Journal of Economic Perspectives—Volume 40, Number 4—Fall 2026—Pages 159–184

The Why and How of Foreign Aid

Dean Karlan

Foreign aid, whether privately or publicly funded and whether short-run humanitarian or long-run development, will not solve all problems yet receives many demands. How can resources be optimally deployed to achieve foreign aid goals and improve development outcomes? Consider two categories of questions for guidance: the why and the how.

The "why" includes: Why tackle a particular problem? Who benefits: the donor, the recipient, or the globe? To be in the national interest of the donor, is it sufficient to align with the will of the voters or must the aid benefit the donor via economic or geopolitical gains? Why address one problem at the opportunity cost of another? Policies must allocate resources across cause areas (health versus poverty), geographies (Guatemala versus Ghana), and time (immediate humanitarian needs versus reducing future suffering).

The "how" includes: How to tackle a particular problem? What is the market failure that underlies the problem, and what size of improvement can reasonably be expected? The "how" takes the "why" as given and drives the pragmatic decisions of whom to engage, what to fund, how to measure, how to decide all of the above, and how much impact to expect from proposed strategies. The "how" also circles back and informs the "why": If the "how" reveals a problem to be truly intractable, one

■ *Dean Karlan is the Frederic Esser Nemmers Distinguished Professor of Economics and Finance, Kellogg School of Management at Northwestern University, Evanston, Illinois. He served as Chief Economist, US Agency for International Development (USAID) from November 2022 to February 2025 and is the founder of Innovations for Poverty Action. His email address is dean.karlan@gmail.com.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.20251476.

may return to the "why" to redeploy resources toward a different, less intractable, problem.

In the language of economics, the "why" asks about allocative efficiency—given scarce resources and an abundance of problems, *cost-benefit* analysis guides choosing which problems to tackle. The "how" in contrast asks about productive efficiency—for a given problem, *cost-effectiveness* analysis guides choosing which strategy most efficiently addresses the problem identified in the "why" step.

Instability in the "why" can hinder development when progress requires long-term partnerships and continued investment. Motivations for publicly funded foreign aid swing with deliberate and informed rethinks, as well as erratic and unplanned winds (or gales, as experienced in 2025). While consensus on the "why" undoubtedly would yield stronger and more durable policy, disparate parties can often agree on a common plan despite different underlying motivations and incentives. For example, supporting a particular country's development effort could be motivated by donor self-interest in strengthening geopolitical power or moral impetus to reduce human suffering. Of course, when prescriptions differ—on the country in which to work, the problem to tackle, or the terms on which to act—then the political process must adjudicate.

The "how" drives whether the aid actually delivers on goals set by the "why." Evidence informing the "how" has improved radically in the past 30 years, with demonstrated impacts both within site (a program generates evidence that helps it iterate and improve) and via extrapolation (the upfront planning of a program is informed by evidence generated elsewhere). For the first, one World Bank (2025, p. 114) estimate of a "trial-and-adopt" iterative process concluded that program impact was often improved by over 50 percent from investing as little as 1 percent of the budget in testing. Returns to investment in evidence are even higher when lessons inform how resources elsewhere get allocated. The Massachusetts Institute of Technology Abdul Latif Jameel Poverty Action Lab and Innovations for Poverty Action report that evidence and policy advocacy from their networks have influenced programs and policies reaching over 850 million people (Dhaliwal 2025). For instance, school-based deworming programs that cost-effectively improve school attendance and future earnings have scaled to reach more than 280 million children (J-PAL 2022a); shifting from in-kind food distribution to electronic vouchers reached more than 18 million eligible households and reduced poverty in Indonesia at a lower cost to the government (J-PAL 2026); and an emissions market reduced air pollution at lower cost to firms and improved air quality for 15 million people in Gujarat, India (J-PAL 2025a). All were supported at some point in their design, testing, or scaling by foreign aid.

Beyond the direct benefits of improving how aid itself is used, three downstream benefits follow from improved evidence on what works and why. First, using aid to improve how low- and middle-income countries use *their own* resources can generate tremendous leverage, in both scale and durability of impact. If $10 million in technical support and funding of innovation to a $500 million ministry of health improves its effectiveness by 10 percent, that likely dwarfs what could have been

achieved by spending the $10 million on direct health services. Second, more effective programs could garner stronger political support, leading to a more stable and, thus, efficient system. Third, lower risk, all else equal, attracts more investment in addressing global problems.

Figure 1 illustrates the "why" and the "how" and their component steps as a decision-making flow chart. The chart inherently oversimplifies, skipping countless nuances and factors, but aims to capture several of the critical questions and how they influence key decisions such as when the "how" feeds back into the "why" and when to use existing evidence versus generate new evidence of effectiveness.

The Why of Foreign Aid: Multiple Motivations, Common Ground

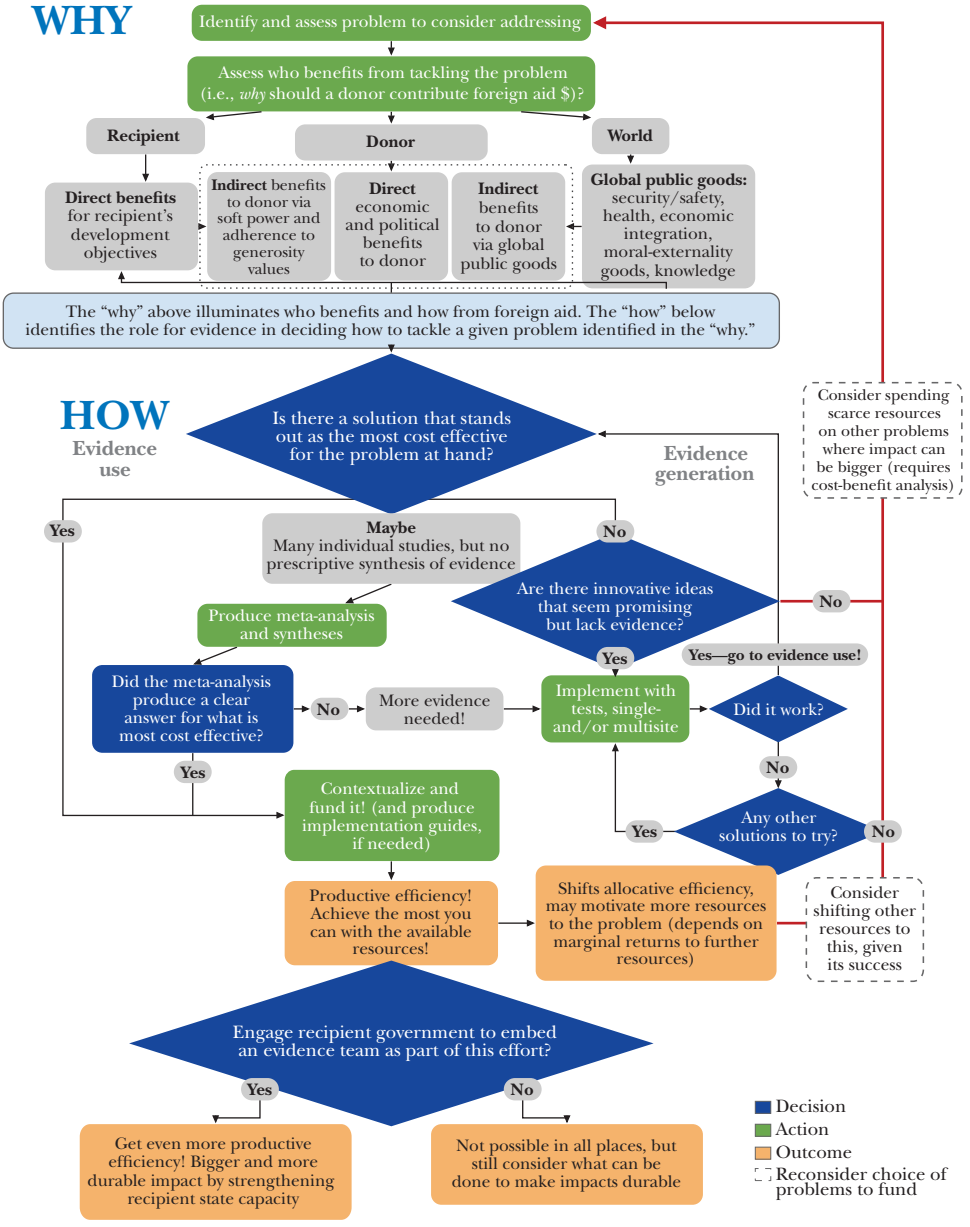
Most of the problems that aid seeks to address require long-term collaboration among many stakeholders. The challenge is to establish a "why" that is meaningful, logically coherent, and empirically supportable, yet broad enough to maintain strength and consistency through natural political ebbs and flows. This is not completely independent of the "how"; a logically coherent and empirically supportable "why" will consider the feasibility of tackling a particular problem.

Strikingly, donor-country political rhetoric consistently (particularly recently) invokes "national interest" as the primary justification for aid, yet leaves its definition vague. American officials have demanded that foreign aid make America safer, stronger, and more prosperous (Rubio 2025), while Germany's national security strategy states that development cooperation is both "in the interest of partner countries and also important for security in Germany" (Federal Ministry for Economic Cooperation and Development 2025). The United Kingdom's then-Secretary of State for International Development Andrew Mitchell (2011), in the same breath, called aid "morally right, it is part of what defines us as the British people… but it is also very much in our national interest."[1] This cross-national chorus reveals that "national interest" has become a phrase capacious enough to justify spending more as well as less. This blurring may not be mere rhetorical slippage, but rather a by-product of the actual overlap in the policies implied by these otherwise disparate motivations.

Categorizing the benefits of foreign aid into three types helps crystallize this conversation: benefits that accrue to (1) the donor, (2) the recipient, and (3) all via global public goods. Each dollar of foreign aid can serve multiple purposes, and the sum and the interactions across these benefits, weighted by the preferences of the

[1] Mitchell's (2011) framing echoes a much older domestic debate in Britain over whether state expenditure, rather than private charitable giving, better represents the authentic moral will of the citizenry. By the mid-twentieth century that debate had largely been settled in favor of the state: in a 1948 opinion poll, over 90 percent of Britons felt there was no longer any role for private charity (Breeze 2005). Prochaska (1988) traces this unfolding across the preceding century, as state provision slowly replaced Victorian-era voluntary philanthropy.

Figure 1
The Why and How of Foreign Aid



Source: Author's rendering.
Note: A simplified decisionmaking flowchart illustrating the component steps of the "why" and "how" of foreign aid.

decision-maker, drive foreign aid decision-making whether implicitly or explicitly (for a model illuminating this point, see Bermeo 2018, chapter 3).

Regardless of the "why," each path is strengthened when the aid itself is more effective at creating benefits for the recipients. A more meaningful gift is more appreciated and more likely to yield reciprocity. For example, effective service provision reduced local insurgent violence against the Iraqi government and Coalition forces in the Iraq war of 2003–2011 (Berman, Shapiro, and Felter 2011). Public opinion polls in 45 countries found that the United States' favorability improved after the launch of the President's Emergency Plan for AIDS Relief and the President's Malaria Initiative, visible and attributable programs found effective at achieving key outcomes (Jakubowski, Mai, Asch, and Bendavid 2019).

Direct Benefits: Economic and Political Returns to Donor Countries

The most direct form of donor benefit is an explicit quid pro quo: buying donor-country products to distribute in recipient countries or facilitating (through subsidy or technical support) trade and investment channels for donor-country firms to the recipient country. For example, the US Agency for International Development INVEST program, which provided support such as transaction advisory services and investment platforms, claimed to have mobilized $1.6 billion in private sector capital for development projects (Brown 2024). Claims of that sort—tallying up investments touched in some way by the consultants paid for by the aid contract—are commonplace in political accountability rhetoric, yet should not be confused as evidence of impact.

Beyond direct trade and investment, foreign aid serves what Bermeo (2017) terms "targeted development:" attempts by donor countries to limit negative spillovers from less-developed countries. Donor nations face challenges that cannot be optimally addressed through domestic spending alone: terrorism, unwanted migration, disease, trafficking in persons and illicit substances, and more. Here, the "quid" is an outcome in the recipient's national interest (for example, airport screening equipment) and the "quo" is an outcome in the donor's interest (for example, reduced terrorism risk). Bermeo shows that in the post-2001 era aid has become more targeted, increasingly aimed at achieving specific domestic priorities through spending abroad.

A direct quid-pro-quo aid effort can serve a domestic redistributive goal. For example, the US Food for Peace Title II program uses tax dollars to purchase American crops, ships those commodities largely via US-flagged vessels, and distributes them in countries experiencing food insecurity.[2] Foreign aid is likely an inefficient approach to domestic redistribution; a program that bought the same food and distributed it domestically would mechanically yield greater domestic benefits. Still, this form of aid could be one "why" behind foreign aid budgets and certainly is more salient and observable to the public than other benefits.

[2] The explicit aims were to "make maximum efficient use of surplus agricultural commodities," "stimulate and facilitate the expansion of foreign trade in agricultural commodities produced in the United States," and transfer "to any nation friendly to the United States in order to meet famine or other urgent relief requirements of such nation" (Agricultural Trade Development and Assistance Act, Pub. L. Nos. 83–480, 1954). When this aid was cut in 2025, many wrote about the potential negative impact on American farmers (Beras and Wong 2025).

Direct benefits need not be tied to the aid itself. In what is best described as hard power, the quid and the quo are immediate, specified, and likely coerced. The United States has often explicitly traded economic assistance for military advantage. For example, the United States forgave $7 billion in Egyptian debt as a reward for that country's participation in the Gulf War of 1990–1991 (Farnsworth 1991). In 2001, the United States waived sanctions on Pakistan and restarted economic assistance as a reward for cooperation in the war on terrorism (Wagner 2001). In 2009, the president of Yemen gave unrestricted permission for US unilateral strikes within the country against al-Qaeda; in 2010, security assistance to Yemen nearly tripled (Booth and Black 2010).

Indirect Benefits: Soft Power Returns to the Donor

The term "soft power" was coined by Joseph Nye, Jr. (1990, 2004) and defined as "the ability to get what you want through attraction rather than coercion or payment—magnets rather than carrots or sticks." A donor country may provide foreign aid expecting benefits that are not explicitly linked to the aid, such as an increase in its influence in global institutions or country-level engagement on specific markets or security matters. A cooperative equilibrium can arise, as in a repeated game in which reciprocity can be sustained without tying each individual act to a specific return.

The soft power rationale for foreign assistance has historical roots. For the United Kingdom, Hilton (2026) traces how humanitarianism became the primary vehicle through which Britain redefined its global role after decolonization, through both private philanthropy and nonprofits such as Oxfam, Save the Children, and Christian Aid, and through public funding. In the United States, President George W. Bush (2002) argued that "we fight against poverty because hope is an answer to terror," a formulation that includes a "targeted" pathway from aid to reduced poverty to greater enfranchisement to less terrorism, and a soft power pathway of increased influence and cooperation internationally.

While "soft power" is descriptively powerful, it can be analytically porous. The very breadth that makes "soft power" politically durable also makes it strategically incoherent, as it encompasses everything, yet commits to nothing specific. This porousness then provides political ammunition for aid expansions and contractions alike. The academic literature likewise can provide fodder either way, because soft power benefits are intrinsically indirect and diffuse and thus more vulnerable to both identification and generalizability challenges. This vulnerability is compounded by political pressure from both ends of the spectrum to bias findings in a desired direction, creating conditions in which "policy-based evidence-making" rules the discourse rather than evidence-based policymaking. Anecdotes, whether causal or not, can almost always be conjured up to make the case for or against foreign aid.

Altruistic Benefits: Generosity as Its Own Reward

Advanced democracies were shaped by ethical commitments such as religious social thought, human rights, and cosmopolitanism that favored the poorest people

over narrow self-interest (Lumsdaine 1993). Laboratory experimental evidence finds that giving itself generates happiness (Aknin, Mayraz, and Helliwell 2017). If the people of a donor country are motivated by some form of altruism, whether derived from philosophical principles or religious values, then a government engaging in humanitarian and development support is acting in its national interest, even if the aid generates no economic returns to the donor country.

Politicians of donor nations often speak of their nations as generous due to their great moral fiber; if so, aid for humanitarian reasons could strengthen community and national cohesion. The generosity motive thus yields returns for the donor nation by satisfying its citizens' stated desire to act altruistically and to be a nation that lives up to its own values (Yoho and Karlan 2025). Indeed, support for foreign aid in donor countries is generally high: for example, a 2014 Eurobarometer survey found that 73 percent of Europeans agreed that tackling poverty in developing countries is a moral obligation of the European Union (Eurobarometer 2015).

Global Public Goods: Benefits to Recipients, Donors, and the World

Under plausible assumptions, investing in global public goods can satisfy both private national interest and public good motivations. In particular, soft power gains are likely larger when the problem is global. If other donor countries follow suit, a cooperative equilibrium can emerge in which the public good is provided at a meaningful scale.

One set of foreign aid global public goods includes investments in reducing negative outcomes for others where transboundary threats pose potential economic or safety harm to donor countries too, such as climate change, contagious disease, forced displacement, war, and instability.[3] A second set comprises normative or moral-externality goods; the moral benefits of living in a world with less suffering are nonrival and nonexcludable. Examples include anti-human-trafficking programs and humanitarian responses to natural disasters. A third set encompasses knowledge public goods, which are typically nonrival and mostly nonexcludable, including scientific, technical, and policy knowledge on topics ranging from vaccines and agronomic innovations[4] to evidence on effective humanitarian and development policy.

This third category provides the conceptual bridge to the "how" of foreign aid. Development knowledge is nonrival: A rigorous evaluation demonstrating that a cash transfer program raises household consumption in Kenya is insightful to policymakers in Bangladesh, Bolivia, and Bihar. Regardless of which "why" motivates a particular donor, investing in the evidence base for what works promotes both national and global interests. The "how" is, therefore, not merely a technical

[3]Examples of tracking and response systems include the World Health Organization influenza work (WHO 2026) and the Intergovernmental Authority on Development's work on East Africa conflict (Intergovernmental Authority on Development 2026).

[4]An example is the Green Revolution, a period during which a series of agricultural innovations dramatically increased cereal output in low- and middle-income countries. The Consultative Group on International Agricultural Research, established in 1971, received multilateral and bilateral funding to create global centers for agronomic research (Byerlee and Lynam 2020).

afterthought to the "why;" it is itself a contribution to the global public good of development knowledge.

Before turning to the "how," a few notes on the sequencing of and relationship between the "why" and the "how." Grand strategies on the "why" established by donor-country leadership (like the executive branch leader, cabinet, and legislature) are often set with little regard for the "how." However, a level down, the "why" can be more responsive to the "how." In particular, if evidence on the "how" suggests that the best approach has a low probability of success, those setting the "why" may reconsider how they prioritized one cause area over another, one country over another, or one tool (for example, development or diplomacy) over another. For example, in 2022, the US government launched a "Democracy Delivers" initiative, which aimed to help young democracies strengthen their markets and institutions, with two broadly stated purposes: provide a global public good of strengthening democratic countries and provide direct benefits to the United States by facilitating trade between those countries and US firms (Blinken 2022). Determining the specific countries and cause areas was left to operating-level decision-makers that vetted ideas and, thus, effectively influenced the "why" as well as carried out the "how." Similarly, the US global aid response to the COVID-19 pandemic was defined as a priority by the highest levels of government and aimed to deliver humanitarian benefits globally (for example, by addressing extreme poverty and food insecurity) as well as direct economic and health benefits to the United States (by mitigating the pandemic's spread and economic fallout), but resources were allocated and later "surged" to countries according to criteria that prioritized need, capacity for success, and existing relationships (Dahl et al. 2024).

The How of Foreign Aid: The Key Role for Evidence

The "how" of foreign aid with respect to evidence can be organized around three interrelated strategies. The first is using existing evidence to choose and design programs, drawing on the accumulated body of rigorous impact evaluations to select interventions with a strong track record and adapt them to local context. The second, and most transformative, is scaling evidence-based policies through recipient governments or firms in low- and middle-income countries and strengthening their capacity to generate and apply rigorous evidence—not merely advising on one program at a time, but building the systems, culture, and institutional infrastructure for evidence-informed policymaking. The third is generating new evidence through single-site evaluations of innovations, coordinated multisite trials, or large does-it-work-at-scale studies. This step includes both impact assessments and finely tuned implementation research. These strategies intertwine: A strong evidence-use process not only identifies what evidence can and should be applied for scale, but also identifies the evidence that would reduce knowledge gaps and better inform future decision-making.

Two omissions are particularly worth noting. First, I omit many critical "how" questions that are outside the (narrowly construed) boundaries of economics,

such as how to legislate development cooperation, how to procure, how to choose funding recipients, how to structure bilateral versus multilateral engagement, and how to avoid fragmented bureaucracies (for three examples of many, see Burnside and Dollar 2000; Easterly, Levine, and Roodman 2004; Carcelli 2023). Second, the focus on cost-effectiveness evidence does not imply foreign aid should focus exclusively on programs that can be evaluated using experimental or quasi-experimental methods. Such methodologies have advantages when studying households, schools, villages, or firms, but are less applicable to many important development questions (for discussion, see Ravallion 2009). These include aid that operates at the country level, where credible comparison groups are difficult to construct; interventions whose effects unfold over decades; budget support or policy dialogue, where the "treatment" is diffuse and the counterfactual speculative; and many (but not all) governance and institutional reforms. This does not mean such aid cannot be evaluated: Natural experiments, modeling exercises, and qualitative case studies have roles to play. The leverage of such efforts can be huge in terms of resources eventually moved or efficiencies created through country-level or global reforms (for more on the role of state capacity, see Honig 2018; Muralidharan 2024). However, donor capacity is all the more important when undertaking aid efforts that are harder to design and measure.

The distinction between cost-benefit analysis and cost-effectiveness analysis is central here. Cost-benefit analysis facilitates adjudication of the "why": it weighs the expected monetized benefits of tackling one problem against another to determine where to allocate resources. Cost-effectiveness analysis guides the "how": Holding the problem constant, it asks which intervention, design, or delivery model produces the most impact per unit of expenditure. Two programs may generate similar impact per household in magnitude and outcome, but differ considerably in cost, thus rendering the cheaper program (all else equal) more cost-effective. The two are complementary but distinct. A program can be highly cost-effective yet still fail a cost-benefit test by solving (albeit efficiently) the less important problem. The "why" debate, adjudicated in part by cost-benefit reasoning, sets the agenda; the "how," guided by cost-effectiveness analysis, determines whether that agenda is actually delivered.

Cost-effectiveness analysis carries a practical advantage within the political economy of aid institutions. Foreign aid bureaucracies, like many large organizations, comprise constituencies loyal to specific problems such as maternal health, food security, climate adaptation, education, and others. Cost-benefit analysis, by asking whether a problem is worth tackling at all relative to alternatives, implicitly threatens those constituencies' raison d'être. Cost-effectiveness analysis is far less threatening: It accepts a group's goals and asks how to do the most good toward those goals with the resources in hand. Rather than asking a team to defend the very existence of their effort, cost-effectiveness evidence empowers them to achieve more with their resources. For example, a global panel of education experts ranked education interventions according to their relative cost-effectiveness at improving learning, giving donors and policymakers direction for what to implement (evidence use) as well as what to evaluate (evidence generation) (Global Education Evidence

Advisory Panel 2023). Thus, cost-effectiveness evidence is typically a more politically and bureaucratically viable path for improving aid. It also avoids the uncomfortable value-laden challenge of monetizing benefits, which economists embrace too readily as obvious and noneconomists decry as amoral (yet do so implicitly).

Building From the Evidence Base to Implementation at Scale

More than 20 years ago in this journal, Easterly (2003) asked whether foreign aid can buy growth and called for more evaluation and better incentives for both funders and recipients. Ten years later, Levine and Savedoff (2015) argued that building knowledge collectively is the future of aid. This period saw a large acceleration in the use of randomized impact evaluations, which was recognized in 2019 with a Nobel Prize in Economics awarded to Abhijit Banerjee, Esther Duflo, and Michael Kremer for their intellectual leadership in building this movement. A randomized evaluation, in its simplest form, randomizes the assignment process of an intervention or policy. The random process, when done with sufficient sample size such that the law of large numbers applies, creates treatment and control groups such that differences between them can be attributed causally to the intervention or policy, and not to outside factors, whether observable (such as weather or macroeconomic conditions) or difficult or impossible to observe (such as individual capabilities or preferences, which typically then could generate a selection bias). According to the International Initiative for Impact Evaluation website, there are now more than 8,000 randomized evaluations in development economics that shed light on the effectiveness of a variety of interventions and help to build external validity, where evaluations successfully identify mechanisms for why interventions work and where multiple evaluations can test for consistency with such theory across multiple contexts.

These efforts have enabled the development and spread of effective public policies reaching hundreds of millions of people. This plethora of evidence motivated the creation of the US Agency for International Development Office of the Chief Economist in 2023 (henceforth, OCE—an office I built and led from 2022 until resigning at the end of February 2025 when it, along with the rest of USAID, was dismantled). The OCE pushed for reform specifically to "use it or produce it"—that is, to either use evidence of cost-effectiveness to decide what to fund, or if the evidence was not deemed sufficient to be fully instructive, then to produce more evidence to better inform future decisions (USAID 2024).

"Use it or produce it" is the central fork of Figure 1: The question "Is there a solution that stands out as the most cost-effective for the problem at hand?" sends a "yes" down the evidence-use column on the left or a "no" down the evidence-generation column on the right, with a third, common, answer of "maybe" in the middle, when individual studies may exist but we lack sufficient evidence (and accompanying syntheses and aggregation) to confidently inform policymakers in a given context whether that solution is the best approach and how to adapt it.

The issue is not merely the quantity of evidence. Having an abundance of randomized evaluations is insufficient for tackling the obvious and necessary question: how does one apply evidence across space and time? How do we address

the "generalizability puzzle" of learning not merely "intervention X worked in context Y," but what about context Y was necessary to generate that result? Bates and Glennerster (2017) offer a framework for moving beyond vague "it depends, context matters" claims by instead clarifying which aspects of context matter for specific interventions so that the discussion does not stop with overly simplistic assertions that "country X is not country Y and, thus, nothing done in X can inform country Y policy." Ultimately, the goal is to combine theory that sheds light on when and how programs can work with data to test those theories (Deaton 2010).

Some new and creative policy innovations originate from nonprofits, some from governments of recipient countries, and some from academic researchers. Whatever the origin, testing and refining these innovations through randomized evaluations can shift the policy landscape. Here are three examples of programs that proceeded from local studies and pilot programs to a broad set of replicated studies and adaptations and adoptions at scale.

Conditional cash transfers, in which transfers are conditioned on school attendance, vaccinations, or health checkups, offer a first instructive example. In Mexico in the 1990s, and later in many other countries, governments worked with researchers to conduct randomized evaluations that consistently demonstrated that well-designed and contextualized conditional cash transfer programs can have positive impacts on health, food security, school attendance, consumption, and income. Subsequent research has learned that unconditional cash transfers have a similarly robust impact on a plethora of primary development objectives (Crosta et al. 2026). These results inspired governments around the world to create and test their own programs, both conditional and unconditional. As of 2016, over 60 governments worldwide have implemented a conditional cash transfer program and 130 have implemented unconditional cash transfers (Bastagli et al. 2016). Hundreds of rigorous evaluations have measured their impacts and helped governments learn what works to optimize program design and improve outcomes (Marliani and Pedro 2025). Recent evidence has gone further to study the macroeconomic effects of unconditional cash transfers, finding a macromultiplier effect of 2.5 (Egger et al. 2022), and efforts are now underway to expand and replicate this study in Malawi.

An education-related success story called Teaching at the Right Level (TaRL) offers a second example. The program builds foundational skills in math and reading by grouping children by learning levels, rather than teaching an identical curriculum to everyone at a given age or grade level. Pratham, one of India's largest education nongovernment organizations, developed TaRL in the early 2000s. Early proof-of-concept evaluations showed its effectiveness, and Pratham and others then tested iterations to understand the effectiveness of different delivery models (for example, implementation by village volunteers, nongovernment organization instructors, or government teachers). With strong repeated evidence of impact (Banerjee, Cole, Duflo, and Linden 2007; Angrist and Meager 2023), the program has now garnered attention from ministries of education for implementation at scale, with several related organizations (Pratham, Abdul Latif Jameel Poverty Action Lab, Youth Impact, and TaRL Africa) now providing technical support to the scaling entities (J-PAL 2022b).

A third example is the "Graduation" social protection approach, originally termed the "Graduation" model by the Bangladeshi nonprofit BRAC. The approach deploys a multifaceted livelihoods program that combines five complementary interventions (or some subset thereof), including a productive asset, training, coaching, savings access, and short-term consumption support, to provide a substantial push toward more secure and sustainable livelihoods for those living in extreme poverty. In 2017, the state government of Bihar in India began piloting a program to adopt the approach, to be implemented by a state agency with technical assistance from a nongovernment organization. The Bihar government subsequently announced plans to expand the Graduation approach in its "Satat Jeevikoparjan Yojana" (Sustainable Livelihoods Program), with the goal of extending the state's previous economic inclusion efforts to 200,000 of the poorest households across Bihar between 2018 and 2023, the first large-scale government adaptation of the Graduation approach (J-PAL 2025b).

Evidence from over 40 randomized controlled trials has shown that the Graduation approach for the very poor can increase income, consumption, productive assets, and other noneconomic outcomes such as mental health and community participation (for the initial randomized evaluations, see Banerjee et al. 2015; Bandiera et al. 2017; for syntheses, see Carter, Jee, Karlan, and Premand 2026; J-PAL 2025b). Ongoing evaluations are testing more specific implementation and design questions (such as which components are needed, how best to deliver the coaching and training, what the optimal bundle of transfers is, and so on). At present, BRAC and other technical service providers, typically nonprofit organizations with experience implementing the model, are working with governments in Ethiopia, India, Indonesia, the Philippines, Rwanda, and South Africa to integrate the model into existing systems at scale to achieve greater, more lasting impact. As of 2024, over 400 economic inclusion programs influenced by the Graduation approach have reached more than 15 million households (70 million people) globally (Partnership for Economic Inclusion 2024).[5]

Not all programs that demonstrate meaningful effects in pilots or in smaller-scale implementation succeed when implemented at scale. Banerjee et al. (2017) highlight examples of programs shown effective in a pilot that failed to deliver similar results at scale. Among these, a program that provided grants to hire extra teachers and reduce class sizes in Kenya raised student test scores when implemented by two different nongovernment organizations, but failed to do so when implemented at scale by the national government. The authors attributed this failure to weaker implementation capacity and adverse political circumstances. While some programs fail to demonstrate predicted impact at scale, others demonstrate greater impact via general equilibrium effects (for example, Egger et al. 2022; Muralidharan, Niehaus,

[5]The global health world is replete with almost countless examples of massive population health improvements stemming from evidence generated from randomized controlled trials. For a list of examples, the Center for Global Development has curated global health success stories at https://millionssaved.cgdev.org/.

and Sukhtankar 2023). The question "does it work at scale," in a sense, however, is a poorly posed question, as it bundles together disparate theoretical issues: organizational management and employee incentives, price effects, and political economy dynamics, among others.

In short, foreign aid can be spent to provide direct services, and there can be benefits from doing so, but when aid is combined with generation of new evidence, it can also influence future deployment of resources, whether public or private, in the recipient country or elsewhere. Generating knowledge—a public good—that identifies cost-effective approaches can further influence how other countries and donors deploy their resources, multiplying impact per dollar of aid and domestic resources alike (Karlan 2015).

Evidence Use and Generation: Strengthening Capacity of Aid Recipients with Evidence Partnerships

Foreign aid budgets may seem large, but are small relative to the budgets of the governments of low- and middle-income countries. Although governments want to implement effective policies, competing demands and limited resources hamper efforts to establish formal, systematic processes for applying and producing evidence. Governments need data to target, manage operations, and evaluate effectiveness of policies and programs, and on many other granular implementation details (akin to the "plumber" analogy in Duflo 2017). However, institutional frameworks that enable experimentation with program designs, pilot testing of new policy approaches, and integration of findings into decision-making are rare.

Formal collaborations with organizations committed to bridging the evidence-to-policy gap can help develop systems and processes that capitalize on this opportunity. The objective is not merely to partner on one research project or scale-up, but rather to change the culture of policymaking so that rigorous evidence becomes one of the key factors considered when choosing programs or policies. One approach is to embed teams dedicated to helping governments use and produce evidence directly within government agencies. Beyond implementing a single project, an embedded evidence lab identifies policy challenges, develops and implements strategies to generate relevant data and evidence, and facilitates the application of these insights. A lab can also strengthen an agency's ability to generate and use evidence by upgrading information systems for research and learning, building officials' capacity to demand and apply evidence, adapting institutional processes and legal frameworks to embed evidence-use practices, and securing the political and financial support needed for long-term sustainability. Here are three examples.

First, the Central Bank of Nigeria's Consumer Protection Embedded Evidence Lab was established in 2021 in partnership with Innovations for Poverty Action.[6] The lab sought to help the Central Bank of Nigeria integrate evidence-based decision-making into its mandate to safeguard financial consumers, who faced

[6]For an overview, see https://poverty-action.org/consumer-protection-nigeria-embedded-evidence-lab.

increasing risks as digital financial services expanded across the country. Nigeria's central bank grappled with a number of challenges in enforcing fair, transparent, and responsible conduct by financial institutions, including siloed consumer and financial data, lack of actionable analysis, and a rapidly evolving financial servicessector.

To address these challenges, the lab focused on improving the ability of the Central Bank of Nigeria to implement evidence-based regulatory policy through a four-pronged approach: (1) build capacity to apply data analysis tools in regulatory work; (2) upgrade administrative data systems to improve quality and reliability so that regulatory decisions are based on accurate and comprehensive information; (3) prioritize monitoring, evaluation, and learning to support decision-making and align with the bank's strategic objectives; and (4) create decision-oriented data products, including market monitoring reports, consumer complaint dashboards and institutional performance assessments. The collaboration improved the capacity of Nigeria's central bank to produce high-quality data, analytical tools, and reports for regulatory decision-making, all key to enhancing its oversight and accountability capabilities. The lab has become an integral part of the strategy of the Central Bank of Nigeria, contributing to multiple policy reforms and building a data-oriented culture.

Second, Peru's Environmental Evaluation and Enforcement Agency established an embedded evidence lab in 2024 to embed systematic data collection, rigorous impact evaluation, and iterative innovation into its regulatory and enforcement functions.[7] As the national authority responsible for balancing environmental protection with private investment, this agency seeks to deploy effective enforcement strategies consistently informed by empirical evidence. The lab uses a five-stage learning cycle: identify policy challenges, design and test solutions, analyze results, engage stakeholders to interpret results, and incorporate lessons into scaled-up approaches. Current work includes using satellite imagery to strengthen municipal compliance with waste management practices, testing interventions to increase use of a citizen reporting app, and developing a WhatsApp chatbot to boost adoption of an environmental emergency reporting tool.

Third, to improve the quality of learning in Rwanda, Innovations for Poverty Action and the Georgetown University Initiative on Innovation, Development and Evaluation partnered with Rwanda's Ministry of Education to conduct a randomized evaluation of a pay-for-performance teacher contracts program, called Supporting Teacher Achievement in Rwandan Schools (STARS), that led to improvements in both teacher effort and student learning (Leaver, Ozier, Serneels, and Zeitlin 2021). Rwanda already ranks highly among sub-Saharan African countries for education access; in 2015, net enrollment in primary education was 97 percent. However, improving education quality, reducing high rates of teacher turnover, and attracting teachers to rural areas remain critical challenges for the country's education sector.

[7]Established by presidential resolution (Resolution 00018-2025-OEFA/PCD 2025), the lab is formally housed within a unit responsible for developing innovative regulatory strategies and staffed by four government officials trained through structured capacity-building and learning-by-doing.

To build on these results, in 2019 the Ministry of Education established an embedded evidence lab to support the scale-up of the pay-for-performance approach and to iteratively test model refinements. With five full-time staff spread across the implementing agencies, the lab has direct access to policymaking processes, data systems, and implementation channels. The scale-up of the Supporting Teacher Achievement in Rwandan Schools program began in 248 schools, expanded to more than 340 schools, and is poised for a national scale-up covering 66,000 teachers and about 3.2 million students. The program includes strengthening data systems, training inspectors and headteachers, and embedding contract management within existing government structures to ensure sustainability. Early experimental results have shown positive treatment effects on teachers' lesson preparation, pedagogy, and consistent reporting of data to a ministry platform.

Scaling the Supporting Teacher Achievement in Rwandan Schools program took years, and securing political support was essential to the lab's sustainability. The lab maintained momentum by identifying small projects that contributed to the scaling effort while demonstrating its usefulness in real time—for instance, using testing data systems built for the randomized evaluation to inform short-term school management decisions on grade repetition and remedial tutoring. Beyond STARS, the lab increased demand for evidence more broadly and influenced institutional restructuring.

Evidence Generation: Innovation Funds

If aid effectiveness varies, and if innovation and evidence can improve it, there is a strong role for rigorous social science: generating evidence as a global public good, identifying cost-effective innovations, and further refining our understanding of effective approaches. This is the right-hand column of Figure 1: Promising ideas that lack evidence get implemented with tests, and the results either feed an implementation guide or send the decision-maker back to look for alternatives.

Development Innovation Ventures (DIV) at the US Agency for International Development demonstrated the potential for high returns from investing in development innovations as global public goods. The approach was broad, defining innovation to include new applications of technology as well as novel business models, delivery methods, production processes, products, or services that were expected to improve development outcomes. DIV also supported applied research and replications of proven innovations in novel environments and used a tiered model to fund three stages of innovation: piloting, testing for impact and cost-effectiveness, and transitioning innovations with rigorous evidence of impact and cost-effectiveness to scale. Like venture capital, most individual investments do not succeed, but a few high-impact innovations can generate exceptional returns that lift the entire portfolio. Kremer, Gallant, Rostapshova, and Thomas (2021) estimates a social benefit-cost ratio of at least 17:1 for the DIV portfolio between 2010 and 2019. The figure is a lower bound by construction: It compares the benefits of a subset of high-reach innovations against the cost of the entire portfolio, including everything that failed.

France's Fund for Innovation in Development, established in 2021, similarly employs a tiered funding model to support prototyping, rigorously evaluating already-prototyped ideas, and scaling those with strong evidence. So far, it has funded 98 projects in 30 countries, amounting to €50.1 million (Fund for Innovation in Development 2025).

Evidence Generation: Going Local, Globally, with Multisite Adaptations and Evaluations

Multisite impact evaluations—which evaluate interventions in multiple contexts with a common strategy—are another public good that foreign aid donors are uniquely positioned to fund. The Graduation approach was subject to a six-country multisite evaluation funded primarily by the Ford Foundation (Banerjee et al. 2015) and involved global coordination with some local adaptations (notable variations included the asset types and savings promotion methods). Multisite evaluations lay a foundation for a "built-in meta-analysis" that yields stronger lessons if coordination reduces unnecessary variation in design and measurement across sites, variation that would otherwise hinder interpreting cross-site treatment effect differences. The multisite design also allows for stronger inferences by expanding the context and populations for which the program is evaluated. Relative to uncoordinated studies, a multisite study can also reduce fixed costs, because sites can be added without having to duplicate all setup costs. Foreign aid donors, which typically fund similar programs in a diverse set of countries and often have in-house research teams, are well placed to identify, coordinate, and fund such studies and to use the results. For example, the World Food Programme and World Bank have conducted a number of multisite evaluations funded through thematic impact evaluation "windows"; among others, these include a four-country evaluation of cash transfers and gender impacts (DIME-WFP 2024) and a four-country study of resilience programs (World Food Programme 2019).

Multisite studies are not without challenges, including adapting interventions and identification and measurement strategies to each context and maintaining implementation fidelity to ensure comparability and accurate measurement. Still, when done well, this type of impact evaluation is a public good that holds significant potential to improve the effectiveness of foreign aid.

The evidence strategies described above—using existing knowledge, building government capacity, generating new rigorous evaluations—are tools for planning and allocating future resources more wisely. Often, budgets for evaluation are set by rough guidelines on a proportion of a program cost, but that is a mistake. Budgets for evaluation (and evidence generation more broadly) should be set by considering the size of the *future* budget that could be more informed and the likelihood that new evidence will inform that decision. Evaluation (and evidence generation more broadly) budgets should be larger when future budgets are larger, larger when evidence gaps are larger, smaller when evidence generation investments are less likely to shrink the knowledge gap, and smaller when political economy or other constraints are more likely to render the evidence unused.

Closing the Academic-Policy Gap: What Researchers Can Do Differently

The academic-policy gap regarding foreign aid is smaller than often claimed, but still too large. The thousands of randomized evaluations conducted in development over the past two decades have identified interventions with strong track records, and that evidence has been applied to scale programs and policies within and across countries—with cash transfers and Teaching at the Right Level programs offering two clear examples. Even so, for many topics, policymakers do not see evidence from elsewhere as prescriptive or helpful locally, and mismatched timelines and incentives are not always conducive to academics producing the outputs most useful to decision-makers. Two gaps are particularly consequential: insufficient synthesis and lack of practical contextualization guides.

The Synthesis Gap

While academics have produced a vast body of impact evidence, much of the knowledge remains fragmented. Systematic reviews and meta-analyses can steer funders toward more effective and efficient approaches by aggregating otherwise isolated evaluation findings into more actionable insights about what interventions and designs are most effective in which contexts. Importantly, evidence syntheses help identify knowledge gaps, which should then guide future allocation of resources to evidence generation and help policymakers know which programs and policies should undergo impact evaluations and which should only be monitored for implementation fidelity. In Figure 1, synthesis is one answer to the "maybe" branch, the common situation of many individual studies and no prescriptive synthesis, and whether the meta-analysis yields a clear answer on cost-effectiveness is the next question asked.

Despite substantial inroads in data sharing, continued effort is needed to close the synthesis gap. Aggregation has its challenges, such as requiring harmonization that may obscure key differences across studies (for example, testing the effect of business training by including a single indicator variable for "taught marketing"). Harmonization is difficult on three levels: outcome measurement, context descriptions, and program features. Selection bias in where and when programs are implemented, studied, and published creates further barriers. Even when data are pooled, careful interpretation is needed to understand whether, why, and how interventions work. More consistent reporting and harmonization of results, data, and measurement practices would facilitate the synthesis needed to identify approaches that consistently deliver results and make it easy for policymakers to take up and scale those "good buys."

Data about costs have improved considerably over the past 20 years, but important gaps remain. Often, evaluation papers focus on the treatment effects and underlying mechanisms for those effects, then render cost reporting to a simple per-recipient figure. An effort to report more granular cost data will help, but this too misses a key point: What matters most are the quantities of inputs and outputs of a program or intervention, far more than the cost of those inputs and outputs. For any policymaker who wants to decide what to do in *their* country, the prices

faced by the evaluated implementer in some other part of the world are far less relevant. Instead, the policymaker needs to know how many employees they will need to hire at what experience and background level in order to provide the desired program to the population they aim to serve. The policymaker can then use the wages in their market to determine the cost of implementing different options. Thus, academic evaluations should get better at reporting the production functions of programs in terms of the specific quantities of inputs and outputs, even more than the costs. Both, of course, would be ideal, so that more research can be done to understand (and thus improve) the production function for delivering products and services as part of public policies. Dioptra,[8] a consortium led by the International Rescue Committee, has made important inroads on this by building a platform for nine large international implementers to share detailed input, output, and price data. These data allow better modeling of production functions (to facilitate planning and cost-effectiveness analyses) and critical operational research that guides implementers on how to lower the cost of delivering the outputs to households, communities, or markets, holding quantity and quality fixed.

Furthermore, creating databases of study results would bridge a key gap in moving from evidence generation to aggregation, translation, and adaptation. Such efforts also would benefit the research process; for example, studies could include smaller control groups if the plethora of data from prior studies provided sufficient information to partially model the counterfactual. A recent collaboration between Northwestern University, the University of California, Berkeley, and Innovations for Poverty Action has launched "DEVIDENCE"[9]—a publicly available database of evidence from causal impact studies in international development—to provide structured and harmonized data on results from randomized evaluations to support evidence synthesis and meta-analyses.

Artificial intelligence (AI) can revolutionize the aggregation, harmonization, and synthesis of evidence, but at the same time, AI tools that deliver black box answers create risks. This guidance for using AI is not particular to policymakers, but also applies to the design of programs and interventions. For example, one chatbot tested in Kenya worked well for those with higher baseline business acumen, but created harm for others. Those with acumen had the judgement to know which advice to take, and those without did not (Otis et al. 2026). Used wisely, AI tools can radically lower cost and shorten time required to synthesize massive amounts of data and results. But if used without appropriate caution, poor policy choices may ensue. Policymakers can already pose questions to AI platforms such as "What is the best way to tackle such-and-such problem?" and AI will deliver an answer. Yet we do not have enough formal meta-analyses to serve as ground-truth data with which to then train and test whether the AI delivers reliable syntheses, nor have we curated sufficient underlying data from which to harmonize study results and generate bespoke quantitative AI-driven analyses and policy recommendations.

[8] For information about the Dioptra tool, see https://www.dioptratool.org/.
[9] For a description of DEVIDENCE, see https://devidence.org/.

Contextualization Gap and Implementation Guides

Adapting effective development programs to local contexts is crucial. Global evidence can point to cost-effective interventions, but failing to account for the local context when designing a program can be wasteful, or even harmful. For example, a financial planning board exercise for farmers in Zambia improved consumption smoothing during the hungry season: Farmers entered the hungry season with one additional month's food supply and achieved 9 percent higher crop revenue from their farms (Augenblick et al. 2026). However, when adapted to Malawi, the program needed to account for farmers' more diversified income streams as well as a scaling model that needed to work with groups rather than individuals to lower marginal costs (Center for Effective Global Action 2025).

Contextualization guides could also help alleviate an "untold" problem inherent in most academic papers. Papers report on the rationale for decisions that are central to the research question posed, like the specific elements that are randomized, varied, or tested. However, academic publications typically omit discussion of implementation decisions that were simply made and held constant: group sizes for interventions, participant recruitment methods, staff hiring processes, and other procedural choices. Such seemingly mundane decisions may influence the effectiveness of the program and its generalizability. Thus, a persistent gap arises between what is reported and what occurred, and between what is shared and what is usable by others. Academics could make it easier for policymakers to adopt evidence-based programs if implementation decisions were reported more thoroughly in accompanying documents.

Contextualization of evidence also requires balancing the "testing theories" imperative of research with less-obviously-theoretical research and/or merely descriptive policy needs. One way to address this is through foreign aid partnerships with local governments for embedded labs, described in the previous section. Breaking the cycle of single-experiment-to-single-paper studies would also enable more rapid-fire operational research to refine policy design. While this may be less rewarding academically than researching big theoretical questions, 100 relatively small and atheoretical tests, considered as a set, could reveal important policy and theoretical insights (for more on the case for operational research in development, see Angrist et al. 2026).

Taken together, closing the synthesis and contextualization gaps is not merely an academic exercise, but a public good in which philanthropists and foreign aid ought to invest (similar to the argument by Levine and Savedoff (2015) that building knowledge collectively is the future of aid). It is the mechanism by which all three strategies of the "how" reinforce one another: Existing evidence informs government action, government partnerships reveal gaps that motivate new evidence generation, and new evidence is most durable when produced by and for the institutions that will use it.

As one recent example of contextualization, the US Agency for International Development's Office of the Chief Economist incorporated guidance on implementing the Graduation approach into two $150 million resilience solicitations in

2024 (the year prior to USAID's dismantling). USAID had regularly made large, five-year awards for Resilience Food Security Activities. Historically, these USAID solicitations specified the outcome(s) to improve and asked applicants to propose interventions (often with some "illustrative examples" provided). Applicants were not asked to provide evidence of cost-effectiveness of their proposed interventions.[10] The lack of specificity, combined with a consensus approach to program design, led to a failure to face implicit-but-real tradeoffs, and thus produced what was pejoratively referred to as the "Christmas tree ornament" problem: Every constituency hung their favorite ornament on the program, whether it meshed well with the others or not.[11] At best, such lack of focus increases operational complexity and overhead costs; at worst, it reduces the resources allocated to the most effective components, hindering overall impact. The Office of the Chief Economist at USAID sought to combat this tendency by bringing forward the evidentiary base and steering program design either toward proven programs ("use it") or innovations alongside new evidence generation ("produce it"). With an embedded senior advisor on the design team, new solicitations became focused and evidence based, and the final plan would have larger impact and a broader reach (although all such awards were eventually cancelled).

As another example of contextualization, a large-scale randomized evaluation in Bangladesh helped identify an effective approach to increase mask-wearing during the COVID-19 pandemic: a community-level mask distribution and promotion program that tripled mask usage. The research team's support to implementing partners, including advising on procurement, locally contextualized communications materials, and a detailed implementation toolkit, enabled rapid scaling of the intervention to reach millions of people across South Asia. Abaluck et al. (2021) produced supplementary material for implementing a similar program, including contextualizing it for other settings. The materials clearly identify four key components of the program to normalize mask wearing and provide detailed guidance for procuring masks, preparing for fieldwork, and conducting surveillance (J-PAL 2021).[12]

This type of implementation and contextualization guidance facilitates rapid adoption and scaling of impactful interventions while also facilitating implementation fidelity and identifying key demand-driven knowledge gaps for further research to address.[13]

[10] As an example, a similarly sized program for Mozambique in 2022 issued such a solicitation (USAID 2021).

[11] Glennerster and Haria (2025) discusses the incentives that drive this complexity and how to address them to use aid budgets to best effect.

[12] For the full implementation toolkit, see https://drive.google.com/drive/folders/1QOcyQzxhXjPvrw1vUt4qVRT4Bisi6HZ.

[13] Another example comes from the What Works Hub for Global Education at https://www.wwhge.org/, with similar work on implementation guides for scaling structured pedagogy.

Conclusion

The "why" and "how" of foreign aid are too often treated as separate conversations—one belonging to diplomats and politicians, the other to program designers and evaluators. However, an ineffective "how" renders the "why" moot, and vice versa. The "why" and the "how" are also rarely as simple as they may sound: many statements are oddly broad or narrow depending on the unspoken fine print.

Of course, foreign aid is no magic wand for development. Aid can enable corruption, elite capture, and clientelism, which can spill over into overall governance (Svensson 2000). Governments may substitute aid for domestic revenue, weakening social contracts and undermining long-run capacity (Feyzioglu, Swaroop, and Zhu 1998; Benedek, Crivelli, Gupta, and Muthoora 2014). Large inflows of foreign currency from donors can appreciate the currency of the recipient country, similar to the way that the "resource curse" can appreciate the currency of commodity exporters, undermining other exports but also allowing for cheaper imports of often essential inputs such as fuel (Ahmad, Pentecost, and Stack 2023). Given the many factors that affect overall economic growth, it is difficult to make strong claims about how foreign aid affects recipient country macroeconomic growth rates (Rajan and Subramanian 2008). It is equally undeniable that many foreign aid efforts produced substantial benefits, with an abundance of evidence of individual programs producing critical, often life-saving, impacts.

Ultimately, as large as foreign aid budgets may seem, they remain small compared to government budgets of aid-recipient countries. The leverage point for sustained impact of aid is not the direct effect of a single demonstration program, but how the knowledge generated helps recipient governments develop the systems, political will, and technical capacity to demand and apply rigorous evidence for the betterment of their own policies and programs. Thus, aid donors should invest explicitly in evidence about foreign aid projects as a global public good—not primarily as an accountability requirement, but as a core strategy to improve future resource use. Also needed is more work on the transportability of evidence—that is, guidance for when lessons from Lilongwe apply to Lima, and when Jakarta in 2015 applies to Jakarta in 2030.

Three forms of humility are warranted. First, epistemic: "Does aid work?" is no more informative than "Does medicine work?" Both benefits and harms depend on the intervention and context. Second, institutional: Donors urging recipient governments to build state capacity should look in the mirror and improve their own capacity to design and deliver aid. Third, rhetorical: Grandiose claims may sound powerful, but lack proper underlying analysis. For example, a country remaining poor after decades of aid does not establish that aid failed; the relevant question is how it would have fared without that aid. Humility should deter us from false rhetoric for the sake of winning political points in public discourse.

What is not in serious doubt is that the tools now exist to do foreign aid dramatically better than it has historically been done. The evidence base is substantial, the methodology is mature, and the institutional models, from innovation funds to

embedded government labs, have proven their viability across diverse contexts. We must sharpen our demand even more for theory and evidence about what works, for whom, and under what conditions. Meanwhile, we have a lot of evidence to work with now and must work to close the gap between what the evidence says and what we do.

■ *I thank Sarah Bermeo, Kim Cernak, Deborah Cohen, Alison Fahey, Juan Manuel Hernández-Agramonte, Cindy Huang, Kelsey Jack, Gabriela Karlan, Chris Maloney, Håvard Nygård, Elina Scheja, Caitlin Tulloch, Anna Vanderkooy, and Christopher Udry for insightful feedback; Hope Cutchins, Elijah Goldberg, Megan McGuire, and Sneha Stephen for both insightful feedback and editing support; and, Nathan Ding and Vanessa Ikirezi for research assistance. All views and any misjudgments are my own.*

References

Abaluck, Jason, Laura H. Kwong, Ashley Styczynski, Ashraful Haque, Md. Alamgir Kabir, Ellen Bates-Jefferys, Emily Crawford et al. 2021. "Impact of Community Masking on COVID-19: A Cluster-Randomized Trial in Bangladesh." *Science* 375 (6577): eabi9069.

Ahmad, Ahmad Hassan, Eric J. Pentecost, and Marie M. Stack. 2023. "Foreign Aid, Debt Interest Repayments and Dutch Disease Effects in a Real Exchange Rate Model for African Countries." *Economic Modelling* 126: 106434.

Aknin, Lara B., Guy Mayraz, and John F. Helliwell. 2017. "The Emotional Consequences of Donation Opportunities." *Journal of Positive Psychology* 12 (2): 169–77.

Angrist, Noam, Tahir Andrabi, Stefan Dercon, Sharnic Djaker, James Habyarimana, Michelle Kaffenberger, Clare Leaver, Akshay Mangla, Nompumelelo Mohohlwane, and Abhijeet Singh. 2026. "An Agenda for Implementation Science in Education." *Science* 393 (6816): 1084–87.

Angrist, Noam, and Rachael Meager. 2023. "Implementation Matters: Generalizing Treatment Effects in Education." Annenberg Institute at Brown University EdWorkingPaper 23-802.

Augenblick, Ned, B. Kelsey Jack, Supreet Kaur, Felix Masiye, and Nicholas Swanson. 2026. "Retrieval Failures and Consumption Smoothing: A Field Experiment on Seasonal Poverty." NBER Working Paper 25430.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2017. "Labor Markets and Poverty in Village Economies." *Quarterly Journal of Economics* 132 (2): 811–70.

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton. 2017. "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application." *Journal of Economic Perspectives* 31 (4): 73–102.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122 (3): 1235–64.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236): 1260799.

Bastagli, Francesca, Jessica Hagen-Zanker, Luke Harman, Valentina Barca, Georgina Sturge, Tanja Schmidt, and Luca Pellerano. 2016. *Cash Transfers: What Does the Evidence Say? A Rigorous Review of Programme Impact and of the Role of Design and Implementation Features.* Overseas Development Institute.

Bates, Mary Ann, and Rachel Glennerster. 2017. "The Generalizability Puzzle." *Stanford Social Innovation Review* 15 (3): 50–54.

Benedek, Dora, Ernesto Crivelli, Sanjeev Gupta, and Priscilla Muthoora. 2014. "Foreign Aid and Revenue: Still a Crowding-Out Effect?" *FinanzArchiv* 70 (1): 67–96.

Beras, Erika, and Wailin Wong, dirs. 2025. "Farmers Will Be Hit Hard by the Dismantling of USAID." *All Things Considered*. NPR, February 27. https://www.npr.org/2025/02/27/nx-s1-5311663/farmers-will-be-hit-hard-by-the-dismantling-of-usaid.

Berman, Eli, Jacob N. Shapiro, and Joseph H. Felter. 2011. "Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq." *Journal of Political Economy* 119 (4): 766–819.

Bermeo, Sarah Blodgett. 2017. "Aid Allocation and Targeted Development in an Increasingly Connected World." *International Organization* 71 (4): 735–66.

Bermeo, Sarah Blodgett. 2018. *Targeted Development: Industrialized Country Strategy in a Globalizing World*. Oxford University Press.

Blinken, Antony. 2022. "Secretary Antony J. Blinken at the USAID Democracy Delivers Event." Speech, USAID Democracy Delivers, New York, September 20. https://2021-2025.state.gov/secretary-antony-j-blinken-at-the-usaid-democracy-delivers-event/.

Booth, Robert, and Ian Black. 2010. "WikiLeaks Cables: Yemen Offered US 'Open Door' to Attack al-Qaida on Its Soil." *The Guardian*, December 3. https://www.theguardian.com/world/2010/dec/03/wikileaks-yemen-us-attack-al-qaida.

Breeze, Beth. 2005. "The Return of Philanthropy." *Prospect*, January 16. https://www.prospectmagazine.co.uk/essays/56671/the-return-of-philanthropy.

Brown, Nora. 2024. "So Long, Farewell: The End of INVEST." *Medium*, September 10. https://medium.com/usaid-invest/so-long-farewell-the-end-of-invest-e1c7df47b407.

Burnside, Craig, and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90 (4): 847–68.

Bush, George W. 2002. "Remarks to the United Nations Financing for Development Conference." Speech, UN Financing for Development Conference, Monterrey, Mexico. March 22. https://www.presidency.ucsb.edu/documents/remarks-the-united-nations-financing-for-development-conference-monterrey-mexico.

Byerlee, Derek, and John K. Lynam. 2020. "The Development of the International Center Model for Agricultural Research: A Prehistory of the CGIAR." *World Development* 135: 105080.

Carcelli, Shannon P. 2023. "Special Interests in Foreign Policy Bureaucracies: Evidence from Foreign Aid." *Journal of Politics* 85 (3): 905–18.

Carter, Samantha, Edward Jee, Dean Karlan, and Patrick Premand. 2026. "Doing Better at Doing Good: Lessons from a Bayesian Meta-Analysis after 20-Years of Poverty Graduation Programs." Unpublished.

Center for Effective Global Action. 2025. "Scaling up a Simple Planning Intervention to Help Farmers Save More of Their Harvest in Malawi." February 5. https://cega.berkeley.edu/feature/scaling-up-a-simple-planning-intervention-to-help-farmers-save-more-of-their-harvest-in-malawi/.

Crosta, Tommaso, Dean Karlan, Finley Ong, Julius Rüschenpöhler, and Christopher R. Udry. 2026. "Unconditional Cash Transfers: A Bayesian Meta-Analysis of Randomized Evaluations in Low and Middle Income Countries." NBER Working Paper 32779.

Dahl, Benjamin A., Beth Tritter, Deena Butryn, Melissa Dahlke, Sean Browning, Richard Gelting, Monica Fleming, et al. 2024. "Global VAX: A U.S. Contribution to Global COVID-19 Vaccination Efforts, 2021–2023." *Vaccine* 42 (S3): 125827.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55.

Dhaliwal, Iqbal. 2025. "The Evidence Effect: Ideas for the Future of Development." MIT Abdul Latif Jameel Poverty Action Lab (J-PAL) blog, May 12. https://www.povertyactionlab.org/blog/5-12-25/evidence-effect-ideas-future-development.

DIME-WFP. 2024. *Transforming Women's Participation via Cash-Based Transfer Programs*. World Bank and the World Food Programme.

Duflo, Esther. 2017. "The Economist as Plumber." *American Economic Review* 107 (5): 1–26.

Easterly, William. 2003. "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives* 17 (3): 23–48.

Easterly, William, Ross Levine, and David Roodman. 2004. "Aid, Policies, and Growth: Comment." *American Economic Review* 94 (3): 774–80.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker. 2022. "General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya." *Econometrica* 90 (6): 2603–43.

Eurobarometer. 2015. "The European Year for Development—Citizens Views on Development, Cooperation and Aid." European Commission, January. https://europa.eu/eurobarometer/surveys/detail/2022.

Farnsworth, Clyde H. 1991. "Egypt's 'Reward': Forgiven Debt." *New York Times*, April 10. https://www.nytimes.com/1991/04/10/business/egypt-s-reward-forgiven-debt.html.

Federal Ministry for Economic Cooperation and Development. 2025. "The German Government's National Security Strategy." German Federal Ministry for Economic Cooperation and Development, June 18. https://www.bmz.de/en/ministry/german-national-security-strategy.

Feyzioglu, Tarhan, Vinaya Swaroop, and Min Zhu. 1998. "A Panel Data Analysis of the Fungibility of Foreign Aid." *World Bank Economic Review* 12 (1): 29–58.

Fund for Innovation in Development. 2025. "FID at a Glance." Fund for Innovation in Development. https://fundinnovation.dev/en/impact#discover.

Glennerster, Rachel, and Siddhartha Haria. 2025. "Radical Simplification: A Practical Way to Get More Out of Limited Foreign Assistance Budgets." Center For Global Development, April 21. https://www.cgdev.org/blog/radical-simplification-practical-way-get-more-out-limited-foreign-assistance-budgets.

Global Education Evidence Advisory Panel. 2023. *Cost-Effective Approaches to Improve Global Learning: What Does Recent Evidence Tell Us Are "Smart Buys" for Improving Learning in Low- and Middle-Income Countries?* World Bank.

Hilton, Matthew. 2026. *Charity after Empire: British Humanitarianism, Decolonisation and Development.* Cambridge University Press.

Honig, Dan. 2018. *Navigation by Judgment: Why and When Top-Down Management of Foreign Aid Doesn't Work.* Oxford University Press.

Intergovernmental Authority on Development. 2026. "IGAD's Conflict Early Warning and Response Mechanism (CEWARN)." IGAD Djibouti. https://igad.int/cewarn-2/.

Jakubowski, Aleksandra, Don Mai, Steven M. Asch, and Eran Bendavid. 2019. "Impact of Health Aid Investments on Public Opinion of the United States: Analysis of Global Attitude Surveys From 45 Countries, 2002–2016." *American Journal of Public Health* 109 (7): 1034–41.

J-PAL. 2021. "Normalizing Mask Usage to Combat Covid-19." October 27. https://www.povertyactionlab.org/case-study/normalizing-mask-usage-combat-covid-19.

J-PAL. 2022a. "Deworming to Increase School Attendance." September 22. https://www.povertyactionlab.org/case-study/deworming-increase-school-attendance.

J-PAL. 2022b. "Teaching at the Right Level to Improve Learning." August 11. https://www.povertyactionlab.org/case-study/teaching-right-level-improve-learning.

J-PAL. 2025a. "Emissions Markets to Reduce Air Pollution." July 1. https://www.povertyactionlab.org/case-study/emissions-markets-reduce-air-pollution.

J-PAL. 2025b. "Targeting the Ultra-Poor to Improve Livelihoods." May 14. https://www.povertyactionlab.org/case-study/targeting-ultra-poor-improve-livelihoods.

J-PAL. 2026. "Digitizing Social Services to Reach the Right People and Reduce Poverty." January 13. https://www.povertyactionlab.org/case-study/digitizing-social-services-reach-right-people-and-reduce-poverty.

Karlan, Dean. 2015. "The Multi-Lateral Development Banks: A Strategy for Generating Increased Return on Investment." Testimony before the US House Committee on Financial Services.

Kremer, Michael, Sasha Gallant, Olga Rostapshova, and Milan Thomas. 2021. "Is Development Economics a Good Investment? Evidence on Scaling Rate and Social Returns from USAID's Innovation Fund." Unpublished.

Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin. 2021. "Recruitment, Effort, and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools." *American Economic Review* 111 (7): 2213–46.

Levine, Ruth, and William Savedoff. 2015. "The Future of Aid: Building Knowledge Collectively." Center for Global Development Policy Paper 050.

Lumsdaine, David Halloran. 1993. *Moral Vision in International Politics: The Foreign Aid Regime, 1949–1989.* Princeton University Press.

Marliani, Lina, and Paula Pedro. 2025. "The Evidence Effect: How LMIC Governments Drive and Scale Effective Poverty Reduction Policies." J-PAL blog, June 5. https://www.povertyactionlab.org/blog/6-5-25/evidence-effect-how-lmic-governments-drive-and-scale-effective-poverty-reduction.

Mitchell, Andrew. 2011. "Interview on Foreign Aid." *BBC News*, March 1. https://www.bbc.com/news/av/uk-politics-12606763.

Muralidharan, Karthik. 2024. *Accelerating India's Development: A State-Led Roadmap for Effective Governance.* Penguin Random House India.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2023. "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India." *Econometrica* 91 (4): 1261–95.

Nye, Joseph S., Jr. 1990. *Bound to Lead: The Changing Nature of American Power.* Basic Books.

Nye, Joseph S., Jr. 2004. *Soft Power: The Means to Success in World Politics.* PublicAffairs.

Otis, Nicholas G., Rowan Clarke, Solène Delecourt, David Holtz, and Rembrand Koning. 2026. "The Uneven Impact of Generative Artificial Intelligence on Entrepreneurial Performance: Evidence from a Field Experiment in Kenya." *Management Science.* https://doi.org/10.1287/mnsc.2024.06909.

Partnership for Economic Inclusion. 2024. *2024 Annual Report.* World Bank.

Prochaska, Frank. 1988. *The Voluntary Impulse: Philanthropy in Modern Britain*. Faber and Faber.

Rajan, Raghuram G., and Arvind Subramanian. 2008. "Aid and Growth: What Does the Cross-Country Evidence Really Show?" *Review of Economics and Statistics* 90 (4): 643–65.

Ravallion, Martin. 2009. "Evaluation in the Practice of Development." *World Bank Research Observer* 24 (1): 29–53.

Rubio, Marco. 2025. "Opening Statement of Marco Rubio to Be Secretary of State." Speech, US Senate Committee on Foreign Relations, Washington, DC, January 15. https://www.foreign.senate.gov/imo/media/doc/6df93f4b-a83c-89ac-0fac-9b586715afd8/011525_Transcript_Nomination%20of%20Marco%20Rubio%20to%20be%20Secretary%20of%20State.pdf.

Svensson, Jakob. 2000. "Foreign Aid and Rent-Seeking." *Journal of International Economics* 51 (2): 437–61.

USAID. 2021. "Request for Applications for a Resilience Food Security Activity in Mozambique." Funding Opportunity Number: 720BHA21RFA00001. https://files.simpler.grants.gov/opportunities/18ffb253-f06b-4906-aff3-9341f7d1ddbb/attachments/6f78adc3-bef2-4112-97ba-16028ca4f6c9/FINAL_Mozambique_FY22_RFSA_RFA.pdf.

USAID. 2024. USAID Position Paper: Cost Effectiveness. USAID.

Wagner, Alex. 2001. "Bush Authorized to Lift Sanctions on Pakistan." Arms Control Association. https://www.armscontrol.org/act/2001-11/press-releases/bush-authorized-lift-sanctions-pakistan.

World Food Programme. 2019. "Resilience in a Changing Climate: Impact Evaluation Window." World Food Programme.

WHO. 2026. "Global Influenza Surveillance and Response System (GISRS)." https://www.who.int/initiatives/global-influenza-surveillance-and-response-system.

World Bank. 2025. *Driving Global Impact: Accelerating Economic Development with Trial-and-Adopt, LEADS, and ImpactAI.* World Bank.

Yoho, Ted, and Dean Karlan. 2025. "A Theory and Evidence Test for Strategic International Aid." *The Hill*, March 27. https://thehill.com/opinion/international/5215080-reimagined-foreign-aid-strategy/.