IE Treatment Arm or Another IE Element: Regarding completed IEs, we ask if any of

■ ■ **Figure 7.15** IE baseline informed program/policy design (channels)



Note: These figures present results to the question, “How did you use the impact evaluation’s first big survey data collection results? (Select all that apply).” The question applies to IEs that have had a baseline.

the treatment arms or another element of the IE were adopted, i.e., continued to be delivered to the target population after the program is finished (this does not necessarily entail a scale-up). Seventy-two percent of government/implementing agencies and 79% of WB operations TTLs reported that some element of the IE was adopted. As figure 7.16 shows, 26% of respondents mentioned that an IE treatment arm was adopted post-IE results, which is close

to the 29% reported by the IE TTLs. The question is framed differently in the MyIE monitoring system, where we only ask IE TTLs about the specific treatment arms. As such, the IE TTL and counterpart responses can only be compared on the blue “adopted arm” bars. In addition, 31% of WB operations TTLs and 7% of government/implementing agencies reported that other elements were adopted, including in particular communication and outreach materials.

SELECTED QUOTES FOR IE BASELINE INFORMED PROGRAM/POLICY DESIGN

The baseline data was used to define the supply and demand incentive thresholds.”

IE Client, Senegal RBF Impact Evaluations

“Baseline results helped better understand which areas most needed infrastructure upgrades. This helped target locations with most acute needs.”

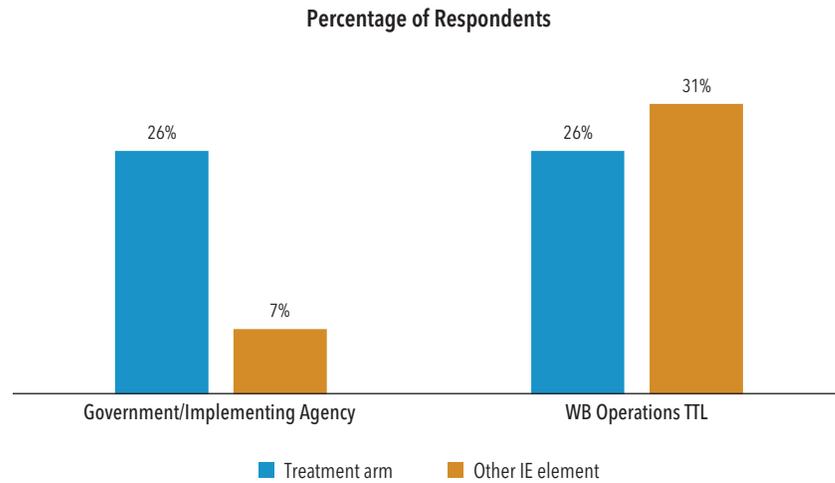
IE Client, Impact Evaluation of Community Infrastructure Upgrading

“The [IE] results are expected to inform adjustments and potential operational/fare interventions that can help optimize the impacts of the BRT on travel by potential public transport riders (and particularly poorer riders and women) as well as on urban development changes. Very good, innovative, and collaborative work by the coordinated DIME/GP team.”

IE Client, Impact Evaluation of the Dar es Salaam BRT System



■ ■ **Figure 7.16** IE element or treatment arm was adopted

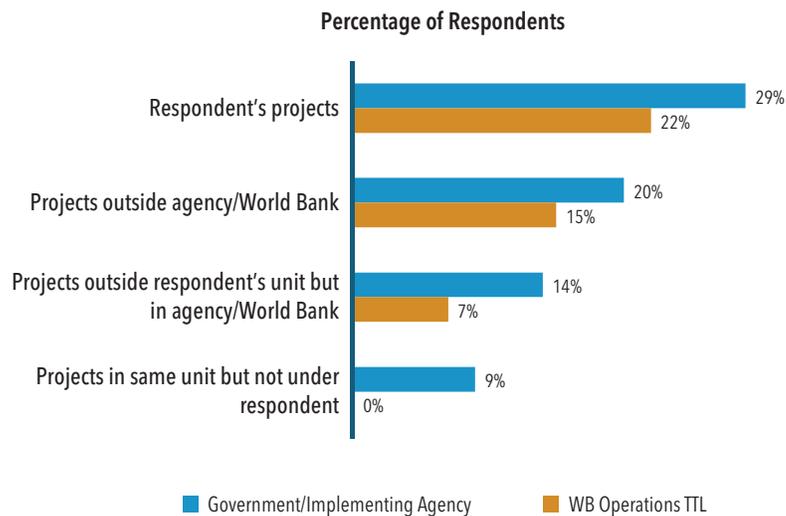


Note: These figures present results to the question, "Can you tell me if the agency/institution continued any of the following interventions after the impact evaluation was completed? (Select all that apply)." The question applies to IEs that are completed.

Influencing Other Projects: This indicator aims to record whether other projects have been influenced by the IE. We are only asking about any projects that the respondent may be aware of and thus we would not consider this an exhaustive list of influence: 71% of government/implementing agencies and 52% of WB operations TTLs reported that the IE influenced

other projects. The average number of other projects influenced per IE surveyed is 1:1 for government/implementing agency versus 0.5 for WB operations TTL. Conditional on "yes" answers, the average of other projects influenced is 1.9 and 1.4, respectively. Figure 7.17 presents results by project type and suggests that IEs are affecting mostly respondent's

■ ■ **Figure 7.17** IE influenced other projects (channels)



Note: These figures present results to the question, "To the best of your knowledge, can you tell me if the impact evaluation influenced other projects, policies, or even impact evaluations inside or outside your unit and/or World Bank? This can be locally, at national level, or even outside the country (Select all that apply)." The question applies to all IEs.

SELECTED QUOTES FOR IE INFLUENCED OTHER PROJECTS

“CONAFOR [The National Forestry Commission for its acronym in Spanish] is now working on the design of the Emission Reduction Initiative in which we want to test an integrated territorial management model (several programs contribute to solving a problem within a polygon). We designed this program with an impact evaluation in mind (treatment, control, baseline, annual data collection, end line 6 years after the baseline). A pilot is being implemented in 5 states.”

IE Client, Mexico Payment for Environmental Services Scheme: A Retrospective Evaluation

“IE data influenced policy at local government level, with local governments setting aside money in their budget to Target programs for 1000 days’ children and mothers with nutrition sensitive interventions.”

IE Client, Sunaula Hazar Din Community Action for Nutrition

“The impact evaluation has inspired the client, MISFA, to conduct its own impact evaluations of the TUP program in other provinces not covered by the study.”

IE Client, Targeting the Ultra-Poor in Afghanistan



projects (29% for the government/agencies versus 22% for the WB operations TTLs), but respondents also report influence on projects outside of the agencies and the World Bank (20% reported by government/agency versus 15% for the WB Operations TTLs).

Motiving Scale-up/-down: 47% of respondents report that the population or geographic coverage of the intervention changed for IEs that are at the final of their cycle. When we consider this policy influence indicator, we tend to focus on how the IE results can shape decision-making and lead governments to extend the program at a wider level or, on the other hand, to stop interventions that had proven not to work at the implementation or impact level. However, in practice, scale-related decisions take place at several stages, including while project implementation is ongoing. Given the multiple layers of and stages for scale-related decisions, we approached the data collection on this question from different angles. We analyze the indicator for those IEs for which results are available (i.e., in last stage of their life cycle), but also try to capture other instances where scale decisions have taken place, even if the IE results are not available yet.

In terms of scale-related decision influenced by the final results of the IEs assessed, the respondents report decisions across multiple sectors and regions. While in Brazil, the impact of the high-intensive financial literacy intervention motivated the expansion of the program to all national public high schools. In Nigeria, the IE results discouraged the government to scale up a one-off/short-term movie intervention due to the lack of impact of their Financial Literacy program. In Senegal, motivated by the IE results, the government replaced HIV community sensitization campaigns with peer-led campaigns, and in Rwanda the LWH Rural Finance IE motivated the promotion of savings products that considerably improve farmers’ production and incomes.

During the open-ended part of the external survey, we collected scale-related decisions for IEs during the implementation and found frequent reports of decisions taken during the implementation of the project. An example is the Kenya Patient Safety Impact Evaluation, where despite not having final results yet, the monitoring data and the work done as part of the IE have already motivated the government to make the decision to scale up to the national level. Many respondents also mentioned

SELECTED QUOTES FOR IE MOTIVED SCALE-UP/DOWN

“Based on the results of the IE, the government decided to scale up commitment savings account products to other project areas. The commitment features were shown to help farmers improve their investments and production beyond a simple mental accounting and financial planning training. The government team subsequently received support from Access to Finance Rwanda, a DFID-supported NGO, to train and support other SACCOs in the country on the new savings products”

IE Client, LWH Rural Finance Evaluation, Rwanda

“Elements from the IEs, including the annual municipal performance survey and the municipal scorecard intervention, are being scaled up under the Bank-assisted Burkina eGovernment project and PACT Additional Financing. This is part of the project restructuring.”

IE Client, Burkina Faso Local Government Support Project

“A follow-on the project supporting LGAs [Local Government Authorities, the lowest level of government in Tanzania] was developed after the CUIP [the Community Infrastructure Upgrading Programme that the IE evaluated]. The IE was a driving factor in securing DMDP [Dar es Salaam Metropolitan Development Project] to follow-on to CIUP.”

IE Client, Impact Evaluation of Community Infrastructure Upgrading, Tanzania

“The most efficient study arm was scaled up. The Senegalese government replaced HIV community sensitization campaigns with peer-led campaigns after the IE. The implementing NGO intensively trained peer leaders for these to train in turn other community members (versus inviting everyone in the community to attending community sessions). The scale-up was national.”

IE Client, Long Term Effects of HIV Community Sensitization Campaigns, Senegal

“CONEF [the National Committee for Financial Education for its acronym in Portuguese] decided to scale up the [financial literacy] program to public high schools in Brazil.”

IE Client, Brazil Financial Literacy

“The IE results suggest not to scale up a one-off/short-term movie intervention as impact seems to persist only in the short term.”

IE Client, Impact Assessment of Financial Literacy in Nigeria

“The project is being scaled up at the national level. We are going to train inspectors in the remaining 44 counties and they will be doing inspections on a daily basis, just like KePSIE [Kenya Patient Safety Impact Evaluation]. And we will adopt almost everything from KePSIE to improve both the quality and legality of services.”

IE Client, Kenya Patient Safety Impact Evaluation

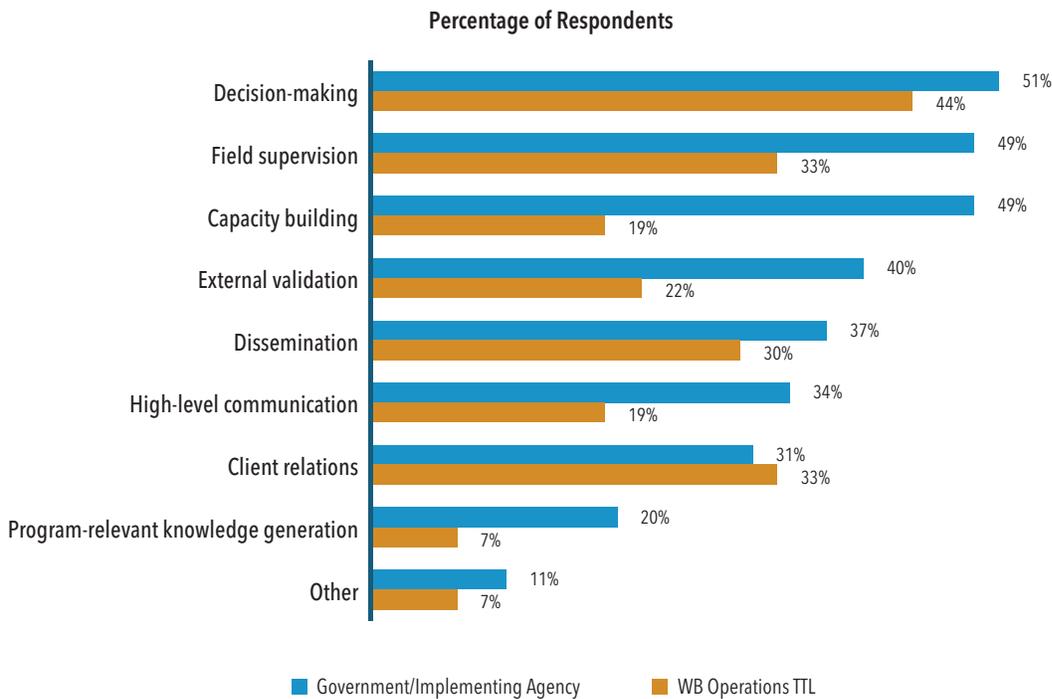
that program scale-up is currently under consideration thanks to the IE results, but that due to budgetary constraints nothing has yet been implemented. Such cases were not counted in the Policy Survey results but are worth mentioning here. For example, in the case of the IE of Payments for Environmental Services in Mexico, preliminary results are helping the implementing agency make a business case for the preservation and possible scale up of the PES program amidst significant cuts in the government budget. In the case of the Speed of Justice IE in Senegal, the IE team developed a computerized tool to track and monitor justice cases (“pop-up”)

that the Ministry of Justice is now hoping to use to scale up to all chambers of the Dakar Court. The dissemination of the IE results is supporting the fundraising efforts.

Adding Value and Improving Further the Potential of IEs to Influence Policy: Perceptions and Ideas from External Counterparts

Consistent with the previous results, 100% of the government/agency respondents and 85% of the WB operations TTLs reported that the IE has added value to their agencies or units. In

■ ■ **Figure 7.18** Ways in which IE added value (channels)



Note: These figures present results to the question, "How did the impact evaluation create any added value to your unit? (Select all that apply)." The question applies to all IEs.

addition, when asked how likely they would be to accept to participate in another impact evaluation, the average was 88% for the government/implementing agencies and 81% for the WB operations TTLs.

A critical question is how all these results interact and map into the perceptions of the main IE counterparts on what creates value from their IEs. The external counterparts are closer to the policy decision-making process and are also critical actors to improve evidence-based policy-making. Therefore, understanding how they perceive the added value of their IEs, and how they could be more helpful becomes central for us. As figure 7.18 shows, government/implementing agency respondents report that their IE helped create value for their programs and units through contributions to better decision-making (51%), better field supervision (49%), building their capacity (49%), providing external validation for

their intervention (40%), better dissemination (37%), and high-level communication (34%). The WB operations TTLs reported that the IE created value mostly through their contribution to better decision-making (44%), field supervision (33%), client relations (33%), and dissemination (30%). Altogether, these results align closely with i2i/DIME's goal to generate high-quality data and evidence to motivate policy change.

Always with the objective to improve, we asked the respondents whether and how they thought their impact evaluation could be more helpful: 97% responded affirmatively to whether their IE could be more helpful, but as figure 7.19 shows, the "how" varies considerably by type of respondent. The top-five responses cited by reporting government and implementing agencies (who also answered in much more detail) are: (i) improving in-country dissemination (43%), (ii) providing more trainings (37%), (iii) making

SELECTED QUOTES FOR WAYS IN WHICH IE ADDED VALUE

“By working in communities as part of the evaluation, the project team was better able to understand needs and challenges in communities.”

IE Client, International Interventions to Build Social Capital: Evidence From a Field Experiment in Sudan

“Having the DIME presence in the field and the DIME TTL providing technical advice was very useful. The DIME presence in the field allowed for more monitoring.”

IE Client, The Angola Local Development Project

“DIME has helped the Ministry to understand better the use of strong evidence in decision-making”

IE Client, Rural Finance Evaluation, Rwanda

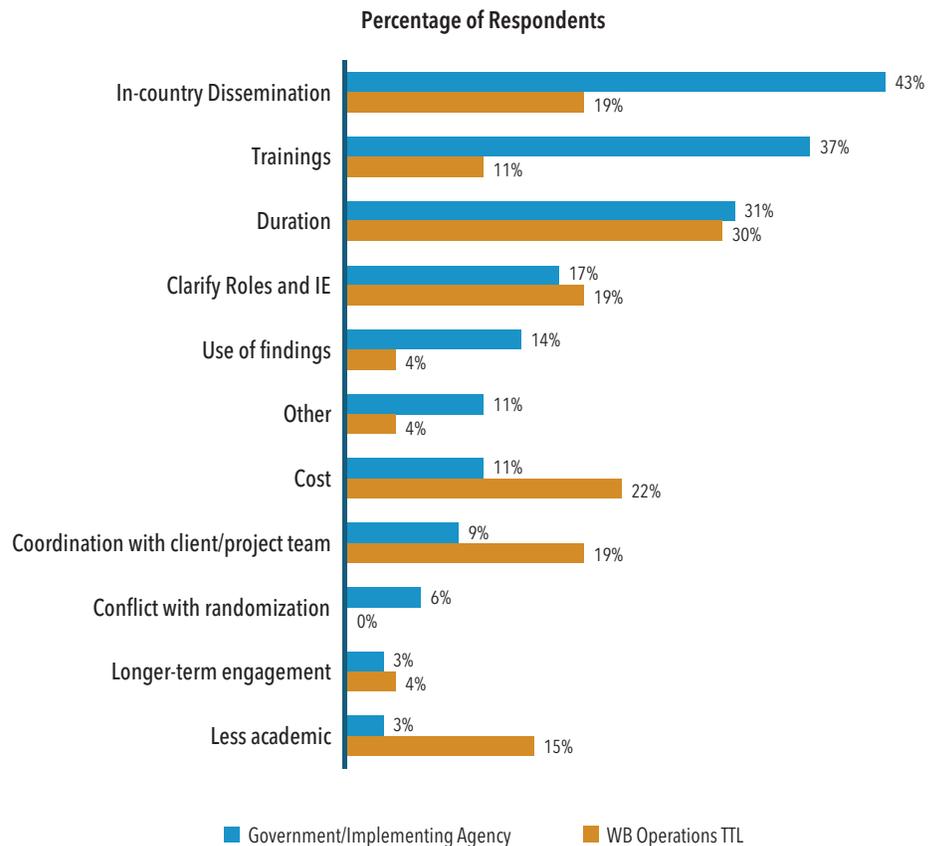
“Because of the results Brazil became a reference of Financial Literacy in OECD and in LA and Caribe”

IE Client, Brazil’s Financial Literacy

“KePSIE helped show us that the system can work. In one sense, joint inspections are a good idea and there’s no question about that. In theory, it should work. But, in practice, can they actually work based on the policy environment?”

IE Client, Kenya Patient Safety Impact Evaluation

■ ■ **Figure 7.19** Ways in which IE can be more helpful



Note: These figures present results to the question, “How do you think the impact evaluation could be more helpful? (Select all that apply).” The choices for this question were unprompted. The question applies to all IEs.

efforts to reduce the duration of the IE (31%), (iv) clarifying the roles of the teams and the IE (17%), and (v) better using the IE findings (14%). The top-5 ranking ways for WB operations TTLs include reducing the duration of the IE (30%),

and four more areas that tied in second place with 19% of responses: reducing the cost of the IE, in-country dissemination, coordination with client/project, and clarifying the roles of the teams and the IE.

■ ■ **Table 7.1** Police influence indicators assessed

Indicator	Question	Definition
Contributed to program/policy M&E	How did the impact evaluation contribute to improving the Monitoring and Evaluation function of your project team or unit?	This question aims to understand how the IE team may have contributed to the Monitoring and Evaluation function of the program or unit. A few examples/options are: <ul style="list-style-type: none"> ■ Better indicators to monitor the project ■ New data systems (with existing or new administrative data) ■ New data collection system (developing electronic tools to collect and/or monitor data) ■ Capacity-building on M&E tools (training)
Helped rationalize program/policy design	How did the impact evaluation lead to reorganizing or rethinking the design of the program that is being evaluated?	This question aims to understand if the impact evaluation helped rethink the project or program. For operational respondents, this may only be true if the IE has implications on project documents (e.g., Project Approval Document), however here are also capturing any de facto changes or improvements. A few examples are: <ul style="list-style-type: none"> ■ Improved efficiency/quality of the operation ■ Thinking on how to deliver the program better ■ Targeting (how to select/reach the target population) ■ Take-up (participation of the target population) ■ New policy-level regulation
Baseline informed program/policy design	How did you use the impact evaluation baseline results?	This question applies to impact evaluations that have had a baseline. We are capturing any potential changes to the program that resulted from the results of the baseline. A few examples are: <ul style="list-style-type: none"> ■ Redefined target population ■ Redefined intervention threshold or parameters (e.g., grant/kind value to distribute) ■ Modified actual form of the intervention
Treatment arm or another element was adopted	Can you tell me if the government adopted any of the following components of the impact evaluation?	This question is asked for completed impact evaluations. The respondent has the option to choose between the different treatment arms preloaded from MyIE, with an additional option to specify any other element of the IE or developed by the IE that may have been adopted. This question is not to be confused with a scale-up of the program. This not about the actual program but a specific treatment arm or elements of the intervention that were adopted post-IE.
Influenced other projects	To the best of your knowledge, can you tell me if the impact evaluation influenced other projects, policies, or even impact evaluations inside or outside your unit/agency and World Bank/Ministry?	This question measures if the impact evaluation design or results influenced other projects. The respondent may not be aware of all the projects that are influenced by the IE but we aim to identify as many as possible. We include any other projects within the agency or WB unit, to the Ministry or World Bank Group and even outside these institutions or government organisms. When possible, we gather information about these other projects or programs.
Motivated scale-up/-down of program/policy	Was the population or geographic coverage of the intervention changed after the impact evaluation began?	This question is asked regardless of the IE phase and aims to capture if any intervention was scaled-up or scaled-down in terms of geographic coverage and/or population size.

Note: All questions have the options, "Other" and "IE did not influence . . ."



CASES OF POLICY INFLUENCE



Stages of Policy Influence

i2i uses rigorous evidence to motivate policy change in the world's poorest countries. At a minimum, this requires connecting lessons from our evaluations to new policy decisions. However, DIME's evaluation model embeds learning into each element of the project cycle: from defining policy, deciding whether to continue a program or not, to the structure of the next phase of learning. Given its continuous nature, the model provides real-time learning through regular feedback loops. Figure 8.1 illustrates the four main stages where IEs influence policy and the proportion of IEs for a sample in our portfolio in each category, based on reports from 2017 DIME Policy Survey including government and implementing agencies; the WB operational team leaders; and the 2017 DIME MyIE Survey, including IE team leaders. The results suggest important influence across all stages: 79% of the IEs in the sample have informed policy design during the initial stages of the project preparation, 52% of IEs have guided policy discussions and decisions, and 61% have informed scale-up/scale-down decisions.¹⁶

Informing Design

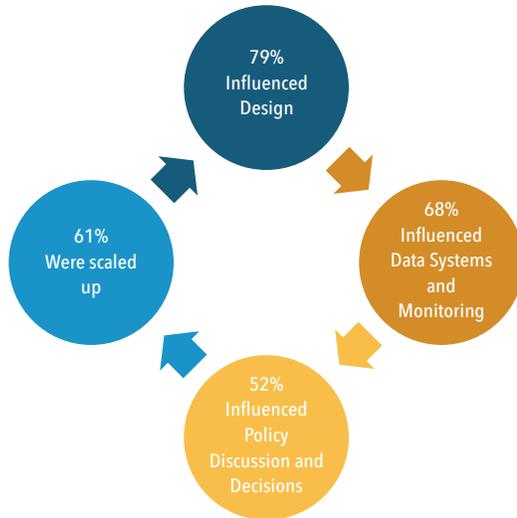
The majority of projects evaluated in i2i participate in IE workshops, where researchers, program officers, and policy makers work together to think through the project's rationale and identify alternative delivery approaches to test hypotheses about what may help the project most effectively deliver on its objectives. For instance, while designing an evaluation of a US\$470 million sewerage investment in Nairobi's slums,

¹⁶The percentages are calculated using the total number of IEs in the portfolio that are at a stage where the policy influence indicator is applicable for the respective subsample. The figure is based on the 2017 DIME MyIE Survey.



■ ■ **Figure 8.1** Policy influence through the IE cycle

Subsample of IEs from the 2017 DIME Policy Influence and MyIE Surveys



the team identified last mile connection costs as a critical barrier to achieving impact: if households were unable or unwilling to connect to the bulk infrastructure, there would be little chance of the program generating any meaningful impacts. As such, the team worked together to experimentally test different subsidy levels for the sewerage connection costs to help the project identify optimal subsidy policies to ensure the program’s success.

Guiding Mid-course Corrections

Early IE results are often able to influence projects while there is still time to adapt the program mid-course. For example, in Côte d’Ivoire, an IE of a public works program found large and significant impacts on earnings for women in the short term, while impacts for men are smaller. Short-term impacts on vulnerable individuals were also found to be stronger than for less vulnerable individuals. This led the program to adjust its geographical targeting while the program was still being implemented to focus on peri-urban areas, as well as raise its quota for women from 30% to 50%. This decision led to an increase in program impacts on women’s 6-month earnings by US\$288.

Informing Scale-up and Scale-down Decisions

To date, evidence generated through i2i has already been used to inform scale-up, scale-down, and adaptations to large development programs influencing the lives of millions of people. Illustrative examples include:



Supporting scale-up. An IE of the MTV drama *Shuga in Nigeria* found that watching this entertainment education drama, which has a potential audience of over 500 million individuals,¹⁷ reduced risky sex. The treatment group was twice as likely to get tested for HIV and chlamydia infections declined by over 50 percent for women. The study helped MTV secure funding to go beyond Sub-Saharan Africa. In 2020, the *Shuga* drama will start seasons in Egypt and India, focusing on gender-based violence (instead of HIV). **In Kenya**, more than 4.5 million people now have access to health facilities with better patient safety standards for an improved regulatory function and inspections system in three pilot counties where the Kenya Patient Safety Impact Evaluation (KePSIE) is taking place. The early successes of the pilot has led the government to scale-up the program nationally starting next year, potentially reaching all of Kenya's 46 million population.

Informing scale-down. In Haiti, an impact evaluation of input subsidies provided to 30,000 farmers found that recipients used the funds as substitutes rather than complements to their own investments in agricultural inputs. The following year, anticipating additional subsidies, farmers again reduced their investments in inputs, in contrast to the program's objective of increasing input use and generating higher yields. Based on these results, the government has scaled back its investments in input subsidies for annual crops and is focusing more on conservation agriculture and agroforestry while they invest in further research to help redefine the annual crop subsidy packages.

¹⁷<https://www.mtvstayingalive.org/blog/2017/07/mtv-shuga-goes-global/>



POLICY INFLUENCE CASES: SUPPORTING LOCAL GOVERNANCE IN BURKINA FASO

DIME pursues deep engagements with countries to embed multiple evaluations within the same country (and at times, within the same project) to build comprehensive evidence on the multiplicative and complementary constraints impeding development in a sector of that country. DIME's engagement with the government of Burkina Faso to improve local and municipal government is an example of this approach within the governance program.

Since 2013, DIME has partnered with the government of Burkina Faso, most centrally with the Ministry of Territorial Administration and Decentralization (MATD) through its Local Government Support Program (PACT), to develop and test innovative solutions to improve the performance and accountability of municipal governments. This work has evolved into a multi-year research collaboration called REGLAB (*Recherche Expérimentale sur la*

Gouvernance Locale au Burkina Faso, or Experimental Research on Local Governance in Burkina Faso). The program involves four large-scale randomized controlled trials (RCTs) looking at: (i) including "citizen observers" in municipal council meetings; (ii) providing local government with municipal scorecards; (iii) providing community-based organizations (CBOs) with a financial stake in municipal performance to incentivize lobbying for better municipal services; and (iv) providing citizens with benchmarked performance measures of their municipalities relative to others.

The design and rollout of these RCTs has influenced projects and policy across all of the core dimensions of the program life cycle.

Influencing project design:

Of the four RCTs described above, three concern the Citizen Observer, Municipal Performance Scorecards, and Incentives for

CBOs interventions, and were jointly conceived with PACT and other Burkinabé stakeholders in 2013, and were integrated into the project's citizen engagement component, which seeks to improve accountability linkages between local level policy makers and citizens. The inclusion of these activities led to the contracting by PACT, of a staff member specifically responsible for the *approches novatrices* (or "innovative approaches") and, overall, greater scrutiny of the way in which the project activities were implemented by PACT staff and members of the IE team.

Intervention development was informed by available evidence and co-produced. For example, the design of the performance scorecards involved year-long consultations with mayors, municipal service providers, line ministries, and other stakeholders. The team identified a set of indicators that are readily accepted as being within the municipal governments' primary responsibilities. The annual performance survey maximizes the integration of existing data sources and complements them strategically. The dissemination of scorecards is under way; and spot-checks found print materials (scorecard posters) available in conspicuous locations at the municipal offices. Beyond the scorecard experiment, the annual performance survey has proven relevant to a larger set of government, civil society, and research stakeholders.



Influencing scale up/adaptation of the program:

The project extension for a new wave of implementation is now being informed by work during the first project phase.

Specific additions to the World Bank's Project Appraisal Document (PAD) for the next phase of the project based on the IE results include:

- **Subcomponent 3.1:** Strengthening local council and community oversight of local government performance. PACT will no longer directly supervise regional NGOs responsible for implementing Component 3 activities, as this arrangement proved difficult to manage and administratively burdensome. Instead, PACT will work with the Association of Municipalities of Burkina Faso, which will coordinate and supervise the work of NGOs.

- **Subcomponent 3.2:** Recognizing and rewarding good local governance practices through competitive sub-grants. The Municipal Performance Scorecard will be used as the basis for COPEGOL, an annual performance competition for municipalities. The scorecard forms the basis for the redesign of this competition with a view to making the evaluation process more independent and transparent, and increasing incentives for municipalities with lower baseline performance to participate (in the previous iteration of the competition, such municipalities were effectively excluded).

- **Subcomponent 3.3 (new subcomponent):** Municipal performance tracking. This subcomponent gives PACT an explicit role in the transfer of the SUPERMUN municipal performance monitoring system (see below) to the Bank-assisted eBurkina Project and in the ongoing oversight and development of the system.

- **Subcomponent 3.4 (new subcomponent):** Partnerships with community-based organizations. The second phase of the Scorecards for Performance experiment will be implemented under this sub-component, where local CBOs will be financially incentivized to work with and lobby their municipal administrations for improvements in municipal performance. This RCT will be implemented in 349 communes, spanning all of Burkina Faso's thirteen regions.

Affecting decisions beyond the focus projects:

A key activity of the engagement in Burkina Faso was the development of an improved data system through the DIME-led development of an annual municipal performance survey, SUPERMUN (for *Suivi de la Performance Municipale*). The survey has been rolled out in 140 municipalities since 2013 and will be expanded nationwide under the Bank-financed **eBurkina project**. The survey provides annually updated information on basic service delivery in health, education, water and sanitation, administrative

services, and on municipalities' institutional capacity. Each indicator is benchmarked by national norms.

The adoption of SUPERMUN by the eBurkina Project illustrates how a "technology" (in this case, the precursor to the SUPERMUN municipal performance monitoring system, developed as part of a learning collaboration with a specific project) can influence policy and program design beyond that project and promote intra-Bank and intra-government collaboration. PACT is implemented by the Ministry of Territorial Administration and Decentralization (MATD) and supervised by the Governance GP, while eBurkina is managed by the National Agency for the Promotion of ICT (ANPTIC) and supervised by the Transport & Digital Development GP.

SUPERMUN was developed for the Municipal Performance Scorecards intervention and because of its congruence with eBurkina's objectives, was adopted by the latter in order to establish a "nationwide, annual municipal performance tracking system that provides reliable, timely open data on the performance and institutional capacity of local government" (eBurkina PAD). ANPTIC and MATD will now jointly oversee and develop the system, while responsibility and capacity for its implementation and maintenance are transferred to ANPTIC. A launch workshop for the collaboration was held with MATD, ANPTIC, and other stakeholders in March 2017.



POLICY INFLUENCE CASES: LABOR INTENSIVE PUBLIC WORK (LIPW) PROGRAMS

Although a number of in-depth country cases exist, the primary framework through which DIME selects impact evaluations is a sectoral one. A sectoral approach means designing complementary interventions across contexts to build a robust portfolio of results to inform strategies in a given sector, such as labor intensive public works (LIPW) programs, to reduce violence. By combining multiple experiments and results across countries and contexts, DIME is building recommendations that can be transported across contexts.

LIPW programs, a popular social safety net model across developing country contexts, provide temporary employment opportunities and earnings to vulnerable households and individuals through infrastructure projects, with the aim to mitigate income shocks/smooth consumption while building community assets. These programs are one of the most common safety net interventions currently implemented, especially in FCV contexts. Despite their ubiquity and substantial budget allocation, rigorous evaluations on the efficacy and effectiveness of these programs are limited. As such, DIME, with the support of i2i, has undertaken a multi-country impact evaluation program of LIPW projects to fill the knowledge gap.

Informing better decision-making

In Egypt, the government has been implementing a LIPW program that targets semi-skilled and unskilled unemployed workers and aims to improve their labor market and economic outcomes, as well socio-psychological ones. The program has two components: i) community infrastructure, which was implemented by creating jobs through the private sector, and ii) community social services, which were implemented through NGOs. Both components embedded an IE undertaken by DIME. Although the first component was found to have no impact on labor market outcomes, the government of Egypt was planning to scale up the LIPW program in a US\$400 million national program only on

the private sector component. When the IE results were communicated to the government, it decided to hold off on scale-up plans and to reconsider their decision to drop the social services component. The government is now waiting for longer-term results from the impact evaluation to inform the final scaled up project design.

In Tunisia, an i2i funded DIME IE evaluated the impact of a Labor Intensive Public Works program for long-term unemployed individuals. The positive IE results informed the scale up of the program; the intervention was extended (almost quadrupled) to a larger target group in the same geographic area of the country.

LIPW Program Impact Highlights

In Egypt, the community infrastructure component of the program found no impacts on labor market outcomes, while the social service component, which provided grants to community organizations **for employing youth, was found to increase employment by 15 percentage points (pp) and incomes by 35%.**

In Tunisia, the IE found positive impacts on outcomes including wage employment, which increased by **8 pp, and incomes, which increased by 26%.**

In Côte d'Ivoire, the IE found that in the short term, total employment, including self and wage, **increased by 12 pp (wage employment increased by 44 pp) and monthly earnings increased by 35%.** In the medium term, there was no impact on total employment, but a **12% increase in earnings.** These impacts are stronger for vulnerable groups.

Paving the Path to Long-Term, Sustained Livelihoods

One of the key findings from the LIPW programs implemented across countries show that these generate employment and earnings in the short term, but are not sustained in the long run. A DIME-commissioned white paper, *Generating employment in poor & fragile states: Evidence from labor market and entrepreneurship programs* by Christopher Blattman and Laura Ralston, found that capital injections, such as cash, business startup grants, capital goods and livestock, stimulate self-employment and raise long-term earning potential.

In Tunisia and the DRC, the findings of the white paper have influenced program design for the introduction of a capital infusion intervention in the form of unconditional cash and/or business startup grants for graduates of the LIPW program, which could help alleviate remaining capital constraints and help program participants achieve sustained livelihoods in the long run.

Regional knowledge transfer informing program design

A LIPW program in Côte d'Ivoire supported transitory employment accessible to unskilled/low skilled youth in urban and peri-urban areas. The program also introduced and tested a range of strategies to help youth transition into more sustainable sources of productive employment, including through entrepreneurship training to help them set-up household enterprises and other types of small businesses. The program embedded an IE to inform decisions concerning targeting of interventions and cost-effectiveness. The IE found that a program that targets the most vulnerable and reduces costs outside transfers would become much more cost-effective, i.e. break even, within about 3 years versus 23 years for running the intervention under the status quo. This has policy implications

for governments on limited budgets aiming to maximize their social protection programs.

The results of the Côte d'Ivoire IE informed targeting mechanisms: it led the program to adjust its quota for women, which initially was 30% and

now is 50%, which has the potential to lead to direct economic benefits. The project also shifted to focus on peri-urban areas. The team is exploring the possibility of including a poverty-targeting filter.

The results of the Côte d'Ivoire IE among other IEs informed the design of a US\$125 million LIPW program (STEP Project) in the DRC to identify how best to deliver it for maximum impact. An important question, however, remains on which mechanisms are at play in effective LIPW programs. As such, DIME influenced the DRC government to test different variations and to experiment with additional interventions, such as trainings and savings mechanisms that aim to lead to sustainable sources of productive employment.





COMMUNICATING LESSONS AND RESULTS

9

The Three Spheres of Influence: Using Experiments to Change Development Practice

Field experiments are revolutionizing the way development research is done. More importantly, they can change development practice.

The central tenet of our dissemination strategy is learning-by-doing through close collaboration and systematic use of data and evidence across three spheres of influence.

First, traditional policy research is done by and among academics skilled in economic analysis, the research then published in reputable journals. Development policy is intermediated by development practitioners, skilled in the practice of policy design and dialogue, then implemented on the ground by governments and other development actors. Seldom in the past have these two worlds interacted in an organic manner to build on each other's comparative advantage and also to address one another's weaknesses. This is our first sphere of influence.

By managing a close collaboration between practitioners and researchers, we secure continued communication of lessons and results that can affect decisions on the ground. The collaboration is aimed at affecting local knowledge (see example on pg. 150), and by informing decisions, the countries' returns to their investments. For example, the latest data from the i2i Monitoring System show that 74% of i2i IEs have already discussed baseline results with the client, and 75% of IEs have discussed final results with the client. Further, around 193 events have been organized in the countries to

disseminate IE results (baseline or final) to policy makers (see chapters below for examples and also chapters 7.2, 7.3 and 8 for more detail on IE results and policy influence).

Second, research is mostly done piecemeal by many independent researchers with little programmatic coordination and guidance. This is our second sphere of influence.

By establishing close collaborations with the Global Practices in the World Bank or with global initiatives, like the GAFSP, DIME develops research programs at the sector level that respond to those practices and initiatives' learning priorities. By producing white papers and summaries of research, and by contributing to learning weeks and many sector-focused events and management meetings, we influence the focus and composition of whole portfolios of operational work. This affects sector practice and understanding (see example on T&C on pg 8 and also chapter 6.2, 6.5 and 6.6 for more detail).

Third, development institutions work in isolation and undertake isolated learning: learning from one another is difficult.

The third sphere of influence leverages cross-institutional collaborations to generate flows of interest and information that can affect global understanding on FCV, transport, or climate change (see examples on FCV and agriculture in this chapter and box on the ieConnect on pg. 106).

The heart of our dissemination strategy is communication through daily collaboration, face-to-face interaction, and operationalization of results and ideas. This is because anything that is only paper-based has long gestation periods



in adoption. In medicine, transformational cures can take years to be adopted, especially when there is no direct profit motive. This is why we complement reports and documents with face-to-face presentations and discussion that can elicit timely policy response.

Stocktaking 2017

Research economists continued to work with project clients and government officials to document evidence during several points of project lifecycles. Outputs produced at this level so far include 30 baseline reports and 27 IE reports. Baseline reports and IE reports stimulate policy dialogue and support the adoption of casual mechanisms based on results. These can influence the scale-up (or scale-down) of a policy at the national level.

As projects and programs mature and produce results, our outreach focus has added to our



large-scale global workshops, different types of events used to share evidence generated at different stages of projects. This is done by bringing together development practitioners, academics, and donors at both global and country levels,

with the latter allowing the outreach to many different country audiences and policy maker levels. DIME workshops, conferences, and events held during the past year included: (1) a global impact-evaluation workshop in Mexico

CROSS INSTITUTIONAL COLLABORATION IN THE AGRICULTURE PROGRAM

The use of evidence in DIME's agriculture portfolio provides an example of how DIME is building connections across global institutions. A core component of the agriculture portfolio is a cluster of evaluations of projects supported by the Global Agriculture and Food Security Program (GAFSP), a US\$1.69 billion multinational program operating in over 35 countries. These evaluations coordinated learning in projects implemented by the World Bank,

the African Development Bank, and the InterAmerican Development Bank, and findings were distributed to audiences including the United Nations High Level Policy Forum, the European Union, the World Food Programme, and many others. The model established by DIME's collaboration has led to new workshops and engagements to bring these approaches to other institutions including a collaboration with the EU in Rwanda (described at

the end of chapter 6.2); and two workshops with the World Food Programme in Rome (November 2017) and in Nairobi (January 2018). The workshops were aimed at encouraging the World Food Programme to learn from the findings of the DIME-GAFSP evaluations and to conduct their own evaluations to test strategies for improving delivery of WFP programs such as gender-based safety nets or school feeding programs.



City on T&C, (2) a Nigeria 3-day event on Health, showcasing how Impact Evaluation research drives innovation and improves outcomes, (3) two global innovative workshops in Washington DC on Identification for Development (ID4D), and on data and evidence for Justice Reform (DEJURE), (4) a global workshop on Transport in Lisbon, (5) a global 3 day follow-up workshop in Rome on Agriculture, (6) workshops and events in Washington DC on (i) Gender Based Violence, (ii) Employment in Fragile, Conflict-Affected, and Violent Contexts, (iii) Agriculture, (iv) Governance Forum and Inside Government Conference, and (v) a Policy Survey Training. In all, 33 DIME

seminars and 5 ieGovern seminars, workshops, and trainings took place over the course of the year, and a total of more 1,000 people attended these events (see table 2.1). And as mentioned earlier, an additional 193 events have been organized in the countries to disseminate IE results (baseline or final) to policy makers.

Further, DIME organizes biweekly seminars in the World Bank, where researchers present papers and results across the portfolio, targeting a wide Bank audience from the GPs and country teams. Only during this last reporting cycle, 22 DIME seminars have been held in the World Bank (73

CROSS INSTITUTIONAL COLLABORATION IN THE FRAGILITY, CONFLICT, AND VIOLENCE PROGRAM



The Fragility Conflict and Violence team has been working with the World Bank to mobilize global knowledge in addressing sustainability challenges in fragile contexts, with the objective of leveraging research in building programmatic approaches. An example is the partnership created between DIME's FCV team and the World Bank's Gender and FCV Cross-cutting Solution Areas; the Social Protection & Jobs Global Practice; and the East Asia Gender and Innovation Lab (GIL). Since 2014, DIME's FCV program has made it a priority to address the fundamental problem underlying the current development approach in fragile states in which programs are often carried out as emergency operations, yet with unclear objectives on whether the goal is to support livelihoods or reduce

violence, or both. DIME's efforts in presenting existing evidence and proposing new research projects that go beyond short-term employment and address capital constraints played a central role in changing the way in which World Bank operations can view their programmatic approach. This has led to the FCV and Gender CCSAs and the Jobs and GIL groups' active participation in DIME's research on cash injection programs in fragile contexts, such as Tunisia, Egypt, or the DRC. The partnership's success fed into two Learning Events on Gender and Employment in FCV, which brought together prominent scholars, practitioners and policy makers to foster, explore, and present evidence on the key mechanisms of capital constraints in FCV contexts.

Beyond internal communication, the FCV program has had a central role in building relationships with other institutions. For example, in October 2017, the German Agency for International Cooperation (GIZ) invited DIME for a 3-day workshop to discuss ongoing work and avenues for future collaboration. The results on LIPW programs were of interest to GIZ and the German Government, who both are aiming to move away from short-term solutions and towards sustainable programs which address the needs of vulnerable populations in the context of instability. The GIZ has decided to work towards building a partnership with DIME to push forward the knowledge frontier in FCV contexts.

in total since program inception). The team is also very active in presenting i2i research externally, such as in academic conferences, M&E practitioner events, and donor partner events. Going forward (and starting with the next reporting cycle), this data will be collected systemically. Table 9.1 provides a few examples for each thematic area.

The DIME team also continued to garner wider policy outreach through (a) policy briefs, (b) World Bank working papers, (c) journal submissions, and (d) DIME newsletter, a quarterly summary of DIME's research and activities. We have produced 26 working papers, 15 publications and 48 policy briefs, providing opportunity for researchers to proffer advice or solutions stemming from research.

■ ■ **Table 9.1** Examples of DIME external presentations

Presenter	Location	Event	Sponsor
Agriculture			
Florence Kondylis	Maputo, Mozambique	Delivering to the New EU Consensus for Development	EU
Paul Christian	Chicago, USA	2017 Annual Meeting of the Agricultural and Applied Economics Association (AAEA)	AAEA
Governance			
Daniel Rogger	London, UK	DFID Chief Economist Seminar	DFID
Bilal Siddiqi	London, UK	Chief Economist's Office; Peace talks seminar series	DFID
Climate Change			
Aidan Coville	Oslo, Norway	Evaluation and Learning for International Sustainable Forest Initiatives	NORAD
Aidan Coville	Washington, DC	American Evaluation Association conference	AEA
Aidan Coville	Massachusetts, USA	North East Development Consortium Conference	NEUDC
Trade & Competitiveness			
Guadalupe Bedoya	Geneva, Switzerland	Launch of the WHO Bulletin theme issue "Measuring Quality of Care"	WHO
Caio Piza	Washington, DC	3ie + IFPRI Seminar Series on Impact Evaluation	3ie/IFPRI
Caio Piza	Lisbon, Portugal	European Economic Meeting	EEA
Caio Piza	North Carolina, USA	Society of Labor Economists	SOLE
Transport			
Theophile Bougna	Toulouse, France	Toulouse School of Economic: Transport Corridors and Their Wider Economic Benefits: A Critical Review of the Literature	TSE
General Cross-sector			
DIME team	Eschborn, Germany	GIZ-World Bank DIME Impact Evaluation Event Transforming Development through Evidence-Based Policy	GIZ
DIME team	Bern, Switzerland	SECO/SDC - World Bank DIME Impact Evaluation Event Transforming Development through Evidence-Based Policy	SECO/SDC
DIME team	Ispra, Italy	EU JRC-World Bank DIME Impact Evaluation Event Transforming Development through Evidence-Based Policy	JRC
DIME team	Brussels, Belgium	EU Commission - World Bank DIME Impact Evaluation Event Transforming Development through Evidence-Based Policy	EU/WB

Note: This is not a complete list. It provides just a few examples for each thematic area that took place during the last year. Going forward, this will be collected systematically across the program and included in the next annual reports.



REFERENCES

- Acemoglu D. and J. Robinson, 2012, *Why Nations Fail: Origins of Power, Poverty and Prosperity*, Crown Publishers (Random House), March 2012.
- Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: Origins of Power, Poverty and Prosperity*. Crown Publishers (Random House).
- Atkin, David. 2017. "Organizational Barriers to Technology Adoption: Evidence from Soccer-ball Producers in Pakistan." *The Quarterly Journal of Economics* 132.3: 1101-1164.
- Banerjee, Abhijit, Eliana La Ferrara, Victor Orozco. Forthcoming. "The Entertaining Way to Behavioral Change: Fighting HIV with MTV." Policy Research Working Paper; *Impact Evaluation Series*. Washington, DC: World Bank Group.
- Benyishay, Ariel, Maria Ruth Jones, Florence Kondylis, Ahmed Mushfiq Mobarak. 2016. "Are Gender Differences in Performance Innate or Socially Mediated?" Policy Research working paper; no. WPS 7689; *Impact Evaluation Series*. Washington, DC: World Bank Group.
- Bruhn, Miriam, and David McKenzie. 2014. "Entry Regulation and The Formalization of Microenterprises in Developing Countries." *The World Bank Research Observer* 29.2: 186-201.
- Cameron, Drew B., Anjini Mishra, Annette N. Brown. 2016. "The Growth of Impact Evaluation for International Development: How Much have we Learned?" *Journal of Development Effectiveness*, 8.1:1-21.
- Campos, Francisco, Aidan Coville, Ana Fernandes, Markus Goldstein and David McKenzie. 2013. "Lessons for Matching Grant Programs from Failed Attempts to Evaluate Them," *World Bank Other Operational Studies* 17041, The World Bank.

- Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, Victor Orozco. 2013. "Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-Out Behavior." Policy Research working paper; no. WPS 6545; *Impact Evaluation Series*; no. IE 99. Washington, DC: World Bank.
- Dercon, Stefan, John Hoddinott, and Tassew Woldehanna. 2005. "Shocks and Consumption in 15 Ethiopian Villages, 1999-2004." *Journal of African Economies* 14.4: 559.
- Dercon, Stefan, and Joseph S. Shapiro. 2007. "Moving On, Staying Behind, Getting Lost: Lessons on Poverty Mobility from Longitudinal Data." *Moving Out of Poverty* 1: 77-126.
- Douglass C. North. 1991. "Institutions," *Journal of Economic Perspectives*, American Economic Association, vol. 5(1), pages 97-112, Winter.
- Duflo, Esther. 2006. "Poor but Rational." *Understanding Poverty*: 367-78.
- Edenhofer, Ottmar, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler et al. ed. 2014. "IPCC, 2014: Summary for Policymakers." In: *Climate Change 2014: Mitigation of Climate Change*. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK and New York, NY, USA: Cambridge University Press.
- Galiani, Sebastian, Paul Gertler, and Ernesto Schargrotsky. 2005. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113.1: 83-120.
- Groh, Matthew, et al. 2016. "The Impact of Soft Skills Training on Female Youth Employment: Evidence from a Randomized Experiment in Jordan." *IZA Journal of Labor & Development* 5.1: 9.
- Goldstein, Markus P; Kenneth Hounbedji, Florence Kondylis, Michael B. O'Sullivan, Harris Selod. 2015. "Formalizing Rural Land Rights in West Africa: Early Evidence from a Randomized Impact Evaluation in Benin." Policy Research working paper; no. WPS 7435; *Impact Evaluation Series*. Washington, D.C.: World Bank Group.
- Kondylis, Florence, Valerie Mueller, and Jessica Zhu. 2017. "Seeing is Believing? Evidence from an Extension Network Experiment." *Journal of Development Economics* 125: 1-20.
- McKenzie, David, and Christopher Woodruff. 2013. "What are we Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *The World Bank Research Observer* 29.1: 48-82.
- McKenzie, David J., and Olga Puerto. 2017. "Growing Markets Through Business Training for Female Entrepreneurs: A Market-Level Randomized Experiment in Kenya."
- McKenzie, David; Assaf, Nabila; Cusolito, Ana Paula. 2015. "Spurring Innovation with Matching Grants: Evidence from Yemen." *Finance PSD Impact*, no. 34. World Bank.
- Milanovic, Branko. 2016. *Global Inequality: A New Approach for the Age of Globalization*. Cambridge MA: Harvard University Press.
- Nguyen, Quang and Trang Nguyen. 2016. "The Role of Manager Perception in Management Practices: Insights from a Randomized Experiment." *Mimeo*.
- Piketty, Thomas. 2014. *Capital in the 21st Century*. Cambridge MA: Harvard University Press.
- Sachs, Jeffrey. 2005. "The End of Poverty: Economic Possibilities for Our Time." *International Journal* 60.3: 849

- Samii, Cyrus, Laura Paler, Larry Chavis, Parashar Kulkarni, and Matthew Lisiecki. 2014a. "Effects of Decentralized Forest Management (DFM) on Deforestation and Poverty in Low and Middle-Income Countries: A Systematic Review." *Campbell Systematic Reviews* 10, no. 10.
- Samii, Cyrus, Matthew Lisiecki, Parashar Kulkarni, Laura Paler, and Larry Chavis. 2014b. "Effects of Payment for Environmental Services (PES) on Deforestation and Poverty in Low and Middle-Income Countries: A Systematic Review." *Campbell Systematic Reviews* 10, no. 11.
- Sen, Amartya. 1992. *Inequality Reexamined*. Clarendon Press.
- Sen, Amartya. 2001. "The Many Faces of Gender Inequality." *New Republic*: 35-39.
- Stewart, Ruth, Laurenz Langer, Natalie Rebelo Da Silva, Evans Muchiri, Hazel Zaranyika, Yvonne Erasmus, Nicola Randall Shannon Rafferty, Marcel Korth, Nolizwe Madinga, Thea de Wet. 2015. "The Effects of Training, Innovation and New Technology on African Smallholder Farmers' Wealth and Food Security: A Systematic Review." *Campbell Systematic Reviews* 11.16.
- Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49.2: 326-65.
- Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5.4: 297-323.
- Wantchekon, Leonard and Piero Stanig. 2015. "The Curse of Good Soil? Land Fertility, Roads, and Rural Poverty in Africa." Working paper.
- World Bank. 2013. *World Development Report 2014: Risk and Opportunity—Managing Risk for Development*. Washington, DC: World Bank.

ANNEXES

**See online version of report for
full annexes**

1. Results Framework
2. List of Impact Evaluations
3. Monitoring Indicators
4. Concept Note Template
5. Concept Note Budget Template
6. IE Helps Deliver Results

PHOTO CREDITS

pages vi, 21, 114: © Simone D. McCourtie / World Bank; pages viii, 68, 136: © Scott Wallace / World Bank; pages 2, 11, 151: © Dominic Chavez / World Bank; pages 6, 24, 38, 40, 43, 54, 62, 144, 149: © A'Melody Lee / World Bank; pages 16, 30, 26, 30, 32, 57, 59, 60, 99, 104, 118, 152, 154: © Barbara Coello / World Bank; page 34: © Daniella Van Leggelo-Padilla / World Bank; page 37: © Stephan Gladieu / World Bank; pages 45, 47, 66, 146: © Peter Kapuscinski / World Bank; page 48: © John Hogg / World Bank; page 53: © Mariana Ceratti / World Bank; page 71: © Jamie Martin / World Bank; page 77: © Dominic Sansoni / World Bank; page 88: © Curt Carnemark / World Bank; page 96: © Rob Beechey / World Bank; page 101: © Hendri Lombard / World Bank; pages 108, 122, 147: © Arne Hoel / World Bank; page 132: © Rama George-Alleyne / World Bank; page 158: © Alex Baluyut / World Bank

