

BROOKINGS

Future Development

A quiet revolution in impact evaluation at USAID

Ariel BenYishay, Rachel Trichler, Dan Runfola, and Seth Goodman Monday, October 8, 2018

USAID is often criticized for not rigorously evaluating its development programs. Like many other aid agencies, USAID has historically prioritized *performance evaluations* that measure the extent to which its programs generate their expected outputs, rather than *impact evaluations* that provide evidence of their net, attributable impacts. Senior USAID administrators have for the better part of the last decade expressed rhetorical support that more impact evaluations should be undertaken. However, high-level interest has not quickly or easily translated into practical changes in the way that that USAID does evaluation.

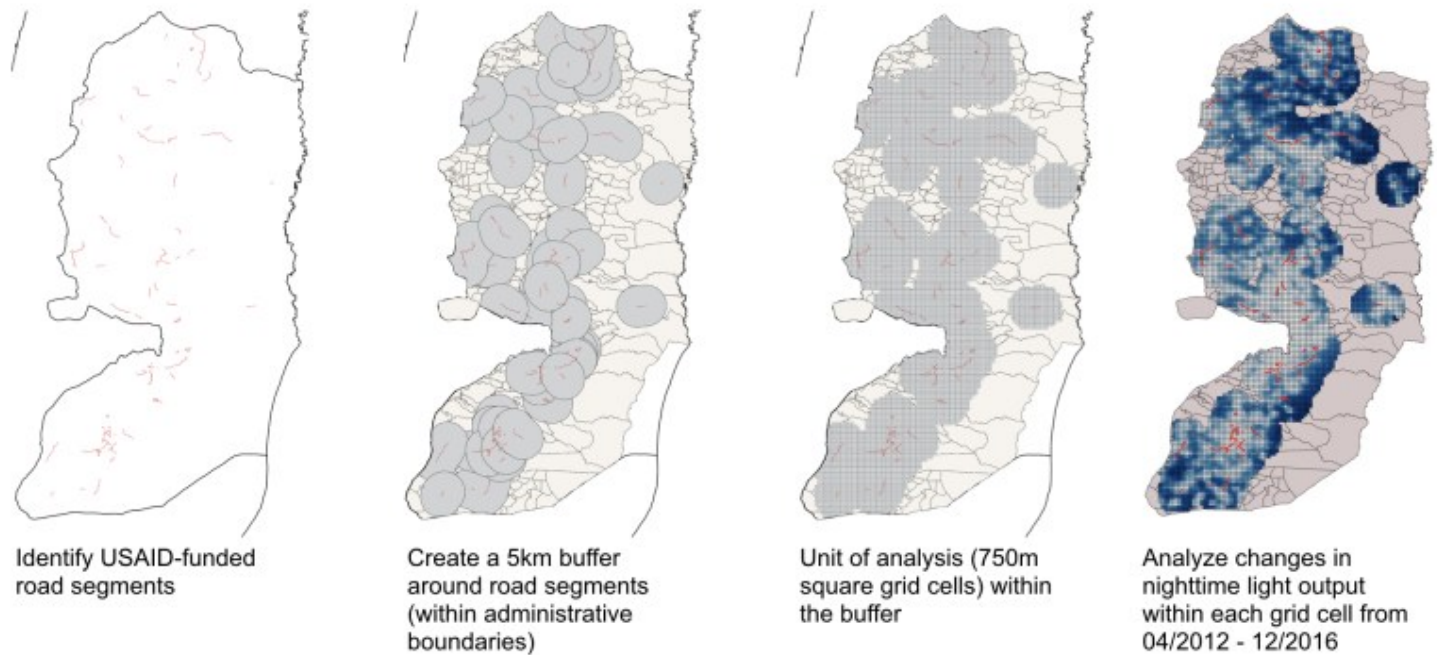
But things are finally beginning to change inside the Ronald Reagan Building and across the agency's global network of field missions. A growing number of USAID-funded programs are being subjected to impact evaluations. These evaluations establish a credible counterfactual, addressing the question of what would have happened in the absence of a given USAID program in order to isolate the change in a development outcome—or a set of development outcomes—that is attributable to that particular program as opposed to other factors. One of the principal reasons why the agency is moving in this new direction is the availability of new data, methods, and tools.

A case in point is USAID West Bank/Gaza's recent \$900 million investment in rural infrastructure. Under the Infrastructure Needs Program II (INP II), USAID funded the construction or rehabilitation of 59 rural road segments in the West Bank. It was interested in subjecting the program to a rigorous impact evaluation. However, a randomized controlled trial (RCT)—the “gold standard” in impact evaluation—was not a viable option. Randomly assigning the placement of rural roads is neither feasible nor advisable. USAID decided to instead commission a geospatial impact evaluation (GIE). GIEs use precisely georeferenced intervention data and outcome data to establish a

counterfactual retroactively, eliminating the need to randomly assign individuals, firms, or communities into treatment and control groups at the outset of a program. Also, many GIEs leverage readily available data like satellite observations, so they can often be implemented at a fraction of the time and cost of traditional RCTs (which usually require custom baseline and endline surveys).

We are members of a research team at [AidData](#) that conducted the evaluation of INP II (available [here](#)). We first identified the exact routes of the road segments supported by USAID and the precise dates when these investments were completed. In order to identify the geographical areas that could plausibly benefit from the improved conditions of these roads, we created 5-kilometer buffers around each of these 59 road segments. We then subdivided these “catchment areas” into 750 x 750-meter grid cells corresponding to our outcome measures. In order to measure a key outcome that the INP II sought to improve, we used a [proxy for local economic development](#) that is particularly useful [in countries where GDP cannot be consistently measured over time and geographic space](#): high-resolution, high-frequency satellite imagery on nighttime light output. More specifically, we used NOAA’s Visible Infrared Imaging Radiometer Suite to develop a monthly measure of luminosity that could be consistently tracked between April 2012 (11 months prior to the first road improvements) and December 2016 (five months after the last road improvements) in each and every 750 meter square grid cell within 5 kilometers of an improved road segment. The maps provided in the figure below illustrate each step of this process.

Geospatial impact evaluation of USAID-funded rural roads program in the West Bank



With the resulting dataset of nearly 400,000 grid cell-month observations, we used a quasi-experimental panel framework to rigorously estimate program impacts due to completion of road improvements (“treatment”). We compared post-treatment nighttime light in each grid cell to counterfactual outcomes obtained from that same grid cell’s own preceding nighttime light levels and trends, as well as the outcomes of grid cells near not-yet-improved road segments. The variation in the timing of road improvements across grid cells and the inclusion of month- and grid cell-level fixed effects at fine geographic levels address concerns about confounding and omitted variables (e.g., the implementation of a nearby employment program, falling electricity costs in a particular area).

The results of the evaluation are encouraging. We find strong evidence that local economic output, as measured by remotely sensed nighttime light output, increased due to the implementation of INP II. In communities that benefited from improvements to multiple road segments, we find that INP II had even larger economic impacts. These findings suggest that USAID and other donors should consider prioritizing rural road improvements in areas with several potential access points to a larger road network.

This impact evaluation of the INP II reflects a broader change in the evaluation landscape. Over the last 20 years, RCTs have dramatically expanded the body of evidence about which types of development programs work, when, and why, but their application has been heavily concentrated in a few sectors. Eighty-three percent of the trials in 3ie's worldwide repository are focused on health, population, and nutrition programs. The fact that rigorous impact evaluations are rarely undertaken in other sectors has stoked demand for a broader toolkit when it is impractical, unethical, or otherwise undesirable to randomize assignment into a development program. GIEs are one way to reach these under-evaluated sectors. They have also opened up new opportunities to rigorously evaluate programs in fragile states where the collection of baseline and endline data is particularly challenging and costly. GIEs are now being used to estimate the effects of municipal governance programs in Colombia and Niger, land tenure programs in Brazil and Ecuador, irrigation programs in Afghanistan and Ethiopia, road programs in Tanzania and Cambodia, and even multi-country project portfolios that seek to reduce deforestation, slow land degradation, and promote biodiversity conservation.

Several groups within USAID—including the Global Development Lab and the Center of Excellence on Democracy, Human Rights and Governance—are now seeking to accelerate this trend and expand the reach of impact evaluation to new sectors and programmatic settings. Other development finance institutions are also moving in this direction, including the Millennium Challenge Corporation, the World Bank, the German Development Bank, the Global Environment Facility, and the Green Climate Fund. The arrival of new tools like GeoQuery, which make it faster, cheaper, and easier for those without GIS expertise to access and use geospatial data on development programs and development outcomes, are helping to fuel this quiet revolution in impact evaluation.

Future Development

This blog was first launched in September 2013 by the World Bank in an effort to hold governments more accountable to poor people and offer solutions to the most prominent development challenges. Continuing this goal, Future Development was re-launched in January 2015 at brookings.edu.

For archived content, [visit worldbank.org](http://worldbank.org) »