Leveraging Knowledge Generation for Policy Impact: Recommendations for the World Bank

AMANDA GLASSMAN, JULIA KAUFMAN, RUTH LEVINE, AND JANEEN MADAN KELLER

Background: CGD’s Working Group on New Evidence Tools for Policy Impact

Since the 2006 release of the Center for Global Development report *When Will We Ever Learn: Improving Lives Through Impact Evaluation*, and building on evaluations of cash transfer programs in the 2000s, there have been nearly two decades of progress in generating and using evidence for public policy decisions and development programs. One area of clear momentum is the steady increase in the number of impact evaluations, a rigorous approach that establishes the attributable net impact of a project or program. Their ability to assess attribution makes impact evaluations uniquely well suited for decision making. A global community of researchers and organizations conducting these evaluations and related evidence activities has also grown substantially in recent years. All the while, notable advances in data and evaluation methodologies and practices have enabled faster, lower-cost, and larger-scale evaluations, expanding the application of impact evaluation tools to new domains. And increasingly, impact evaluations are paired with complementary quantitative and qualitative information that help derive policy-relevant inferences.

Nevertheless, the COVID-19 pandemic put a spotlight on an unfinished agenda and underscored the need for high-quality, timely, and context-specific evidence. Across sectors, decision makers within governments, aid

KEY RECOMMENDATIONS FOR THE WORLD BANK

1. Embed impact evaluation and related evidence resources across the Bank’s operational structure and develop sectoral, regional, and country learning agendas
2. Allocate dedicated resources to routinely finance rigorous evaluation
3. Strengthen and centralize tracking and publication systems for data and evidence
4. Develop formal mechanisms to promote evidence-to-policy partnerships, capacity strengthening, and demand generation
agencies, multilateral organizations, and NGOs have not yet fully harnessed the value of evidence for better public policies. Numerous challenges limit evidence use. On the demand side, impact evaluations may lack relevance to policy decisions and fail to respond to the priorities, timelines, and questions of decision makers. On the supply side, decision makers often lack institutional incentives and funding to generate and act on relevant evidence. Current funding models further compound these challenges by contributing to misaligned incentives between policymaker needs and academic researchers.

In response to these challenges and building on progress to date, CGD launched the Working Group on New Evidence Tools for Policy Impact to develop a renewed agenda for the next generation of investments in impact evaluation and related evidence systems to enhance their value for real-world decision making. The working group brought together a diverse set of policymakers and experts to review recent progress and examine how to address remaining obstacles to the use and utility of evidence for global development.

The working group’s final report highlights how far the field has come in addressing persistent critiques about the scale, generalizability, and policy utility of impact evaluation. It also offers recommendations to the development community on “what and how to fund to deliver on the promise of impact evaluation and bolster the broader evidence ecosystem” as two intertwined goals. Specifically, the working group proposes five ways to improve impact evaluation funding and practice: (1) design evaluations that start from the policy question and decision space; (2) harness technology for timely, lower-cost evidence; (3) advance locally grounded evidence-to-policy partnerships; (4) enact new incentives and structures to strengthen evidence use; and (5) invest in evidence leaders and communities. The working group’s reinvigorated agenda aims to optimize the benefits and full potential of impact evaluation for improved social and economic well-being around the world.

To illustrate the application of this agenda to specific development funders, the working group developed detailed recommendations for three key audiences with strong existing foundations for evaluation and evidence use: philanthropies, USAID, and the World Bank. This brief lays out how the World Bank can more systematically embed impact evaluation and related data and evidence functions into programming, implementation, and country support. The recommendations are based on working group discussions, consultations with experts currently and formerly at the World Bank, and background research over the course of the project. Accompanying briefs directed to philanthropies and USAID can be found at www.CGDev.org/evidence-to-impact.

Leveraging the World Bank’s role in knowledge production for greater impact

Multilateral development banks, such as the World Bank, work closely with country governments to articulate policy goals, design and implement policies and programs, and assess their effectiveness. The World Bank is both a funder and implementer of evaluations globally and is one of the largest players in the field. Its extensive role in knowledge production gives the World Bank a comparative advantage in generating and using evidence, and its focus on serving client governments often exemplifies the principles of demand-driven evaluation. The critical role of evaluation—and
specifically impact evaluation to attribute causality—in assessing program effectiveness and promoting operational learning is a well-established institutional principle. Still, less than five percent of Bank projects have been subject to formal impact evaluation methods since 2010.\(^3\)

Over the coming years, there is significant scope to improve how evidence and evaluation functions, notably for impact evaluation, are structured within the Bank and to scale up efforts to generate greater value for policy use, mobilize additional financing, and achieve economies of scale. The 2021 Strategic Framework for Knowledge calls for the World Bank Group to better align financing with global knowledge to generate development solutions in a timely, contextualized, and integrated way.\(^4\) As the Bank’s leadership operationalizes this framework, now is the time for a reinvigorated commitment to evidence generation and use.

The World Bank currently funds and implements impact evaluation through a range of approaches. One modality is through thematic trust funds. For instance, the Development Impact Evaluation (DIME) group, funded by the Impact Evaluation to Development Impact (i2i) trust fund, conducts research in 60 countries with 200 partner agencies, leveraging a $180 million research budget against $18 billion in development finance across sectors.\(^5\) Further, the Global Education Evidence Advisory Panel, funded by a trust fund, recommends “smart buys” in the education sector, systematizes impact evaluation evidence and carries out new evaluations to inform countries’ decisions about where to allocate their budget and reform efforts.\(^6\) Another example is the Health Results Innovation Trust Fund (HRITF) that financed rigorous impact evaluations alongside health operations to promote learning about results-based health financing. HRITF supported large-scale evaluations in 28 countries, helping generate global knowledge about the possibilities and limitations of results-based financing models for health.\(^7\)

Historically, the trust fund model has provided dedicated resources for evaluation and enabled collaboration between research and operational staff, contributing to improvements in Bank operations. For example, the i2i trust fund demonstrates that modest investments in data and research, including impact evaluation, can increase the effectiveness of Bank operations by large margins, while also building country capacity for evidence use.\(^8\) The i2i trust fund, which serves as the umbrella facility for impact evaluation, also provides a Bank-wide governance structure with dedicated guidelines for technical quality, ethics, transparency, reproducibility and data privacy and security. Yet the trust fund model poses challenges, including fragmentation of evaluation activities across the Bank and vulnerabilities to insufficient and volatile aid flows due to dependence on specific external donors.\(^9\)

Beyond dedicated trust funds, impact evaluations are directly embedded in World Bank financed operations. But whether or not impact evaluations are directly linked to operations and used to assess their effectiveness depends on requests from client governments and/or the interest of individual Task Team Leaders (TTLs). In some cases, Development Economics (DEC) staff and other researchers are encouraged to conduct research that serves operational needs, although publication in peer-reviewed journals is typically a primary incentive for research staff. TTLs can motivate and request funding for evaluations that are directly linked to operations but interest in impact evaluation varies across TTLs, and management incentives focus on project approval rather than evidence generation and use. Staff turnover further compounds challenges for longer-term evaluation horizons and follow-up policy uptake.

The importance of designing impact evaluation in response to key policy moments as well as demand from policymakers (who can make related policy decisions based on that evidence) has not yet translated into routine and widespread practice. Further, beyond informing project-specific policy decisions in a particular country, evaluation results may also be useful in other country contexts, and thereby represent a critical global public good. Yet the variable, ad hoc approach
taken across the World Bank makes it difficult to understand the breadth of evaluation activities underway and further hinders the use of evidence for more impactful lending and related policy dialogue and decision-making, as well as evidence synthesis on what works in specific policy areas. The potential scope of the World Bank’s role in producing globally- and policy-relevant knowledge remains unfulfilled.

Several other challenges persist related to the broader set of data and evidence activities. For example, World Bank policy requires dataset publication, but compliance is low. To date, Bank-supported data and evidence activities have not consistently linked to government decision points. Further, impact evaluations do not always include intervention costs and efficiency analyses, optimize use of administrative data and other existing data sources, or prioritize capacity strengthening through country-based evaluation organizations. Though pacing impact evaluations to align with national budget cycles can be challenging, other evidence can be brought to bear in a more systematic and timely manner. Indeed, the Bank is taking encouraging steps to move this agenda forward.

The World Bank could make a major contribution towards systematically building in-country capacity to conduct evaluations and use evidence for decision-making. But its knowledge production function is not financed with this objective in mind. While some specific country lending operations include financing for government evaluation units, most research is structured to take place predominantly in-house or via one-off consultancies. As a notable exception, DIME uses a coproduction model throughout the research process aimed at capacity strengthening with both governments and local researchers; for example, see their recent call for expression of interest from local researchers for the Partnership for Economic Inclusion’s Impact Evaluation Collaborative. And the newly launched Global Evaluation Initiative managed by the Independent Evaluation Group seeks to develop country-owned monitoring and evaluation capacities. CGD analysis suggests spillovers from World Bank project monitoring and evaluation systems can be leveraged to improve client governments’ own monitoring and evaluation.

### BOX 1. HEALTH RESULTS AND INNOVATION TRUST FUND EVALUATIONS

A recent example illustrates the difficult balance between speed, policy relevance, and rigor in impact evaluations. In 2007, 36 impact evaluations were commissioned of performance-based financing projects in health connected to a set of World Bank operations financed by HRITF. Because many evaluations were slow to execute, pilot interventions were extended or ended before results were available, hampering their policy value. Available findings were systematized in 2021 and published by DEC in May 2022 (five years after evaluations first rolled out, with an acknowledgement that delays were due to country implementation). Analysis predominantly took place outside of operational support to ongoing World Bank projects, and the extent of feedback to policymakers is not documented in the public domain. DEC’s independence and rigor were important since findings suggested that the approach needed adjustment. The sequence of events illustrates the difficult balance between real-world timelines, technical rigor, and operational engagement and relevance.

This brief proposes four actions the World Bank and its shareholders should pursue to turn the page from ad hoc to more systematic evaluation funding approaches that make a real difference for government systems and policies. Overall, the Bank should institutionalize the use of evaluation to determine whether projects have their intended impact on improving lives as part of relevant approval and accountability processes. The potential rate of return is immense: a $1 million impact evaluation could save hundreds of millions in mistargeted lending.
1. Embed impact evaluation and related evidence resources across the Bank’s operational structure and develop sectoral, regional, and country learning agendas

World Bank country staff should collaborate with select national ministries to develop country-owned learning agendas that target priority knowledge gaps in specific sectors and/or regions. The Bank’s leadership should mobilize a dedicated resource envelope to implement this agenda (see recommendation 2).

The DIME model, where coproduction enables capacity strengthening among partners, could be mainstreamed across the Bank and applied to other high-priority areas. This model has demonstrated its track record in improving program design and adaptation. Another relevant model is the Africa Gender Innovation Lab (GIL) where centralized analytical functions and resources are deployed and disbursed to regional offices to respond to policy priorities and client questions. A system of regional and/or sector-focused impact evaluation labs embedded within the wider operational structure of the Bank can streamline evidence dissemination from researchers and TTLs to policymakers in client countries. Trust funds have historically funded such efforts, including the examples described earlier, but the idea would be to scale up and expand support for such labs across a wider range of thematic topics.

Specifically, sector impact evaluation programs should be scaled up. This involves deploying a team research and operational staff to collaboratively develop data and analysis, including through impact evaluation, to inform future sector-specific operations. In tandem, country research programs, which help develop the underlying data ecosystem and enhance capacity to use advanced analytics to improve efficiency in government processes, should be expanded. Increasing the scope of these efforts would help reduce the costs of generating and disseminating knowledge by fostering long-term relationships with client governments and other in-country evidence partners, and in turn help overcome challenges related to generating and sustaining demand for high-quality evidence. The Bank should also set up a Bank-wide internal quality assurance group to track progress against learning agenda implementation and course correction with managers.

2. Allocate dedicated resources to routinely finance rigorous evaluation

The Bank needs a range of evaluation structures that can inform policy and lending operations to make timely adjustments, while also protecting the integrity and independence of the research to remain credible. Since impact evaluation currently depends on insufficient—and at times uncertain—external trust fund resources and fragmented operational interest, new dedicated resources are needed to undertake evaluations more strategically and systematically across the Bank’s portfolio. The World Bank’s leadership could consider different financing mechanisms, each with their own benefits and drawbacks.

There are several options for financing impact evaluation and related evidence activities. One option in low-income countries is to designate a share of the regional International Development Association (IDA) window for data and evaluation functions, including impact evaluation. Where feasible, this funding could be used to support national evaluation and evidence entities in both the public and private sectors. Dedicating a relatively small portion of IDA resources to evaluation related functions would enhance the overall impact of World Bank finance, and therefore represents good value for money.

Another option in both low- and middle-income countries is to establish a defined allocation of IDA and IBRD lending towards data and research for all Bank-financed operations. Integrating these essential functions across all lending operations would help generate data and evidence that is operationally relevant and responsive to a client country’s decision-making needs. However, this approach could also
introduce some risks; bundling knowledge production and lending functions may raise concerns related to conflict of interest when governments are both judge and jury of evaluation results, necessitating the need for an approval structure where researchers review the available evidence base prior to loan approval. For example, the Africa GIL model provides an initial tranche of $50,000 for evaluation design, with the remaining funds released once a detailed concept note is reviewed against specific criteria (e.g., description of a counterfactual and power calculations) by several researchers and operationally linked staff.

Third, the Bank should also continue to produce research and evidence, including through trust funds, in addition to systematically integrating funding for data and evaluation-related activities into operations. As such, trust funds should be adequately funded to support evaluations and related evidence functions that are not necessarily project specific, including studies that measure longer-term outcomes beyond the time horizon of a specific operation, are regionally focused, and/or relate to global public goods.

Other experts propose embedding knowledge management within the Bank’s Governance Global Practice, which could pursue joint ventures with other global practices as a way to contribute to the shared goal of strengthening in-country evidence functions and capacities. Such efforts could build on the collaboration between DIME and the Governance Global Practice to improve access to administrative data and other analytical and digital tools to enhance uptake of evidence and data for real-time policy guidance. Another option offered by some experts is to institutionalize the use of evaluation as an economic safeguard in line with the Bank’s other environmental and social safeguard policies.

3. Strengthen and centralize tracking and publication systems for data and evidence

Amid the disparate but widespread evaluation efforts being undertaken, the Bank should move toward (1) tracking all evaluation surveys and impact evaluations conducted—from design to completion—and their costs; (2) making survey data, impact evaluation documentation, and findings publicly available in a timely manner; and (3) rethinking the knowledge adaptation and adoption model. This is especially important in light of staff turnover and project cycles.

A new World Bank publications rule that requires data and code to be uploaded to a centralized hub is a step in the right direction. The Bank could go further and commit to making all data immediately public by default (with exceptions requiring a vetted justification) and invest in a data portal to decrease the cost and increase the speed of navigating and acting on these datasets, building on the Bank’s work in financing the Living Standards and Measurement Surveys. For example, datasets currently published in the DIME portal represent just a subset of all World Bank evaluations. Operationalizing these policies will require the right incentives in combination with enforcement by managers, and should not be left to the preferences of individual TTLs or researchers. Under the status quo, incentives to release data are weak, though in some instances there are requirements from trust fund partners. Another option is to offer performance review credits for posted datasets.

Data repositories can help decrease the cost and increase the speed of knowledge generation. Tapping into tax data, procurement data, and other administrative data sources would also help make evidence generation faster and cheaper, but requires greater investments in data collection, quality, and infrastructure (including national statistical systems) to link these data sources to each other and make them accessible and usable for researchers and policymakers. In this vein, the World Bank’s new Global Data Facility is an exciting step. Further, a new dedicated facility called Microdata and Evidence for Government Action (MEGA) will serve as a global repository of government-owned administrative microdata with the aim to standardize and improve data quality and usability. This initiative could enable government use of data to improve the efficiency of government processes and services, and in turn facilitate the use of data by Bank staff to inform projects and policy dialogue. Yet harnessing these benefits will require significant political commitments from
governments to provide data in real time. The World Bank should also lead the way in expanded partnerships between international organizations and private companies to facilitate use of third-party data for development research, building on the recent Development Data Partnership.  

4. Develop formal mechanisms to promote evidence-to-policy partnerships, capacity strengthening, and demand generation

More equitable evidence-to-policy partnerships would increase the quality, relevance, and visibility of the World Bank’s research and evaluation work within client governments, national universities, and civil society groups, as well as at the global level. While the Bank already pursues numerous partnerships with external research organizations that play a critical role in capacity strengthening and demand generation through DIME and other efforts, there is scope to expand. The Bank currently contracts country-based survey firms through one-off consultancy agreements, including through vendor shortlisting and open competition. Where feasible, the Bank should shift to providing medium-term financing for select organizations with the potential to build their evaluation capacity and expand their client base to interested governments and other research funders.

The World Bank could also co-invest with partner research organizations in local evaluation firms to further develop the enabling ecosystem for evidence generation and use, though current procurement rules would likely need to be changed. Further, the Bank can provide policy lending to support government-based or semi-autonomous evidence-to-policy initiatives, which are proximate to information on policy windows and ideally structured by design to ensure research questions align with government priorities. Lastly, creating an internal roster of local partners in client countries—including universities, researchers and research consortia, and survey firms—would be a helpful resource for staff as they seek to expand the scope of impact evaluations and related evidence activities across the Bank’s portfolio.

Endnotes


Could be through regional units or global practices. Examples of Sector Programs include ieConnect program (Transport GP/DIME), Bureaucracy Lab (Governance GP/DIME), Partnership for Economic Inclusion (SP) GP/DIME, Compel (TIC GP/IFC/DIME/DECRG), and GIL Federation (GIL, DIME, Regions). Examples of country research programs include Nigeria human development to address the COVID reversal; Rural Transformation programs in Rwanda and Mozambique to explore market approaches to rural and economic transformation; the Senegal Green program to showcase the value of innovative data and analysis to tackle leading causes of particulate matter and methane emissions across a portfolio of infrastructure, urban and waste management operations; and Kenya Digital to leverage innovative digital technologies and data to tackle key development issues across multiple sectors. The EU supports the country impact evaluation program approach in four SSA countries.


16. Could be through regional units or global practices.

17. Examples of Sector Programs include ieConnect program (Transport GP/DIME), Bureaucracy Lab (Governance GP/DIME), Partnership for Economic Inclusion (SP) GP/DIME, Compel (TIC GP/IFC/DIME/DECRG), and GIL Federation (GIL, DIME, Regions). Examples of country research programs include Nigeria human development to address the COVID reversal; Rural Transformation programs in Rwanda and Mozambique to explore market approaches to rural and economic transformation; the Senegal Green program to showcase the value of innovative data and analysis to tackle leading causes of particulate matter and methane emissions across a portfolio of infrastructure, urban and waste management operations; and Kenya Digital to leverage innovative digital technologies and data to tackle key development issues across multiple sectors. The EU supports the country impact evaluation program approach in four SSA countries.


19. One example of such an effort is the Millennium Challenge Corporation’s investments in Morocco to set up a research partnership with a local think tank to conduct rigorous evaluations of Moroccan labor market programs and policies. As background, the Government of Morocco and MCC are implementing a first-of-its-kind evaluation partnership as a part of the $461 million Morocco Employability and Land Compact. MCA-Morocco, the local implementing entity, signed MCC’s first Cooperation Agreement, a funded partnership within a country program. This first MCA-driven partnership agreement has brough Nobel prize-winning economic analysis approaches from MIT and Harvard to partner with a Moroccan think tank and created an Employment Lab to conduct rigorous research into Moroccan labor market programs and policies. This research is coupled with training and capacity building to key Moroccan policymakers to promote evidence-based decision-making. See more at Millennium Challenge Account Morocco. “Mission,” https://mccmorocco.ma/mission.


