



# The rise of randomized controlled trials (RCTs) in international development in historical perspective

Luciana de Souza Leão<sup>1</sup> · Gil Eyal<sup>1</sup>

Published online: 15 May 2019  
© Springer Nature B.V. 2019

## Abstract

This article brings a historical perspective to explain the recent dissemination of randomized controlled trials (RCTs) as the new “gold standard” method to assess international development projects. Although the buzz around RCT evaluations dates from the 2000s, we show that what we are witnessing now is a *second* wave of RCTs, while a *first* wave began in the 1960s and ended by the early 1980s. Drawing on content analysis of 123 RCTs, participant observation, and secondary sources, we compare the two waves in terms of the participants in the network of expertise required to carry out field experiments and the characteristics of the projects evaluated. The comparison demonstrates that researchers in the second wave were better positioned to navigate the political difficulties caused by randomization. We explain the differences in the expertise network and in the type of projects as the result of concurrent transformations in the fields of development aid and the economics profession. We draw on Andrew Abbott’s concept of “hinges,” as well as on Bourdieu’s concept of “homology” between fields, to argue that the similar positions and parallel struggles conducted by two groups of actors in the two fields served as the basis for a cross-field alliance, in which RCTs could function as a “hinge” linking together the two fields.

**Keywords** Expertise · Fields · Hinges · Philanthro-capitalists · Policy evaluation · Randomistas

Few things have shaped development economics and policy evaluation as much as the turn to randomized controlled trials (RCTs) over the past two decades. Initially restricted to a handful of researchers, the number of RCT evaluations has grown

✉ Luciana de Souza Leão  
ldl2112@columbia.edu

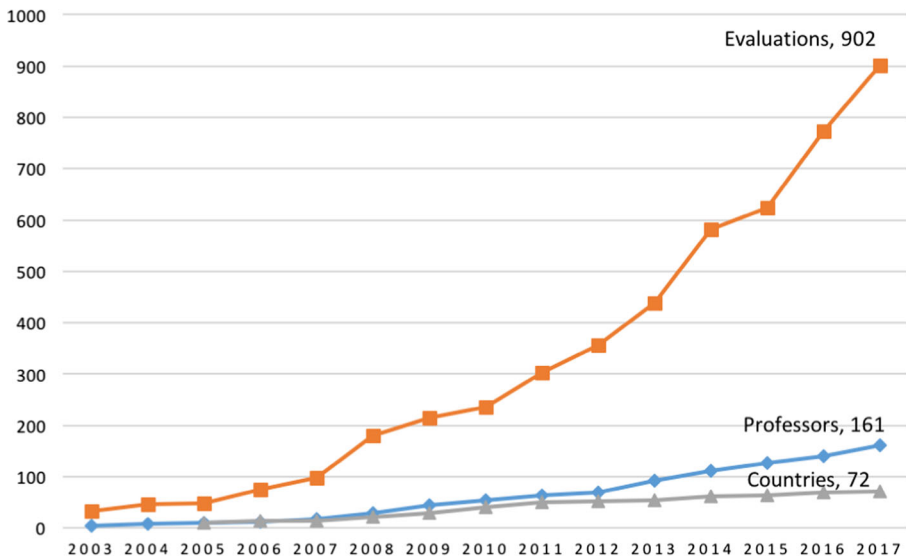
Gil Eyal  
ge2027@columbia.edu

<sup>1</sup> Department of Sociology, Columbia University, 501 Knox Hall, 606 west 122nd street, New York, NY 10027, USA

exponentially over the past years. As can be seen in Fig. 1, when the Abdul Latif Jameel Poverty Action Lab at MIT (J-PAL) was founded in 2003, it consisted of 4 affiliated Professors and conducted 33 projects. By 2017, there were 161 affiliated Professors and they were involved in 902 evaluations in 72 countries. The honors bestowed on J-PAL affiliates, also known as *randomistas* (Deaton 2006), testify to the prominence of RCTs: since 2003, they were the recipients of five John Bates Clark Medals and three McArthur “genius” awards, and have commanded considerable attention from the media and foundations. As the charismatic co-leader of J-PAL, Esther Duflo puts it: “[RCTs] became this brand.... It became a thing—in academia and outside organizations” (quoted in Parker 2010).

Similar to medical trials, RCT evaluations in development economics attempt to measure the impact of different projects on their participants by comparing them with a randomly selected no-treatment control group. A 2006 malaria prevention experiment (Cohen and Dupas 2010) can serve as an illustration. The question was whether to distribute bed nets for free or to require a small co-payment by recipients. The researchers selected 5 health clinics in rural Kenya to distribute bed nets for free, 11 other clinics distributed the bed nets for varying small fees, while 4 additional clinics did not distribute bed nets at all. After 3 months, researchers compared the control and treatment groups in terms of demand for bed nets and actual use. They found that charging even a small price reduced demand considerably, while receiving bed nets for free increased usage. This finding led to a policy prescription of providing bed nets for free. Experiments like this are now performed in multiple international development policy areas.

What could explain the rise of RCTs in international development? *Randomistas* tend to present it as due to the intrinsic merits of their method, its ability to produce “hard” evidence as compared with the “softer” evidence provided by case studies or



Source: J-PAL (2018)

Fig. 1 J-PAL annual growth (2003–2017)

regressions (Banerjee and Duflo 2011; Glennerster and Takavarasha 2013). They compare development RCTs to clinical trials in medicine, implying that their success is due to the same “gold standard” status in the hierarchy of evidence: “It’s not the Middle Ages anymore, it’s the 21st century ... RCTs have revolutionized medicine by allowing us to distinguish between drugs that work and drugs that don’t work. And you can do the same randomized controlled trial for social policy” (Duflo 2010).

This explanation does not pass muster and need not detain us for very long. Econometricians have convincingly challenged the claim that RCTs produce better, “harder” evidence than other methods (Rodrik 2008; Deaton and Cartwright 2016). Their skepticism is amply supported by evidence that *medical* RCTs suffer from numerous methodological shortcomings (Demortain 2011, pp. 53–57), and that political considerations played a key role in their adoption (Carpenter 2010; Marks 1997). These objections accord with the basic insight of science studies, namely, that the success of innovations cannot be explained by their *prima facie* superiority over others, because in the early phases of adoption such superiority is not yet evident. It is only evident in retrospect, thus representing a version of the retrospective fallacy. To explain success, argue science studies scholars, one has to examine the rhetorical and political strategies by which champions of these innovations recruit allies and convince audiences of said superiority (Latour 1987; Pinch and Bijker 1984; Barnes et al. 1996).

The few social scientists who attempted to explain the recent spread of development RCTs have generally followed this insight. They explain the success of RCTs as due to the “rhetorical ... and organizational strategies” employed by the *randomistas* to problematize development knowledge, to offer “RCTs as a means to reduce uncertainty” and to transform contested development questions into seemingly technical problems (Berndt 2015; Donovan 2018, p. 29; Pritchett in Ogden 2016). The agency of this “thought collective” is the key explanatory mechanism in these accounts (Donovan 2018, pp. 34–36; Pritchett in Ogden 2016, pp. 139–142).

While it represents an advance over the self-congratulatory accounts of the *randomistas*, we think that this account is incomplete and ill-specified. Methodologically, it lacks a means of testing the hypothesis and evaluating the significance of different causal factors. Theoretically, it fails to recognize that the key problem is to explain the creation of an *enduring link between fields*. It fails to appreciate the resistance faced by those who attempt to build this link. And it puts too much of the explanatory burden on the foresight and interested strategizing of the actors.

In what follows, we use the comparative method to derive and test a more complete explanation for the success and spread of development RCTs. While the buzz around RCTs certainly dates from the 2000s, the assumption—implicit in both the *randomistas*’ and their critics’ accounts—that the experimental approach is new to the field of international development—is wrong. In reality, we are witnessing now a *second* wave of RCTs in international development, while a *first* wave of experiments in family planning, public health, and education in developing countries began in the 1960s and ended by the early 1980s. In between the two periods, development programs were evaluated by other means (USAID 2009, p. 16). We treat the sequence of first and second waves as cases of “reiterated problem-solving” (Haydu 1998): Instead of asking, “why are RCTs increasing now?” we ask, “why didn’t RCTs spread to the same extent in the 1970s, and why were they discontinued?” In other words, how we explain the success of the second wave must be consistent with how we explain the failure of the first.

The comparison demonstrates that the recent widespread adoption of RCTs is not due to their inherent technical merits nor to rhetorical and organizational strategies. Instead, it reflects the ability of actors in the second wave to overcome the *political resistance to randomized assignment*, which has bedeviled the first wave, and to forge an enduring link between the fields of development aid and academic economics. To advance this argument, the first section of this article develops a theoretical framework that combines Abbott's (2005) concept of "hinge" with Bourdieu's concept of homologies to account for how fields become durably linked. It is followed by a brief presentation of our data and methods. The third section compares the two waves of RCTs as "hinges," and shows that the political resistance to randomized assignment is much less significant for the second wave because researchers are answerable to a different audience from that in the past. Resistance is less significant also because second wave RCTs typically evaluate interventions that are much shorter and smaller in scale than in the past. The fourth section demonstrates that these differences stem from homologous transformations in the fields of development aid and academic economics. These transformations created the conditions for an alliance, across field boundaries, between economists and the leaders of private foundations. The fifth section details the elective affinities that facilitated this alliance. In conclusion, we summarize our findings and underscore the contribution of the article to research on how fields become durably connected.

### The theoretical framework: RCTs as a hinge between fields

The problem common to both the first and second waves of RCTs was *how to turn foreign aid into a "science" of development*. Since foreign aid is about the allocation of scarce resources, the decisions of donors and policy-makers need to be legitimized. One way of doing so is by recruiting academic experts, whose advice serves to defuse criticisms and legitimate decisions as efficacious and rational. For their part, academics also stand to gain from such recruitment in the form of political influence and material resources. Yet, interests alone are not enough to bring the two together and forge this science of development. The advice offered by academic experts could easily be criticized as subjective, biased, and detached from real-world considerations, while policy-related research is often viewed as below disciplinary standards.

This problem is a sub-species of a more general sociological question about how distinct fields become durably connected to one another, despite the fact that they are governed by conflicting logics. This question is at the core of any sociological investigation of policy-oriented forms of expertise, as institutionalized, for example, in think tanks or the research departments of global financial institutions (Medevtz 2012; Babb 2009). Given that the academic field is governed by a distinct logic and set of incentives, to create a durable link connecting it with the field of development aid—a field governed by more practical and often political considerations—is to take on a delicate and precarious task that can only happen under specific conditions (Murray 2010).

This challenge can help explain the preference for RCTs in both waves. As a form of "mechanical objectivity ... based completely on explicit rules," rather than expert judgment (Porter 1995, pp. 6–8), RCTs are calculated to put to rest suspicions of

subjectivity and bias, while also appearing to adhere to the highest standards of disciplinary rigor. They seem, therefore, to be able to bridge between the conflicting logics and demands of the development aid and academic fields. As J-PAL's co-leader, Abhijit Banerjee (2007, pp. 115–116) says, “the beauty of randomized evaluations is that the results are what they are: we compare the outcome in the treatment with the outcome in the control group, see whether they are different, and if so by how much.” By saying that the results are “what they are,” Banerjee means that they are objective, unbiased, independent of any subjective assumptions, and therefore academically rigorous and trustworthy, as well as legitimate from the point of view of actors in the field of development aid.

If considered merely as a rhetorical strategy, mechanical objectivity should have worked similarly for both waves. To explain why it ultimately failed in the first wave but is now successful in the second, we need a different way of thinking about how fields become durably connected. As originally articulated by Bourdieu (1975, 1977), however, field theory is mostly silent on this question, because of its near equation of “fieldness” with autonomy and hence distinctness (Eyal 2013). The recent interest in interstitial fields can be seen as a response to this myopia and is a welcome start (Medevtz 2012; Stampnitzky 2013; Panofsky 2011). A more direct approach, however, is offered by Abbott's (2005) concept of “hinge.” Abbott suggests that ecologies or fields<sup>1</sup> can become linked to one another through the construction of “hinges,” by which he means “issues that provide ... dual rewards,” competitive “strategies that work well in one ecology as in the other,” thus enabling an alliance among actors across boundaries (ibid., p. 255).

The hinge metaphor is valuable for our purposes because it stresses that, for separate fields to become linked, a *mechanism* needs to be built. A rhetorical strategy would not suffice. A hinge is a device with different parts that must “hang together” for the mechanism to work. We suggest to think about it as an expertise network that connects actors across field boundaries in order to accomplish a task (Eyal 2013). For the hinge to enable the simultaneous pursuit of dual rewards in two different fields, one would need to overcome resistance created by the tension between the incommensurable logics of the fields. In this way, the concept of “hinge” gives us purchase over the empirical problem of why RCTs were discontinued in the first wave, though they are successful now.

For RCTs to work as a hinge, there must be *randomized assignment to treatment and control groups*. This is how potential biases are removed; how the results are certified as rigorous according to disciplinary standards; and how the policy statements produced by RCTs inspire trust among donors and decision-makers. Yet, there is nothing simple or straightforward about randomized assignment to a no-intervention control group, as is demonstrated, for example, by the contentious history of medical clinical trials (Carpenter 2010). Because development interventions usually involve a form of social assistance, any attempt to assign people randomly to a “no-intervention” control group incurs strong political resistance from participants, the implementing

<sup>1</sup> Since Abbott (1988, p. 343) himself suggested that his concept of “task area” is similar to Bourdieu's “field,” and since the distinction between “ecologies” and “task areas” is immaterial for our purposes, we use “ecology” and “field” here interchangeably. We follow Mudge and Vauchez (2012) in incorporating Abbott's concepts into field analysis and treating the two approaches as compatible and complementary.

bureaucracy, and politicians. How does one get people to participate in such an exercise of their own accord, without certainty that they would get something out of it?

As Gueron (2017, p. 5) says, this makes RCTs an especially “tough sell”: “why would any politician or administrator chance adverse publicity, a potential lawsuit, bureaucratic backlash, or even staff revolt?” Even when randomized assignment does take place, legal mandates to provide assistance can exacerbate all the known problems of “substitution bias”—i.e., the possibility that individuals in the control group are participating in other available policy options (Heckman et al. 2000). The attempt to implement RCTs in international development, therefore, could succeed only under very special conditions, and it requires “political skills and savvy” (Gueron 2017, p. 5). The tension between the differing logics of the academic field—in which the RCT is an “intervention” being evaluated—and the field of development aid—in which it could be construed as the provision of “assistance” to some but not to others (Rayzberg 2019)—make it exceedingly difficult to build a functioning hinge.

This point is ignored by most other attempts to explain the rise of RCTs, with the exception of Rayzberg’s (2019) insightful analysis. A comparison between the two waves of RCTs is perfectly suited, however, to highlight its significance and develop its implications. In the third section of this article, we show that differences in the composition of the network of expertise required to carry out RCTs, and in the characteristics of the projects evaluated, largely explain why RCTs were not able to function as a hinge in the first wave and why they were discontinued—because the political resistance to randomized assignment outweighed its advantages. This was because donors and decision-makers in the first wave tended to be governmental agencies and thus politically exposed; and because first wave RCTs typically evaluated long-term, broad social policy *programs* with significant consequences in terms of health, education, and material inequality. Consequently, political resistance to randomized assignment was much stronger in the first wave. In contrast, the composition of the expertise network in the second wave—especially the role played by philanthrocapitalists<sup>2</sup> and global NGOs—renders it far more insulated from political pressures, as well as better endowed with resources and more tightly controlled. Similarly, RCTs in the second wave are typically smaller, of shorter duration, and they evaluate well-bounded “interventions” that are less consequential in terms of allocating of scarce resources. Consequently, they are able to function as a hinge linking the academic and development aid fields.

Abbott’s approach, however, is limited when it comes to the question of why and how hinges emerge. Given the tensions between the two fields, it is not obvious why actors would be motivated to try to build a hinge. Abbott (2005, p. 255) pays little attention to this question, saying only that “synchronic and diachronic patterns within and between ecologies create possibilities for alliances between actors and locations across the borders of ecologies.” In effect, Abbott must be seen as saying that hinges emerge 1) when internal dynamics in both fields change in tandem; 2) to create objective possibilities for alliance across boundaries; 3) and to foster a shared

<sup>2</sup> The term “philanthro-capitalists” has been coined by Bishop and Green (2009) to describe new, large foundations, established by successful individuals from the worlds of finance and technology, who “apply business techniques and ways of thinking about their philanthropy” (ibid., p. 6). The often-cited example is the Bill and Melinda Gates Foundation (created in 2000), though the term captures a broader movement in the philanthropy sector (Reckhow 2013).



perception of these possibilities. The explanatory burden is high. The imputation of interests, even together with the identification of changing conditions of possibility, does not suffice if one cannot also account for how actors' perceptions of these interests and conditions change and become aligned. This is something for which Bourdieu's (1977) praxeological approach, his concepts of "homologies" and "strategy without a strategist," is much better suited (Wacquant 1992, p. 25). In this approach, the successful construction of hinges is not the result of consciously formulated strategy, but of predispositions and perceptual schemas that are shared across boundaries because they were formed in the course of conducting parallel struggles in homologous fields. Actors' perceptions of their interests are refracted and shaped by the relational structure of oppositions in their field, while the homology between these relational structures works to create "elective affinities" across boundaries and bring them closer together.

The fourth and fifth sections of this article, therefore, supplement the concept of "hinge" with an analysis of the homologies between fields and the elective affinities to which they give rise. The fourth section traces the transformations in the fields of development aid and academic economics, to demonstrate how second wave RCTs became the "hinge" linking them. As the field of development aid fragmented with the end of "the Washington Consensus" (Rodrik 2006; Babb and Chorev 2016), RCTs served newcomers like the philanthro-capitalists to challenge the managerial style of the older foundations (e.g. Rockefeller, Ford) and governments, who relied on expert judgment. As the field of the economics profession was disrupted by struggles over causal identification, RCTs offered young development economists a means of shielding themselves from the anomic effects of these struggles, while challenging the leadership of the field over its reliance on "priors." Thus, what led to the alliance between these two groups was not any rhetorical strategy or rational appraisal of their shared interests, but the fact that they were both relative newcomers conducting homologous struggles against the established orthodoxies of their fields at a moment of relative disorganization.

While others have also attributed the emergence of second wave RCTs to transformations in the fields of economics and development aid (Donovan 2018; Pritchett 2016), our comparative framework allows us to isolate the latter as the decisive, necessary (though not sufficient) condition. In both waves there were social scientists trying to conduct RCTs, but only in the second wave was there an ecology hospitable to their efforts. The key difference between the two waves is the fragmentation of development aid, the fact that the field is no longer dominated by Official Direct Assistance (ODA) between governments. As long as the most significant audiences for RCTs were the leaders of state and semi-state agencies, as well as politicians and bureaucrats in developing countries, their embrace of the technocratic consensus of the day and their political exposure spelled a much weaker interest in RCTs and a tendency to revert back to reliance on expert judgment. In contrast, because the philanthro-capitalists work with global NGOs rather than governments, they provide the *randomistas* with relative immunity from the political resistance to randomized assignment. The *randomistas* themselves, therefore, are not a necessary condition. An alternative scenario in which private donors ally themselves with a different group of experts conducting RCTs is entirely conceivable. The *randomistas* are a contingent factor, a "historical switchman" that happened to be in the right place at the right time.

For this very reason, however, their distinctive, contingent characteristics play a crucial role in shaping the final result, which justifies the attention we pay to their formative experiences.

In the fifth section, we analyze the elective affinities between the worldviews of the *randomistas* and the philanthro-capitalists to show how homologous struggles lead to the formation of an alliance across field boundaries. We draw on Daston and Galison's (1992, p.82) analysis of objectivity as a profoundly *negative* concept, whose meaning is derived from whatever facet of *subjectivity* is problematized as dangerous or misleading. This means that an ideal of objectivity, like the preference for RCTs, typically constructs an image of the virtuous expert, characterized by a set of ascetic values (ibid., pp. 83, 122), through rejection of qualities that thereby come to be understood as "subjective." From Bourdieu's point of view, this must be understood as a strategy seeking to change the balance of power within a field. The qualities problematized as "subjective" are likely to be the virtues claimed by the orthodoxy of the field. Indeed, the orthodoxy does not consider them subjective at all, but constitutive of what Porter (1995, pp. 3–4) calls "disciplinary objectivity," namely the specialized knowledge held by expert communities. In contrast, the new virtues of mechanical objectivity are championed by heterodox challengers seeking to delegitimize their opponents. We show that the *randomistas* and the philanthro-capitalists share an enthusiasm for measurement and "trust in numbers" rooted in their common attack on the "biases" of expert judgment. They share a distinct preference for leverage strategies, whereby a small intervention is strategically deployed to achieve much larger ends, because of their common criticism of the ideological attachment of government planners to "programs." And they agree on an ascetic, self-limiting vision of the virtuous expert as a "choice architect" practicing "libertarian paternalism."

## Data and methods

To compare the two waves of RCTs in international development, we used two analytical strategies. The first step was to construct two datasets, each a sample from the total population of studies in each period. Then we were able carry out a comparison of these datasets.

To obtain a sample of studies from the first wave, we consulted extensive reviews published by the World Bank in collaboration with the Population Council (Cuca and Pierce 1977; Searle 1985), as well as web-based repositories of published RCTs produced by the Campbell Collaboration and 3ie—the International Initiative for Impact Evaluation. To correct for a possible bias towards over-representation of studies supported by these organizations, we also searched major academic databases in health, economics, and public policy provided by platforms such as EbscoHost and the Central Trials Registrar from the Cochrane Library.<sup>3</sup> Additional bibliographies were located by following citations in this initial list (Riecken and Boruch 1975; Boruch et al. 1978; Bauman 1997).

<sup>3</sup> We also conducted interviews and exchanged emails with key academic figures from the first wave, as well as with Trials Search Coordinators from the Campbell Collaboration and the Cochrane Library to look for possibly missing repositories and reviews from the period.



This search strategy yielded a population of 114 experimental studies conducted from 1953 to 1986 in developing countries. By reading the abstracts or bibliographic annotations about these studies, we determined that only 60 qualified as experimental studies comparing treatment and control groups (half of which conducted between 1966 and 1973), while the other 54 were quasi-experimental studies.<sup>4</sup> Finally, we obtained a corresponding report or publication for each of the 60 experimental studies, either from published sources or from the archives of the World Bank, USAID, and the Population Council. Each report was read by one of the researchers and coded for multiple aspects of the evaluation (duration, size of sample, implementing partners, region, and unit of randomization).

Obtaining a sample of second wave RCTs was more straightforward because the online library of the MIT Poverty Lab provides a complete list of RCTs conducted by J-PAL and affiliates. Our analytical sample was defined on January 13th, 2016. Of the 625 RCTs that were listed in J-PAL's library that month, we excluded 100 studies that were not conducted in developing countries. We then drew a random sample of 100 RCTs to be analyzed. From these, we excluded all RCTs that were still on-going or for which we could not identify a corresponding publication or policy report. This left us with a final sample of 63 RCTs. As with the first wave sample, for each RCT, we read its corresponding academic paper or policy report to extract the relevant information. We then crosschecked data from these publications with detailed information about funding partners and experimental design provided in J-PAL's website.<sup>5</sup>

Additionally, the account of the second wave relies on three years of participant observation with *randomistas* by one of the authors, Luciana de Souza Leão. During this period, the author participated in the fieldwork of two RCTs related to microfinance in Peru (2007), as well as in the fieldwork, data analysis, and publication of one RCT in the field of financial education in Brazil (2010–2012). Since we do not possess similar firsthand information about the first wave, we rely on fieldwork information that appeared in archival sources, as well as our interview with Robert Boruch, a key leader of the first wave (on August 31st, 2016).

The second strategy was to analyze the controversy regarding RCTs that appears in academic and policy publications during the second wave. The publications analyzed include articles in economics journals, transcripts and PowerPoints of international development conferences, opinion pieces, and blog posts of key development economists. This secondary literature was particularly useful to examine the ways that advocates and critics narrate the contemporary success of RCTs, as well as to identify the elective affinities between the *randomistas* and the philanthro-capitalists.

## Comparison of the two waves of RCTs

In both waves, academic entrepreneurs similarly set out to organize RCTs and to convince relevant audiences that the experimental method would turn foreign aid into a “science of development.” While this effort faltered in the first wave and RCTs were

<sup>4</sup> Experimental designs that did not clearly explain how the control group was assigned, or that used average data from the population as control, were considered quasi-experimental.

<sup>5</sup> A list with all the studies in our two samples is available upon request.

ultimately replaced by other evaluation methods, it seems—at least from the present vantage point (2019)—to have been successful in the second wave. In this section, we analyze the differences between the two waves along two dimensions—the composition of the expertise network necessary to carry out RCTs and the characteristics of the projects evaluated—to explain why the second succeeded where the first failed. In accordance with the theoretical framework developed earlier, we place particular emphasis on two factors that explain why political resistance to randomized assignment is less significant for the second wave—allowing RCTs to function as a hinge between fields—than it was for the first. First, while the prestige of second wave economists, the resources they command, and the much tighter coupling of the network they constructed, all play a role in explaining the final result, the key difference seems to be their strategic alliance with private foundations as compared with the reliance of first wave researchers on governmental agencies. Second, the shorter duration and limited scope of second wave RCTs similarly explain why the resistance to randomized assignment does not pose as significant an obstacle for them, as it did for first wave RCTs evaluating long-term, large-scale programs.

### Who: participants in the expertise network of RCTs

The most obvious difference in the composition of the networks of expertise (Eyal 2013) involved in RCTs is the disciplinary affiliation of the leading researchers. Table 1 reports the disciplinary affiliation of authors of the papers/reports in our samples of first and second wave RCTs. While academic economists predominate in the second wave authors (80%), there were no economists among the 76 first wave authors for whom we identified a disciplinary specialization.<sup>6</sup> Instead, roughly 20% of these 76 hail from other social sciences or affiliated disciplines (psychology, sociology, population studies, and statistics), with an even larger group (~30%) hailing from public health (physicians, epidemiologists, etc.). These academics typically also held administrative or research positions at the Population Council or were hired as consultants by the World Bank or USAID. The table also lists a smaller group (18.3%) consisting of staff members of these organizations whose disciplinary affiliation was not identified. The leadership of the second wave, therefore, is more cohesive and more autonomous than the first wave's. This is important because of what it means for negotiating with funders, obtaining resources and maintaining tight control over other parties, and for being able to convert policy-related work back into academic and scientific capitals.<sup>7</sup>

Field experiments in development economics, however, are not conducted by senior researchers alone. While they are the personified “authors” of the evaluation, the RCT is a product of the full expertise network described below. A complex organizational effort is required to coordinate the activities of multiple parties and, most importantly,

<sup>6</sup> There is a large number of missing values (unknown disciplinary affiliation) in our first wave data (28% of authors), but we do not think this casts doubt on the finding. Given the decades that have passed, it is not surprising that it is more difficult to identify the disciplinary affiliation of first wave authors. The unidentified, however, are unlikely to be US-based economists because this is the category easiest to identify. We suspect that most were local researchers.

<sup>7</sup> Of 32 development papers in the top 5 economics journals in 2015, 10 were RCTs, up from zero in 2000 (Duflo 2016). No first wave studies, in contrast, appeared in high-visibility disciplinary journals. The main venue was *Studies in Family Planning*, an interdisciplinary journal published by The Population Council.

**Table 1** Disciplinary affiliations of authors in first and second waves of RCTs

	First wave <sup>a</sup> (%)	Second wave <sup>b</sup> (%)
Academic economists	0	80.1
Other social scientists	18.3	6.4
Public health and medicine	29.8	4
Staff, unknown discipline <sup>c</sup>	18.3	0.5
Business & finance	0	4.5
Math & statistics	2.9	0
Others	3.9	1.5
Unknown	26.8	3

<sup>a</sup> Percentage from sample of 104 authors from 55 studies

<sup>b</sup> Percentage from sample of 202 authors from 63 studies

<sup>c</sup> Staff of foundations, NGOs, and governments

to control the control group and prevent attrition. This effort is made even more complex by the fact that these RCTs are implemented in remote areas of extremely poor countries, in the absence of administrative data available in more developed countries, and having to negotiate language and cultural barriers (Teale 2014). In what follows, we compare the two waves in terms of the necessary components of such a network: 1) Coordinating Center; 2) Field Staff; 3) Implementation Partners; and 4) Funders.<sup>8</sup>

The coordinating center of the second wave is a novel organizational entity, the so-called “Poverty Lab.” It is a complex organizational structure with central offices in prestigious academic institutions in the United States and regional offices throughout the Global South. These offices are run by a research team composed of a mix of senior and junior professors. As can be seen in Table 2, the composition of senior researchers at the coordinating center closely tracks the composition of authors in our second wave sample: 93% have PhDs in Economics or Applied Economics, two-thirds (65%) teach in Economics departments, while the others are in Business (15%) and Public Policy Schools (9%). Most of these professors, moreover, have done consulting work at international development organizations (85%), especially the World Bank (25%). They formulate the research question and experimental design, construct the questionnaires, analyze the data, and publish the papers. No less importantly, they negotiate with the implementing partners and funders.

In contrast, most first wave authors held administrative or research positions in large US-based nonprofits that were tightly linked to US foundations and government. Roughly 20% of authors in Table 1 were employed as full-time staff, while many others held a part-time position as consultants or advisers at non-profits such as the Population Council or SSRC’s special committee on social experimentation, or at USAID (Riecken and Boruch 1975, p. ix). In short, the coordinating center of the first wave, where operating procedures and reports were put together, was located within the very same institutions that also provided the funding, technology, and research teams. It

<sup>8</sup> Another necessary component are the policy beneficiaries themselves, but we do not possess data that would allow us to compare them.

**Table 2** Disciplinary affiliations of J-PAL affiliated professors (research team at coordinating center)

	Current department <sup>a</sup> (%)	PhD field <sup>a</sup> (%)
Economics	65.1	85.3
Business school	14.7	0
Applied economics	3.7	7.8
Political science/politics	4.6	2.9
Public policy/government	9.2	3.9
Education	0.9	0
Global health	1.8	0

<sup>a</sup> Percentage from sample of 109 affiliated professors (JPAL 2016)

was a nexus of institutions closely aligned with the US government, wherein there was a definite affinity between “development” and the extension of US influence (Heydemann and Kinsey 2010, p. 222).

This is important to understanding the differences between the two waves. Researchers at the coordinating center in the 1960s and 1970s did not enjoy the benefits of the organizational structure of Poverty Labs, specifically the autonomy (and symbolic capital) associated with Ivy League universities. This novel organizational entity allows second wave *randomistas*, as they themselves admit (Rotemberg 2009), to attract their own sources of funding and implementing partners and to negotiate with these from a position of strength, relatively insulated from political pressures. Consequently, they were able to secure large amounts of funding not just for specific projects, but also for infrastructure such as the creation of regional offices and employment of large field staff. This, in turn, made them extremely attractive partners for NGOs.

The actual RCT, however, is conducted by field staff located at the project site. Here, once again, there is a stark difference between the two waves. The contemporary Poverty Lab employs at its regional offices a large group of recent graduates from US universities, who mediate between the coordinating center and a variety of local actors whose cooperation is necessary for implementation. In January 2016, there were 123 research assistants employed in J-PAL’s seven global offices. As can be seen in Table 3, most were trained in Economics, but many majored in other disciplines related to development studies and public policy.

**Table 3** Disciplinary affiliations of J-PAL junior research staff at regional offices

	Field of graduate degree <sup>a</sup> (%)	Field of undergraduate degree <sup>a</sup> (%)
Economics	35	41
Development studies	19	0
Public policy	20	5
Social sciences/ humanities	11	15
Engineering	9	25
Business	7	8
Other	0	5

<sup>a</sup> Percentage from sample of 123 research staff (JPAL 2016)

First wave field teams, in contrast, even when led by academic researchers, typically lacked this intermediate layer of academically trained auxiliary staff. The few initiatives to create regional centers came to naught because funders and universities were not keen on making the necessary investments (Interview with Boruch). Research assistants were obtained by partnering with local universities and research centers. This meant that first wave researchers were much more reliant on local implementation partners and local bureaucracies and thus more exposed to political resistance to randomization.

The field staff must collaborate with implementation partners in order to carry out the experiment. To implement an RCT in education policy, for example, one has to secure the cooperation of the Ministry of Education and school districts; one may need teachers to administer the intervention, or perhaps contract with an NGO to conduct it. The local staff of NGOs or government agencies are the most crucial group to control, because they have to differentiate their day-to-day practices to create and maintain the division between treatment and control groups. As can be seen in Tables 4, 65% of the implementation partners of first wave studies were central and local government agencies, together with local universities and nonprofits. International nonprofit organizations accounted for a mere 6.3%. If we exclude USAID and The Population Council from the total—because strictly speaking they were not implementation partners, but the coordinating center—the representation of domestic partners would rise to 75%. At the field site, it was mostly the local government staff in partnership with researchers and students from local universities that implemented interventions, administered questionnaires, and collected data. Far more sensitive to local political pressures, they were the most likely source of substitution bias and resistance to randomization: “experimental control was not a high priority for Salvadoran administrators who were trying to deal with a major educational reform ... the political necessity of introducing the reform as a package outweighed th[e] preference [for RCT]” (Hornick et al. 1973, pp. 274–276; see also USAID 2009, p. 18).

**Table 4** Implementing partners in first and second waves of RCTs

	First wave <sup>a</sup> (%)	Second wave <sup>b</sup> (%)
For-profit organization	0	39.4
International nonprofit	6.3	18.1
The Population Council & USAID	12.6	0
Other US government agency	2.1	0
US University	13.7	2.1
National government	21	11.7
Local government	7.4	8.5
Local university	25.3	3.2
Local nonprofit	11.6	17

<sup>a</sup> Percentage from sample of 95 partners from 52 studies

<sup>b</sup> Percentage from sample of 94 partners from 61 studies

In contrast, *randomistas* rely on local bureaucracies to a much lesser extent and instead draw on a new set of implementation partners that mostly did not exist earlier: global NGOs, for-profit organizations (primarily banks involved in micro-finance schemes), and local survey firms. Together they comprise 75% of implementation partners. This means that fieldwork teams in the second wave arrive at a setting that is better prepared in terms of infrastructure for data collection, as compared with the first wave. It also means that the coordinating center and regional offices in the second wave are in a better position to control the implementation of the experimental design and to ignore political resistance to randomized assignment. Working with these NGOs, rather than governments, “is key to enabling randomization and overcoming ethical concerns” because NGOs do not “have to pretend to serve everyone” (Ogden 2016, pp. xx-xxi).

Finally, implementing RCTs in developing countries requires a significant amount of funding. The donors should be understood as an integral part of the RCTs’ network of expertise since without their input the task would not be accomplished and there would not be a functioning hinge. Moreover, their expectations play an important role in shaping research. As can be seen in Table 5 below, the main funding sources for first wave studies were the Population Council and USAID, to which we can add also the “old” Foundations, Ford and Rockefeller (the Population Council being a subsidiary of Rockefeller). The coordinating center of the network, therefore, represented 48% of the total funding sources. Local governments, universities, and nonprofits provided another 20% of funding sources. In contrast, the main funding sources for second wave studies are international organizations and NGOs, which constitute a third of funding sources (with the World Bank accounting for more than half of these), as well as other US foundations (17.5%). Chief among the latter is a new breed of foundations including the Bill and Melinda Gates (6%) and the McArthur (5%) Foundations.

While superficially it may seem that private foundations play the same role in the two waves, there are several key differences that set the new foundations radically apart from their predecessors. Most importantly for our purposes, as Heydemann and Kinsey (2010, p. 222) explain, Ford and Rockefeller foundations worked closely together with the US government, and their “activities ... [were subordinated] to the foreign policy priorities of the state.” They even “drew their leadership and senior staff [from] men who had gained managerial experience in government agencies [...], who carried into their positions commitments ... to advance the policy priorities of the state.”

This state-foundations alliance (and revolving door) has been replaced, in the second wave, by a university-foundations alliance. The foundations led by philanthro-capitalists purposefully distance themselves from US-government influence and do not see themselves as an auxiliary arm of US foreign policy. They explicitly aim to replace the previous ethos with business oriented-models (see Reckhow 2013 for a summary), where they make strategic investments and alliances with academic centers. For this reason, the numbers in Table 5 actually underrepresent the impact of these foundations, since they do not account for the significant *infrastructural* funding they provided for Poverty Labs or their host universities (which constitute another 20% of funding sources). Typically, it was not the lab, but the university, which appeared as award recipient in foundations’ reports. Innovations for Poverty Action (IPA), for example, appeared as awardee for a total of \$31,500,000 from the Gates Foundation in the last ten years (Gates Foundation 2016), but even larger sums were awarded to



**Table 5** Funding sources in first and second waves of RCTs

	First wave <sup>a</sup> (%)	Second wave <sup>b</sup> (%)
International organizations <sup>c</sup>	19.6	33.3
The Population Council & USAID	36.9	5
Other US government agency	4.3	6.7
Other OECD governments	0	7.5
Ford and rockefeller foundations	10.9	0
Other US foundation	4.3	17.5
US University	2.2	20
Local sources <sup>d</sup>	19.6	4.2
For-profit organizations <sup>e</sup>	2.2	5.8

<sup>a</sup> Percentage from sample of 46 sources from 27 studies

<sup>b</sup> Percentage from sample of 120 sources from 52 studies

<sup>c</sup> Including World Bank, International Planned Parenthood Federation, UNICEF

<sup>d</sup> Local governments, nonprofits, universities, and think-tanks

<sup>e</sup> Banks, financial institutes, and private companies

Yale University, where IPA is based. Thus, the total for IPA must have been significantly higher. This is true as well for J-PAL, which describes its own history as consisting of a series of awards from private foundations that allowed “J-PAL to grow significantly” and to establish regional centers. The creation of J-PAL Africa in 2010, for example, was made possible “with the support from the William and Flora Hewlett Foundation.” (JPAL 2016) Additionally, new sources of funding are the government agencies of other OECD countries.

This transformation in the composition of funders was the decisive, necessary condition making possible the multiple ways in which the second wave network differed from the first’s and thus accounting for its success. In the first wave, the coordinating center composed of state agencies, as well as foundations closely tied to the US government, also provided the bulk of the funding. Additional funding and implementation partners came from the public sector in developing countries. In contrast, the contemporary network reflects a process of pluralization of funding sources (Babb and Chorev 2016), wherein a key role is played by a new set of private foundations. This difference is important not because private capital is nimble while state agencies are bureaucratic and slow, but because it significantly modified the political context in which RCTs operate, reducing the obstacles to randomization and the sources of substitution bias. It thus allowed the turning of RCTs into a functioning hinge between the fields of academic economics and development aid.

### What is being evaluated?

The second dimension of difference between the two waves is that contemporary evaluations are focused on small, short-term *interventions*, while the object of evaluations in the 1960s and 1970s were relatively large-scale, long-term social policy *programs*. By “small,” however, we do not mean the size of the sample, but the

character of the intervention. Typically, second wave RCTs do not aim to assess whether an overall policy works or not, but to evaluate the effect of one specific, limited “nudge” or “plumbing detail” of the policy (Duflo 2017) on a selected outcome of interest, e.g., effect of sending text message reminders on micro-finance clients’ repayment rates (Karlan et al. 2014; Duflo et al. 2007, p. 3; Ravallion 2009). As Duflo (2017, p. 4) says, the interventions tested often involve “taking care of apparently irrelevant details, such as the way the policy is communicated or the default options offered to customers,” or they involve “logistical decisions” that hitherto had been treated “purely mechanically.” Samples of second wave RCTs, therefore, can be quite large (e.g., 75 elementary schools, 30,000 students [Miguel and Kremer 2004]), precisely because the intervention itself is limited. The sample size in these cases is more a function of the improved ability to run surveys than an indicator of the scope of the study.

We compare, therefore, the first and second waves not in terms of sample size, but the type of intervention and its duration. In 29 of 63 studies in our second wave sample, what is being tested is not the policy itself, but a specific “nudge” meant to overcome a cognitive or behavioral obstacle to policy uptake (Berndt 2015, p. 8; Duflo 2017, pp. 4, 11). In 19 other studies, the intervention deals with a “plumbing detail” or “design of the tap” problem by providing information or brief training Duflo (2017, pp. 4–5). In 15 other studies, the intervention remains small by piggybacking on an existing government or NGO program.

When the intervention is limited in this way, its duration is quite short and can range from a 5–10 min meeting with a microcredit officer to a 2–3 hrs financial literacy workshop, typically accompanied by follow-up measurements of the financial behavior of “the poor” for a few months (Drexler et al. 2014). As can be seen in Tables 6, 52.6% of second wave studies belong to this category, where the duration of intervention is no more than 1 month (in fact, 22 of these studies, or 35% of second wave sample, involve merely a few hours of training, workshop, watching videos, etc., during a one-day visit). This was true for only 4.9% of first wave studies. Similarly, while there were 38.6% of second wave RCTs that lasted a year or more, the corresponding figure for the first wave is 63.4%, including studies that extended for 8 or 9 years.

The longer duration of first wave studies goes a long way towards explaining why randomized assignment was not as widespread as in the second wave, and why it was ultimately abandoned. Having to contend with the continued resistance of the aforementioned Salvadoran administrators, Hornick et al. (1973) report that “failure of the randomization procedure undermined the validity of the experiment, while

**Table 6** Duration of intervention in first and second waves of RTCs

	First wave (%) <sup>a</sup>	Second wave (%) <sup>b</sup>
1 month of less	4.9	52.6
2–12 months	31.7	8.8
1 year or more	63.4	38.6

<sup>a</sup> Percentage from 41 studies

<sup>b</sup> Percentage from 57 studies

administrative difficulties affected the comparison (...) With the failure of the experiment, a less rigorous design was adopted.” “True” RCTs in our first wave sample had an average duration of 16 months, while the average duration of field experiments without random assignment was 29 months. A review of 96 family planning experiments from the same period found a similar trend, implying the presence of substitution bias: “comparability [*between control and treatment groups*] decreases with the passage of time. The longer the time span the greater the likelihood that other factors will intrude. (...) In such cases the advantage of having controls is greatly diminished” (Cuca and Pierce 1977, p. 35).

Not only were experiments in the 1970s much longer, but more importantly they attempted to evaluate whole delivery systems in broad social policy areas. Often, they compared not intervention with no-intervention, but different levels of intervention to determine which is most cost-effective. For example, a study by the Population Council (1986) had physicians travel to remote community clinics to insert IUDs, provide gynecological services, and treat clients with reported side effects from contraceptive use. The study randomly assigned clinics to 1, 2, or 4 physician’s visits per month to determine the optimal level of treatment. It ran for several years and provided reports every 6 months.

Likewise, the geographical scope of first wave experiments was much larger than contemporary ones. It was common to conduct experiments that took place over the entire national territory of countries such as Barbados, Nicaragua, and Taiwan. One study of family planning conducted in India employed 6500 field-workers who visited 2.4 million households in 28,000 villages and towns (Cuca and Pierce 1977, p. 123). The largest second wave RCT, in contrast, involved 20,858 students in 386 schools (Borkum et al. 2012).

Put differently, the expertise networks in the two waves of RCTs differed not only in terms of the actors involved, but also in terms of what actually was being evaluated by RCTs. In the first wave, the object of evaluation was a wholesale program of long duration. Such evaluations required a significant component of monitoring, answerable to Federal criteria originally established for domestic programs, bringing into play questions of implementation, maintenance, fatigue, equity, and political considerations (Berk et al. 1985). In practice, this meant that the experiments were expected to provide information about the long-term “impact” of the program, and thus had to contend with the legal and political barriers that the program might have to face in future expansion. The question of “operational validity”—namely, how to replicate the experiment’s results “in a larger environment ... includ[ing] the issues of resource requirements and acceptability of the experimental approach on a wide scale”—was front and center (Cuca and Pierce 1977, p. 7).

In the second wave, in contrast, the object of evaluation is typically a short-term intervention, explicitly striving to be “clever,” i.e., well-bounded and easily measurable. “The interventions are designed to answer a specific practical problem in a specific context; for example, how to get teachers to come to school more often, how to help farmers to save more, how to convince parents to get their children immunized” (Duflo 2006, p. 3). There is comparatively little interest in how the intervention might be “scaled up,” often explicitly leaving it to other actors, especially governments, who may or may not choose to get involved. This contrast in the duration and scope of RCTs provides additional insight into why RCTs were discontinued in the

first wave, but are thriving in the second wave. It means that the problem we identified earlier—the political resistance to randomized assignment—was less significant in the second wave.

The need to evaluate large scale, long-term programs during the first wave—itsself a function of the inter-governmental nature of development aid—ultimately meant that randomization had to be abandoned or marginalized in favor of long-term monitoring of implementation. First wave experimenters, therefore, were told that “not only the content, but also the methodology must be adaptable to a variety of environments.... Although ideally the design should conform to the requirements of a true experiment, conditions might have to be modified for certain purposes ... where a quasi-experimental design might well suffice” (Cuca and Pierce 1977, pp. 12–13). There is no justification, argued Dennis and Boruch (1989, p. 301), “for evaluation in general or to randomized field experiment in particular, unless the results are likely to be useful.” This often meant tailoring the design of evaluation to “meet the ethical demands of the setting” and conducting a randomized experiment only if certain “threshold conditions” were met. Since “their objective was to inform policy,” researchers in Barbados “did not use a no-treatment control group” because it was “politically inappropriate” (ibid., p. 302). Ultimately, this is why at the end of the first wave, RCTs were not so much discontinued, but came to be understood as “research” rather than evaluation. Not being able to overcome the political resistance to randomized assignment meant that RCTs could not function as the “hinge” between the academic and development aid fields and were assigned a mostly academic (“research”) value. From the point of view of key actors in the field of development aid, “rigorous field trials” were perceived as taking too long and encountering too much resistance, so priority was given to “quick turnaround, cost-effective evaluation tools” that could be employed in continuous monitoring of programs as they were being implemented (USAID 2009, pp. 16–18).

Moreover, the very nature of the alliances among USAID, the Population Council, and the governments of developing countries meant that at stake in the results of field experiments was the larger question of the role of foreign aid in international development. Hence the switch, over time, toward emphasizing implementation research and long-term follow-up, and away from randomization.

*Randomistas* in the second wave face the same problems whenever their studies take too long. A study of microfinance in India became messy over time as many people in the control areas took loans offered by competing microfinance companies. After three years, “competitors were everywhere,” and Duflo had to settle for an imperfect measurement strategy (quoted in Parker 2010). Most of the time, however, *randomistas* address the potential resistance to randomized assignment by *going small and short-term*. When the intervention is small—a text message or a “free cooking stove”—it is relatively trivial to persuade people to participate in a control group, and the “researcher might, with a clear conscience, randomize the order in which people are supplied” (ibid.).

An additional benefit of RCTs being marketed as answering small, well-defined questions is that one is able to avoid the controversies regarding the long-term advisability of development aid. Local NGOs who consent to participate in RCTs, do so because short-term interventions seem less politically risky than full-fledged programs. Hence, family planning field experiments are a tiny proportion of J-PAL’s portfolio, since they tend to be politically fraught, raise suspicions regarding the overall aim of development aid, and require a long-time horizon. When they fail, they call into

question the whole recent history of development aid. By comparison, short-term nudges do not bring into view the overall role that NGOs and private foundations play in the development aid matrix. The main implication of a negative finding is that you should try again by varying a different aspect of the “behavioral game” being played (Banerjee and Duflo 2011; Banerjee et al. 2015).

If challenged about the limited scope of their experiments, *randomistas* counter that they are not only meant to find out what works and what does not, but also to “test economic theory.” “We’re not just evaluating a program and attaching some larger thoughts at the end, but finding ways to use evaluation to explore theories” (Duflo, in Parker 2010). Accordingly, a recent review of the use of experiments in economics states:

This current generation of field experiments oftentimes has more ambitious theoretical goals than social experiments (which largely aim to speak to policymakers); modern field experiments in many cases are designed to test economic theory, collect facts useful for constructing a theory, and organize data to make measurements of key parameters.... (Levitt and List 2008, p. 19)

Paradoxically, what allows contemporary projects to be short is precisely this talk about a “more ambitious theoretical agenda”. We should take it with a grain of salt. It is not more ambitious, but *differently wired*. *Randomistas* avoid the political debate about development by limiting themselves to testing behavioral hypotheses that often could be quite innocuous (“if you give textbooks to school children, they get better grades”). What makes them “more ambitious” is that they are connected back to the discipline of economics, not to a social program’s bureaucracy. *Randomistas* are not concerned about the limited and short-term nature of RCTs, because they consider each evaluation as merely one little piece in a greater puzzle, part of a slow accumulation of knowledge. Ultimately, they promise, the knowledge accumulated will be brought back to influence policy. What economists “bring to the table,” is not only evaluation expertise, but “prior evidence and theories that help them to predict what should work” (Duflo 2006, p. 3). This means, essentially, that whether the field experiment is considered part of theory building, thus appealing to disciplinary audiences, or is understood to elucidate a “plumbing detail,” thus appealing to policy-makers and NGOs, is kept strategically ambiguous (Banerjee et al. 2015, pp. 21–22; Rayzberg 2019). In this way, and unlike the first wave, the small, short-term RCT evaluating a policy/theory-relevant “nudge” is able to function as a hinge, linking the fields of academic economics and development aid. It permits foundations and NGOs to legitimize their selective giving by its “measurable impact,” while academic economists are able to frame their results as relevant to disciplinary concerns.

## The homologous transformations of development aid and economics

In this section, we draw on published accounts of the histories of international development and academic economics to describe the parallel transformations in these two fields. As each field underwent a crisis that destabilized its status quo, an opening was created for newcomers to mobilize RCTs as a heterodox strategy challenging the

field's orthodoxy. Building on the conceptual framework we developed earlier, we show how the homology between the struggles conducted in the two fields created an elective affinity between the positions and strategies of the *randomistas* and the philanthro-capitalists, leading them to settle on limited, short-duration RCTs as the hinge linking their parallel struggles. The limited and short-term nature of these RCTs, as we saw, minimized the political resistance to randomized assignment, while providing dual rewards in the two fields: apparently objective and effective short-term solutions to the problem of triage in foreign aid, combined with long-term opportunities for theory building in development economics. In this way, RCTs became a hinge durably linking the two fields.

To be clear, there were institutionalized links between the two fields prior to the rise of RCTs (Rayzberg 2019). There was circulation of personnel between the two, with recent PhDs employed as applied economists in the World Bank and similar organizations, and former World Bank officials appointed as professors. There were organizational sub-units, which replicated the format of units from the other field (e.g., the Development Research Group at the World Bank mimicking an academic department; academic institutes combining research with policy-oriented activities).<sup>9</sup> There were multiple forums where a robust discussion about development was conducted among experts from both academia and aid organizations, who shared a body of expert knowledge acquired through common training, constituting what Porter (1995) calls “disciplinary objectivity.” Yet, as concurrent transformations destabilized both fields, they also destabilized the authority of disciplinary objectivity, thus threatening to rupture the most important link between them. This was the context in which a turn to mechanical objectivity, represented by RCTs, could become a profitable strategy for sub-groups in both field, and could serve as a “hinge” linking their parallel strategies.

### The fragmentation of development aid

At any point in time, the field of development aid is composed of all actors involved in designing, implementing, and funding development aid projects. These include foundations and their personnel, as well as NGOs, national and local governments, multi-lateral organizations, etc. The key argument of this section is that, while during the first wave of RCTs (1960–1970), the field was organized around a dominant coalition, it has become fragmented during the second wave, and it is this fragmentation that serves as a condition of possibility for the success of second wave RCTs. In the earlier period, the field of development aid was composed mostly of bilateral (e.g., USAID) and multi-lateral (e.g., World Bank, UNESCO) organizations together with national governments in developing countries (Krueger 1995). Private foundations, such as the Rockefeller Foundation, were relevant actors, but as we saw, they essentially acted as auxiliary arms of US foreign policy (Heydemann and Kinsey 2010, p. 222). Thus, the field centered on the dominant alliance among US agencies, the old foundations, and national and local governments in the developing world. In the second wave, in contrast, the field

<sup>9</sup> These replicas of a unit from one ecology (a profession, a discipline, a political party,) within another are what Abbott (2005, pp. 265–269) calls “avatars.” For example, applied economics is an avatar of academic economics in the “policy and advice arena.” While an avatar is “an institutionalized hinge,” over time it can become independent and even compete with its original creators.



became fragmented, as the percentage of projects by the dominant alliance diminished in favor of a new type of private foundations (the philanthro-capitalists), global NGOs, and new country donors.

The fragmentation of the field of development aid was a complex process, which involved not only the entry of new actors and the relative weakening of old ones, but a related transformation of the “development imagination” and a loosening of the previously tight linking among resources, norms, and ideas that characterized the field in the era of the “Washington Consensus” (Rodrik 2006; Babb and Chorev 2016). By the late 1990s, the increasing attention to the problems of bad governance and corruption among recipient governments polarized the field between those in an “optimist” faction, led by Columbia University professor Jeffrey Sachs (2005), who were still proposing large-scale, internationally coordinated efforts to combat global poverty; and others in a “pessimist” faction, led by NYU professor and former World Bank official, William Easterly (2007), who argued that development aid does more harm than good, and that poverty can be tackled only by returning agency to poor countries. This was more than just “aid fatigue.” It signaled a profound crisis of disciplinary objectivity, widespread skepticism among development experts about their own ability to guide development aid. Some even proposed to end development aid all together (Moyo 2009).

This weakening of the orthodoxy organized around ODA and disciplinary objectivity served as an opening for other actors—global NGOs, new country donors, and the philanthro-capitalists—to enter the development aid field in the late 1990s, as evidenced by the rapid growth in the relative share of private international assistance (Babb and Chorev 2016). No less importantly, it allowed the philanthro-capitalists to play the role of a heterodoxy and to exert an influence far beyond their actual share in total disbursement. The pessimist critique has led donors (including governments and multilateral organizations) to channel greater flows of aid to new global NGOs with a local footprint, and to the private sectors of developing countries, both of which became natural allies of the philanthro-capitalists. In contrast to the development enterprise in the 1960s and 1970s, donors increasingly rely on private transnational groups as contractors and intermediaries, including for-profit development organizations (Watkins et al. 2012)—USAID and the EU, for example, disbursed 30% of their budgets through private for-profit groups in the year 2000 (Cooley and Ron 2002).

This fragmentation is significant less for the actual amount of money disbursed by philanthro-capitalists,<sup>10</sup> than for its impact on what Babb and Chorev (2016) call the “development imagination,” tilting it in the direction of business-oriented norms (quick turnaround, measurable results, and professionalized aid management). These new norms are emphasized not only by the philanthro-capitalists, but also by government agencies and multilateral organizations (Adams 2016). The World Bank, to cite one example, has been rebranded by its current President as a “development consultancy

<sup>10</sup> The existing estimates are not yet reliable enough to determine the relative share of development aid coming from private foundations. The Development Assistance Committee (DAC) estimates that, since 2002, aid to developing countries from Private Voluntary Organizations had been three times as large as ODA. This number, however, includes not just direct development investments, but also private bank lending and remittances, so it is an over-estimate. More accurate estimates are likely in the future since in 2010, the Gates Foundation became the first private aid donor to report to the DAC, encouraging other foundations to do the same (OECD 2011, p. 4).

agency,” and enjoined “to think of ourselves now as strategic advisors, honest brokers who link capital looking for a greater return to countries looking to achieve their higher aspiration” (Kim 2017). Similarly, many United Nations agencies, USAID, and UK’s Department of International Development have invested heavily in new Monitoring and Evaluation Units and development indicators (Babb 2009; Watkins et al. 2012).

This transformation of the development imagination is evidence that the orthodoxy—composed of the leadership of multilateral organizations and government agencies disbursing or receiving ODA, together with the development experts serving as their advisers—has lost its hold. The field became pluralized, as alongside and relatively independent of the inter-governmental aid that was dominant during the first wave, there emerged, as Krause (2014, pp. 4–5) suggested, a quasi-market where “the good project” is a quasi-commodity produced by global NGOs and consumed by private donors who demand “measurable results.” Consequently, “the pursuit of the good project develops a logic of its own that shapes the allocation of resources ... relatively independently of beneficiaries’ need” (ibid.). To overcome what Swidler and Watkins (2017) called the principal-agent problem of “altruism from afar,” or what Krause (2014) calls the “triage problem” of how to allocate limited resources among many competing needs, donors increasingly rely on issuing short-term, renewable contracts for discrete aid projects, requiring NGOs to bid competitively and to demonstrate concrete results (Berrios 2000). The result is a thoroughgoing transformation of the type of aid disbursed from inter-governmental *programmatic aid*, such as long-term funding for the whole education sector of a country, to *project aid*, given for a specific intervention with a short time frame.

These changes served as a hospitable environment for the second wave of RCTs. On the one hand, the polarization of the development aid debate invited a centrist strategy. This is indeed what the *randomistas* did by combining the pessimist critique of large developmental projects, with the activist attitude of the optimists. In this way, they handed further ammunition to, and solidified their alliance with, the “impatient optimist” Bill Gates and other philanthro-capitalists. Their championing of RCTs as a form of mechanical objectivity was predicated on problematization of the subjectivity of development experts and government bureaucrats (i.e., an attack on the disciplinary objectivity claimed by the orthodoxy), their tendency to become attached to big programs because of ideological preconceptions. According to Duflo (2011), ideology simplifies the causes of poverty and therefore dictates a preference for comprehensive programs rather than small-scale, tailored interventions: “Programs are often borne in ideology ... [that] the poor are entrepreneurs, or they are starving or they are slothful. [Programs] are conceived in ignorance of the reality of the field, and then they persist because once they exist there is a consistency for them to just continue. I think we need to fight against that.”

To counter this subjectivity, *randomistas* emphasized not only the mechanical objectivity of RCTs, but also their small scale and practical nature. The development debate, they said, is couched at the wrong level. The big, philosophical questions such as whether development aid is fundamentally helpful or not, or what the root causes of global poverty are, cannot be answered. The debate is futile and leads to “stagnation and inertia” (Karlan and Appel 2011, p. 5). What RCTs can offer, conversely, are answers to small, practical, and topic-specific issues (Banerjee and Duflo 2011, p. 13).

*Randomistas* thus offered RCTs as a means of bringing *closure* to the heated controversy in the development aid community over how to address global poverty. At the same time, they also offered a longer-term vision in which small, short-term studies generate a virtuous cycle of knowledge accumulation. The second point is, therefore, that the fragmentation of the development aid field and the emergence of the quasi-market for “good projects” offered an especially hospitable environment for RCTs, provided that they remained small and short-term. It is hard to overestimate the impact of this transformation on the scope for using RCTs. Given that the success of randomization is inversely correlated to the length of the experiment, the short-term nature of current projects, itself dictated by the episodic nature of funding in the quasi-market, emerges as a key condition of possibility for the success of the second wave. The inter-governmental ties characteristic of the first wave, in contrast, dictated a much longer timeframe. The fact that researchers are now working with NGOs “who did not have to pretend to serve everyone,” rather than governments, means that the political resistance to randomization is minimized: “if you strongly believe that the program you are running will benefit people, it would arguably be unethical to deny that program to some people in order to create a control group. But if you cannot serve everyone anyway, it is, again, arguably fairer to determine who is served via randomization” (Ogden 2016, pp. xx-xxi; see also Glennerster 2015).<sup>11</sup>

The different parts of the hinge now work together seamlessly linking the two fields. The RCTs provide the private donors and NGOs with precisely what they need to pursue their heterodox strategy in the field of development aid—“clear goals,” “measurable results,” a demonstration that they are being “effective altruists.” NGOs are encouraged to present evidence about the effects of the project according to narrowly delimited aims (Babb and Chorev 2016, p. 95), and are discouraged from taking into consideration the broader effects that a development intervention might have (Krause 2014). In contrast, during the first wave, evaluators could not abstract away from the broader, longitudinal impacts of foreign aid, because the goal was precisely to stimulate sectorial, macro changes over the long run (Sommer 1977; Freeman et al. 1980). RCTs’ promise to deliver an unbiased (i.e., untainted by expert judgment) measure of the impact of a specific intervention is key to translating and aligning the interests of donors and NGOs, while the insulation from political pressures that the latter provide guarantees that the *different* parts of the “hinge” are held together and the link between the two fields is durable. In contrast, any attempt to assess the longer-term effects of a development program is likely to drive them apart.<sup>12</sup>

<sup>11</sup> The problems faced by the Inter-American Development Bank (IADB), which works primarily with governments, underscore this point, demonstrating how hard it is for governments to randomize benefits. Despite explicitly requiring all loans to undergo impact evaluations, IADB was able to conduct RCTs only in 26% of its loans and had to resort to a quasi-experimental design in the remainder. IADB (2017) representatives reported that RCTs were seen as “imposed on country governments, which are reluctant to appropriate RCTs by themselves.”

<sup>12</sup> Pritchett makes the same point, though perhaps more bluntly: “The only people for which the RCT movement is in fact a tool for the job are philanthropists.... From the charity perspective, there’s a nice confluence between the methodological demand for statistical power and of being able to tweak at the individual level. I can give this person food, but not that person. (...) I’m not trying to affect the government; I’m not trying to affect national development processes.” (in Ogden 2016, p. 142)

## The consequences of anomie in academic economics

Having demonstrated how short-term RCTs figured into heterodox strategies in a destabilized field of development aid, let us move now to the concurrent set of transformations in the field of the American economics profession in order to highlight how RCTs figured into the homologous situation faced by young economists. In the late 1980s and early 1990s, the field of economics was undergoing an “empirical turn”: the end of the hegemony of theoretically-oriented formal modeling and the proliferation of empirically-oriented tendencies (Angrist et al. 2017), of which the most relevant for our purposes were the debates in development economics about causal identification and the rise of behavioral economics.

In development economics, the new, empirical mood appeared in the form of a shift away from its previous focus on macroeconomic theories of international trade, human capital, fiscal policies and their interrelations, towards an intensive interest—some characterize it as an “obsession” (Jonathan Murdoch, in Ogden 2006, p.51)—with the question of causal identification, namely “how ... to separate out the causal impact of a specific policy or factor from potential confounding factors” (Michael Kremer, *ibid.*, p.1). Multiple observers, both critics and adherents of RCTs, cite the ensuing debate as the formative context for the rise of RCTs (Lant Pritchett, *ibid.*, p.140; Deaton 2010). They describe a milieu, especially at Harvard and MIT, where young development and labor economists were trained to find an “instrumental variable” (IV), with which to neutralize confounding factors and thus pinpoint the cause with a high degree of confidence. At the same time, they were also trained to be highly skeptical of this exercise, as they “watched every empirical paper ... picked apart based on causal claims,” while some collaborated with statisticians to formulate an even more radical critique of IVs (Murdoch, *ibid.*, pp. 51–53). From our point of view, the crucial point about these methodological debates is that they disrupted business as usual and destabilized disciplinary objectivity. Observers describe an anomic situation, where “I can sit in a seminar in Cambridge and whatever instrument you propose, I can concoct a story in which your instrument is wrong.” (Pritchett, *ibid.*, p. 141). This anomic situation increased the uncertainties and pressures concentrated in the position of new entrants into the field (PhD students and young faculty members), disrupting the normal process of reproduction. Graduate students were “having to fight these battles over the validity of their instruments,” and were finding it harder to publish their work (Murdoch, *ibid.*, p. 55). In this context, RCTs offered a means of escaping this predicament, or in Murdoch’s terms, of “creating a new kind of instrumental variable. Not one that you stumble across. You create it. Carefully and deliberately (...). Randomization provides the golden ticket, the Holy Grail from an IV perspective. RCTs are a machine for creating credible instrumental variables” (*ibid.*).

RCTs, in short, are not the much-touted “gold standard,” they are a much narrower “golden ticket,” namely a hall pass that shields a young scholar from the destructive effects of an anomic situation where the consensus underlying disciplinary objectivity has collapsed. Similar to RCTs’ role in bringing “closure” to heated debates in the development aid field, they proved to be a very powerful tool for young economists to enter the booming empirically-oriented portion of the economics field (Angrist and Pischke 2010).

The parallel rise of behavioral economics in the 1990s, attacking the “neo-classical” *homo economicus* paradigm (Mullainathan and Thaler 2000), served as an additional impetus for the rise of RCTs. While initially unrelated to the causal identification debates, in retrospect it is possible to see similarities between the two movements. Both were revivals of longstanding disputes within the economics profession—behavioral economics harked back to Herbert Simon’s 1940s research program (Heukelom 2012); the causal identification debates echoed the debates in the Cowles Commission (Ogden 2016, p. 140). Both mobilized allies from outside the discipline—statisticians and cognitive psychologists, respectively—to challenge the orthodoxy of the field. Both had a strong academic base at Harvard University and MIT. Both represented parallel heterodox offensives to pluralize economics by opening this notoriously insular discipline to imports from other disciplines, thereby creating alternatives to the prevailing neoclassical orthodoxy (Santos 2011).

Both movements, finally, because they strove to pluralize the field, destabilized the established criteria by which new work was evaluated, thereby disrupting the normal process of reproduction. When behavioral economists criticized the lack of realism in standard micro-economic theory, they opened the proverbial “floodgates.” They too were vulnerable to the criticism that their investigations lacked realism, because they were conducted in the “fake” environment of experimental labs (Guala 2007). At the same time, they were vulnerable to the counter-attack of micro-economic theorists that their research is purely descriptive, without theoretical value, since it is unable to explain how markets operate despite the limited rationality of the participants. These critiques of behavioral economics provided an opening for the *randomistas*. They offered young scholars a research paradigm that could claim to be far more realistic than both laboratory studies and microeconomic theory. At the same time, they touted the theoretical contribution of their field studies, designed to shed light on how “the poor” in developing countries actually make decisions in natural settings and to demonstrate how nudges and ecological features can increase the rationality of their decision-making (thereby demonstrating their relevance to economic theory) (Berndt 2015).

Developing countries could then be seen as ideal testing grounds for some of these theories.... There may be more to learn about human behavior from the choices made by Kenyan farmers confronted with a real choice than from those made by American undergraduates in laboratory conditions. (Duflo 2003, p. 9)

The two movements—causal identification and behavioral economics—intersect, for example, in Esther Duflo. She was introduced to work on “natural experiments” by one of her dissertation advisors, Joshua Angrist (Parker 2010), a key figure in the causal identification debates. At the same time, she drew on the work of behavioral economists to test whether cognitive biases are responsible for poverty traps (Duflo et al. 2011; see Berndt 2015, p. 8). This contingent combination shaped the distinctive character of second wave RCTs and contributed to their success, because it provided experimenters with a toolkit of small, short-term intervention—construed as “nudges”—while allowing *randomistas* to frame these experiments as contributing to economic theory. Young scholars conducting field experiments could defend their work as realistic, empirically rigorous, and theoretically relevant: “You’re not just learning

about what this particular program does in this particular place, but understanding human behavior better” (Rachel Glennerster, quoted in Parker 2010).

When the orthodoxy of the field counterattacks, pointing out the lack of external validity of RCTs and the limited theoretical value of the results (Nobel Laureate Angus Deaton went as far as to liken the *randomistas* to researchers testing “the idea that parachutes are useful to people who jump out of planes”), the response of the *randomistas* is extremely telling. When speaking to disciplinary audiences, they readily admit that randomized assignment does not eliminate all the sources of error in causal inference and that it is not an optimal research strategy if the expert “places little weight on persuading her audience.” Yet, they say, if the expert is faced with an audience consisting of “stakeholders with veto power ... whose priors may diverge from ... [one’s] own,” and if it is of paramount importance to persuade this audience, then the need to communicate as little “bias” as possible outweighs other considerations. In this situation, when experts are faced with “an adversarial audience who may be able to veto [their] choices,” and who is leery of the reliance on expert judgment, then “randomized experiments allowing for prior-free inference become optimal” (Banerjee et al. 2016, pp. 2, 11–15).

One cannot fail to hear here an echo of their formative experiences, during which they encountered the “adversarial audience,” sitting “in a seminar in Cambridge.” Yet, it is also clear that in the present context, the combative audience they now had in mind were the philanthro-capitalists. In essence, they were justifying their preference for RCTs by reference to the mistrust of experts prevailing, as we saw earlier, in the field of development aid where, in the wake of the pessimist critique and collapse of the Washington Consensus, there is “loss of hope in development” and lack of trust in traditional development expertise (Krause 2014, p. 42). If you want to build an enduring link to this other field, they seem to be saying, you must take into account that your project has to be sold to donors who now “trust in numbers” (Porter 1995) much more than they trust in expert judgement; who in fact perceive the theories and experience of experts to constitute “bias.” RCTs, as a strategy that minimizes bias, is best suited to build this link. This defense of RCTs could not have been formulated during the first wave, when there was relative optimism about the possibilities of development, about the power of multilateral organizations to do good, and about the cogency of development expertise (Krueger et al. 1989).

Facing towards the field of development aid, RCTs served to translate and coordinate the interests of the new coalition with private foundations and global NGOs, while promising to bring closure to polarizing controversies and restore the objectivity of development expertise. Facing towards the economics profession, field experiments allowed young economists to navigate the anomic situation they faced, while parrying the twin accusations of unrealism and pure description. By certifying that a certain intervention “works,” RCTs reassure private foundations and global NGOs that the project—the quasi-commodity that links them—has produced a measurable difference, thus validating their exchange, while leaving questions of “scaling up” for later. They also reassure young economists that they are contributing to disciplinary knowledge. What they no longer do, however, is precisely what was meant in the past for a program to “work,” namely evaluate whether it could be implemented on a large scale, taking into account political and administrative constraints (Dennis and Boruch 1989, pp. 301–302).



## Elective affinities

The objective basis of alliance between the *randomistas* and philanthro-capitalists, as the previous section established, is the homology between their positions and strategies in their respective fields. They were both newcomers, a heterodoxy gearing to change business as usual. The subjective basis of their alliance, however, is a set of values and images, which function as elective affinities attracting the two sides to one another (Eyal 2000). The ascetic values and image of the virtuous expert encoded by the mechanical objectivity of RCTs are intelligible by way of contrast with the prevailing ethos of their opponents, the orthodoxies of their field, whose approach to development aid was thereby problematized and rejected as subjective and biased. By the same token, these images allow the two groups to recognize each other as natural allies.

i) *Trust in Numbers*: the most obvious affinity linking *randomistas* and philanthro-capitalists is their emphasis on *measurement*. It clearly expresses the most important factor that brought them together, namely their parallel attacks on the orthodoxies of their fields, problematizing expert advice as subjective and biased. Bill Gates's 2013 annual letter on behalf of the Gates Foundation was titled "Why Measurement Matters" (Gates Foundation 2016). It was a passionate argument that "You can achieve amazing progress if you set a clear goal and find a measure that will drive progress toward that goal in a feedback loop." This image of a feedback loop from measurement to policy is noteworthy for what it leaves out—theory, experience, expertise, the knowledge that the development aid community has presumably already accumulated—in favor of what is presented as an almost mechanical *tâtonnement*, trial-and-error (the analogy to machine learning seems apt). This preference for measurement is premised on a problematization of theory or expert judgment as "bias," preconceived, subjective opinion. The virtuous expert should have none: "One of my great assets of being in this business, or maybe I've developed it over time, is I don't have many opinions to start with [...] I have one opinion—one should evaluate things—which is strongly held. I'm never unhappy with the results. I haven't yet seen a result I didn't like" (Duflo, quoted in Parker 2010). There is clear affinity between Duflo—who has acquired in the causal identification debate the ascetic predisposition to trade-off precision for a virtuous, unbiased measure—and Gates, who plays the role of the adversarial, skeptical audience that wants nothing of the expert's priors, only this unbiased measure that can "drive progress" by itself.

This trust in numbers and mistrust of experts is dramatized and celebrated in a set of institutional rituals—TED Talks, conference, and award nights—that bring the philanthro-capitalists and *randomistas* together. They present themselves as disruptors who are challenging the orthodoxies of their fields, while reaching out to allies across boundaries. The *randomistas* present themselves as critics of armchair economics: "to find out what works on the ground, you need to climb down from the ivory tower and do some serious legwork in the places you are trying to help" (Duflo, quoted in Benko 2013). Thinking of economic research as hands-on accords with the worldview of philanthro-capitalists like Gates, who is critical of academics "doing nothing more than

teaching two classes a semester and pumping out armchair advice in academic journals” (Gates 2014). The philanthro-capitalists similarly present themselves as poised to “revolutionize philanthropy, making nonprofit organizations operate like business, and creating new markets for goods and services that benefit society” (Edwards 2017, p. 1). Thinking about philanthropy as business, as rational economic action, comports with the worldview of economists. Ultimately, however, it is their emphasis on measurement that makes them welcome guests in these forums. It is the *raison d’être* of outfits such as the Institute for Effective Altruism (IEA) and *GiveWell*, which draw on the results of RCTs to publish shortlists of recommended NGOs to which donors should give. The donors can “see for themselves what works,” no longer needing to rely on suspect expert advice. Or, seen from the other side, these outfits have turned donors into this “adversarial audience,” to whom only the *randomistas* are now linked. Unsurprisingly, topping the shortlists are the poster children projects of J-PAL (Givewell 2017). Invited guests include foundation leaders, J-PAL members, and allied behavioral economists such as Cass Sunstein, whose talk title “From Behavioral Economics to Public Policy” neatly captured the alliance between economists and donors.

ii) *Leverage*: while measurement cuts the expert-qua-advisor out of the science of development aid, their joint enthusiasm for “leverage” demonstrates how *randomistas* and philanthro-capitalists premise their alliance on cutting out government planners as well. In *randomistas’* eyes, attachment to “programs”—large-scale, comprehensive, one-size-fits-all, typically government-administered programs—reflects ideological bias. The virtuous expert, in contrast, is agnostic, targeted, and clever, approaching each situation anew (“without many opinions to start with”), looking for the right lever of change. There is a strong affinity between the strategy of “going small” practiced by *randomistas*; the “nudges” favored by behavioral economists; and the idea of “leverage” dear to philanthro-capitalists. The latter distinguish themselves from the older foundations, whose management style they criticize, precisely by emphasizing their concern to deploy “philanthropic resources more strategically,” taking a “market conscious” and “impact oriented” approach to giving “driven by the goal of maximizing the ‘leverage’ of the donor’s money” (Bishop and Green 2008, pp. 6, 152). They do not consider themselves donors, but “social investors” or even “venture philanthropists” (Frumkin 2003). This means the opposite from funding wholesale social programs. It is “a portfolio approach, experimenting with lots of different ideas that, if successful, might be scaled-up by other institutions, including governments” (ibid.). Leaving implementation and scale-up to others, means that the virtuous expert no longer occupies the same position as the government planner. Clearly, “leverage” is meant here not in the financial sense, as when a small amount of money brings in much more, but in the sense that this small amount of money, correctly invested, could ultimately have a much larger impact, either because its success as proof of concept attracts other actors who will scale it up or because it is targeted at a “lever” of change.

*Randomistas* and behavioral economists cast themselves as the appropriate partners sought by philanthro-capitalists by offering leverage in both these senses. RCTs, with their promise to provide evidence on “what works,” have

the power to certify (and de-certify) projects, thereby attracting more donors and acting as levers for scaling-up.<sup>13</sup> More interestingly, the conceit of “nudges” is that people can be greatly influenced by small changes in context, such as which food items are at eye level on a shelf (Thaler and Sunstein 2009, pp. 1–2). Small actions creating a big impact is the textbook definition of “leverage.” Nudges are levers of change. As Gates (2011) says in his review of Banerjee and Duflo’s *Poor Economics*: “small tweaks can sometimes turn failing interventions into effective ones.”

iii) *Libertarian paternalism*: the third and final affinity between the worldviews of the *randomistas* and the philanthro-capitalists is clearly shaped by their joint centrist strategy in the development debate. Nudges and similar behavioral interventions provide leverage because they activate the autonomous, self-organizing powers of the actors themselves, especially their power of choice. Nudges, thus, signify a certain combination of the technocratic rationalism of the “optimists,” tempered with respect for the autonomy of the actors involved, characteristic of the “pessimists.” It is precisely this combination that Thaler and Sunstein (2009, pp. 4–6) aim at with the seeming oxymoron of “libertarian paternalism”: “We argue for self-conscious efforts, by institutions in the private sector and also by government, to steer people’s choices in directions that will improve their lives.” This combination is calculated to appeal to foundation leaders, who would like to be “effective altruists,” yet are worried about government intervention and coercion. Nudges offer them the possibility to exercise “gentle power” (ibid., pp. 8, 11). Behavioral economics is thus offered as a liberal art of government, an art of leading people gently towards rational and self-interested action (Berndt 2015).

The preference for nudges over commands implies a vision of the virtuous expert as a “choice architect” (ibid., p. 3). This vision, which originated in behavioral economics (Santos 2011), encapsulates the centrist strategy of *randomistas* and philanthro-capitalists. On the one side, there is the paternalism of “programs” and the social planner, which they attribute to the heyday of the older foundations and to an older generation of social scientists. On the other side, there is the unbridled libertarianism of free market reforms and of the theoretical micro-economists, who accord the expert the role of constitution-giver and mechanism-designer (Eyal 2000). Choice architects, by contrast from social planners, do not tell people what to do. Yet, unlike mechanism-designers, they do not have as much faith in the power of the market to provide actors with the best incentives to choose rationally. Thus, to induce Kenyan farmers to save in order to buy fertilizer later, Duflo et al. (2011, p. 2353), demonstrate that “a paternalistic libertarian ... approach of small, time-limited discounts could yield higher welfare than either laissez-faire policies or heavy subsidies.” Behavioral economists *qua* choice architects work to augment the rationality of economic actors, so as to direct the market process towards collectively rational and equitable goals. RCTs serve to test which organization in the context of

<sup>13</sup> For example, total donations to the four charities shortlisted by *GiveWell* increased from \$3,000,000 in 2010 to \$110,000,000 in 2015. Conversely, when RCT evidence is inconclusive, projects lose their recommended status and funding, as happened to Development Media International (GiveWell 2017).

choice works best to “nudge” individuals towards more rational choices, without eliminating their freedom of choice.

There is an affinity here with the worldview of foundation leaders who hail from the information technology sector: “choice architects can make major improvements to the lives of others by designing user-friendly environments” (Thaler and Sunstein 2009, p. 11). What else are software engineers doing but “designing user-friendly environments” that imperceptibly guide the users towards certain choices? The metaphor of architecture is pervasive in the information technology world, and it is apt. A building (like a computer interface) channels the choices of its users by its physical structure. The likelihood of certain choices is “programmed” into the building by its architect. If choice is inescapable and indeed desirable, the alliance of *randomistas* and philanthro-capitalists seems to say, you might as well do it in a “smart” way, you might as well “hack” the process to make it more user friendly and to favor rational choices.

## Conclusions

In this article, we are arguing that the contemporary success of RCTs is better understood as a product of historical and institutional processes that have changed the political and scientific context in which RCTs are implemented, rather than as evidence of their “gold standard” quality. By jointly mobilizing the concepts of “hinge” and “homology between fields,” we show how the fragmentation of the development aid field and changes in the economics profession made RCTs answerable to new audiences and allowed *randomistas* greater leeway to bypass the political resistance to randomization. The nudge-type, short-term RCTs serve as the “hinge” for the alliance between *randomistas* and philanthro-capitalists and therefore enjoy broad appeal. In our concluding remarks, we would like to address both practical and conceptual implications of our findings.

A recent controversy about an RCT evaluating the educational benefits of deworming in Kenya (Miguel and Kremer 2004) can illustrate the advantages and limitations of “going small” for international development research and practice. In this RCT, *randomistas* provided deworming medication for kids in treatment schools. The intervention, therefore, was short-term and did not address the entrenched problems of educational underfunding or limited labor market prospects. It was rather a sort of “hack” of the development process that was easy to administer and measure. Initially, researchers reported improvements in test scores and school attendance not only for treated children, but also spillover effects for kids that did not receive treatment. This led *randomistas* to prescribe deworming medication as a cheaper and more effective solution to improve educational outcomes in poor countries—a prescription that was followed by multiple donors and that had affected 200 million children by 2015 (JPAL 2016). Unfortunately, in 2014, after the data were made available online, multiple replications invalidated the main finding and led to heated debate about the use of RCTs to test and prescribe social policies (see Humphreys 2015).

From our point of view, however, the main implications of this controversy in particular and of the widespread use of RCTs in general is that “going small” in

development is often tantamount to “searching under the street light,” i.e., to evaluating only what can be easily and quickly administered and measured. Do we really need an RCT to know that if children are less sick, they are more likely to go to school, and less likely to get other kids sick? Evaluating this sort of intervention can generate impressive results and cement the alliance with global foundations interested in demonstrating the measurable impact of their giving. Yet, it also avoids the much harder questions, such as the underfunding of educational systems in developing countries, or how to motivate poor children to go to school when their labor market prospects are dim. In other words, the “targeted” interventions characteristic of second wave RCTs obscure the harder and more politically fraught task of addressing the complex mechanisms reproducing poverty or of assessing the overall impact of policies.

The conceptual toolkit that we use to arrive at these conclusions also has broader implications that extend beyond the study of RCTs. So far, sociologists studying how distinct fields become durably connected to each other tended to place either the strategic action of actors or the structural constraints of fields at the center of their theories, instead of thinking about how the two interact and condition one another given the contingencies of the historical case at hand. Typically, a focus on structural constraints, especially in field theory, tended to underestimate the possibilities for creating a durable link; while a focus on the strategic action of actors underestimated the obstacles presented by the contrasting logics of the two fields. By combining the concept of “hinge” with field analysis, however, we have provided other sociologists with a more balanced approach to the question of how fields become durably linked. Specifically, we are arguing that an adequate answer must demonstrate not only that a strategy offers “dual rewards” in both fields, but also how it overcomes the resistance stemming from the tensions between them. No less importantly, we argue against imputing undue causal weight to the interests and consciously formulated strategies of the actors involved, while we argue in favor of a relational analysis of how “strategies without a strategist” are formed and become coordinated in the course of conducting parallel struggles in homologous fields. Our article is geared to demonstrate the utility of historical comparison in pursuing and validating this explanatory approach. This approach provides a way to study connections among fields and the emergence of policy-oriented expertise without overloading the causal significance imputed to the strategic intentions and “social constructions” of the actors involved. Through this demonstration, we aim to stimulate more research in this direction.

**Acknowledgements** The authors want to thank XingJian Li for her meticulous work in this project, as well as Margarita Rayzberg, Diana Graizbord, Moran Levy, Joan Robinson, Josh Whitford, and Diane Vaughn for their feedback. Previous versions of this article were presented at the 2014 Society for Social Studies of Science Meeting, the 2015 Social Science History Association Meeting, the 2016 and 2018 American Sociological Association Conference, Columbia University’s SKAT workshop, Sciences Po, and the Federal University of Rio de Janeiro. We also thank the participants at these presentations for their helpful comments and suggestions.

## References

- Abbott, A. (1988). *The system of professions: An essay on the division of expert labor*. Chicago: University of Chicago Press.
- Abbott, A. (2005). Linked ecologies: States and universities as environments for professions. *Sociological Theory*, 23(3), 245–274.

- Adams, V. (2016). *Metrics: What counts in global health*. Durham: Duke University Press.
- Angrist, J., & Pischke, J. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2), 3–30.
- Angrist, J., Azoulay, P., Ellison, G., Hill, R., & Feng Lu, S. (2017). Economic research evolves: Fields and styles. *American Economic Review*, 107(5), 293–297.
- Babb, S. (2009). *Behind the development banks: Washington politics, world poverty, and the wealth of nations*. Chicago: The University of Chicago Press.
- Babb, S., & Chorev, N. (2016). International organizations: Loose and tight coupling in the development regime. *Studies in Comparative International Development*, 51(1), 81–102.
- Banerjee, A. (2007). *Making aid work*. Cambridge: MIT Press.
- Banerjee, A., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. United States: Public Affairs Book.
- Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1), 1–21.
- Banerjee, A., Chassang, S., & Snowberg, E. (2016). Decision theoretic approaches to experiment design and external validity. *NBER Working Paper 22167*.
- Barnes, B., Bloor, D., & Henry, J. (1996). *Scientific knowledge: A sociological approach*. Chicago: University of Chicago Press.
- Bauman, K. (1997). The effectiveness of family planning programs evaluated with true experimental designs. *American Journal of Public Health*, 87(4), 666–669.
- Benko, J. (2013) The hyper efficient, highly scientific scheme to help the World's poor. *Wired Magazine*. Retrieved on May 2, 2017, <https://www.wired.com/2013/11/jpal-randomized-trials/>. Accessed 17 Oct 2018.
- Berk, R., Boruch, R., Chambers, D., Rossi, P., & Witte, A. (1985). Social policy experimentation: A position paper. *Evaluation Review*, 9(4), 387–429.
- Berndt, C. (2015). Behavioural economics, experimentalism and the marketization of development. *Economy and Society*, 44(4), 567–591.
- Berrios, R. (2000). *Contracting for development: The role of for-profit contractors in U.S. foreign development assistance*. Westport: Praeger.
- Bishop, M., & Green, M. (2008). *Philanthrocapitalism: How the rich can save the world*. New York: Bloomsbury Press.
- Borkum, E., He, F., & Linden, L. (2012). The effects of school libraries on language skills: Evidence from a randomized controlled trial in India. In *NBER Working Paper 18183*. Cambridge: National Bureau of Economic Research.
- Boruch, R., McSweeney, J., & Soderstrom, J. (1978). Randomized field experiments for program planning, development, and evaluation—an illustrative bibliography. *Evaluation Quarterly*, 2(4), 655–695.
- Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the Progress of reason. *Social Science Information*, 14, 19–47.
- Bourdieu, P. (1977). *Outline of the theory of practice*. Cambridge: Cambridge University Press.
- Carpenter, D. (2010). *Reputation and power—Organizational image and pharmaceutical regulation at the FDA*. Princeton: Princeton University Press.
- Cohen, J., & Dupas, P. (2010). “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment.” *The Quarterly Journal of Economics* 125(1), 1–45.
- Cooley, A., & Ron, J. (2002). The NGO scramble: Organizational insecurity and the political economy of transnational action. *International Security*, 27(1), 4–39.
- Cuca, R., & Pierce, C. (1977). *Experiments in family planning – Lessons from developing world*. Baltimore: The Johns Hopkins University Press.
- Daston, L., & Galison, P. (1992). The image of objectivity. *Representations*, 40, 81–128.
- Deaton, A. (2006) Evidence-Based Aid Must not Become the Latest in a Long String of Development Fads, Pp. 60–61 in *Making Aid Work*, edited by Abhijit Banerjee. Cambridge: MIT Press.
- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2), 424–455.
- Deaton, A., & Cartwright, N. (2016). Understanding and misunderstanding randomized controlled trials. *NBER Working Paper Series*, 22595.
- Demortain, D. (2011). *Scientists and the regulation of risk: Standardising control*. Cheltenham: Edward Elgar.
- Dennis, M., & Boruch, R. (1989). Randomized experiments for planning and testing projects in developing countries—Threshold conditions. *Evaluation Review*, 13(3), 292–309.
- Donovan, K. (2018). The rise of Randomistas: On the experimental turn in international aid. *Economy and Society*, 47(1), 27–58.



- Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2), 1–31.
- Duflo, E. (2003) Poor, but rational?. *MIT Working Paper 747*.
- Duflo, E. (2006) Field experiments in development economics. Lecture delivered at the 2006 world congress of the econometric society. Retrieved June 1, 2017, <https://economics.mit.edu/files/800>. Accessed 17 Oct 2018.
- Duflo, E. (2010) Social experiments to fight poverty. Lecture delivered at TED conference in February 2010. Retrieved June 1, 2017, [https://www.ted.com/talks/esther\\_duflo\\_social\\_experiments\\_to\\_fight\\_poverty](https://www.ted.com/talks/esther_duflo_social_experiments_to_fight_poverty). Accessed 17 Oct 2018.
- Duflo, E. (2011) The Power of Data in Decision Making. Lecture Delivered at the Center for Effective Philanthropy in May 10–11, 2011. Retrieved June 1, 2017, <http://cep.org/programming/national-conferences/2011-conference/>. Accessed 17 Oct 2018.
- Duflo, E. (2016) Randomized controlled trials, development economics and policy making in developing countries. Lecture delivered at the World Bank conference “the state of economics, the state of the world in June 2016. Retrieved June 1, 2017, <http://pubdocs.worldbank.org/en/394531465569503682/Esther-Duflo-PRESENTATION.pdf>. Accessed 17 Oct 2018.
- Duflo, E. (2017). The economist as plumber. *American Economic Review: Papers and Proceedings*, 107(5), 1–26.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of Development Economics*, 4, 3895–3962.
- Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *American Economic Review*, 101(6), 2350–2390.
- Easterly, W. (2007). *The White's men burden: Why the West's efforts to aid the rest have done so much ill and so little good*. New York: Penguin USA.
- Edwards, M. (2017). The Emperor's new clothes. In M. Moody & B. Breeze (Eds.), *The philanthropy reader*. New York: Routledge.
- Eyal, G. (2000). Anti-politics and the Spirit of capitalism: Dissidents, monetarists and the Czech transition to capitalism. *Theory and Society*, 29(1), 49–92.
- Eyal, G. (2013). For a sociology of expertise: The social origins of the autism epidemic. *American Journal of Sociology*, 118(4), 863–907.
- Freeman, H., Rossi, P., & Wright, S. (1980). *Evaluating social projects in developing countries*. Paris: OECD Development Centre.
- Frumkin, P. (2003). Inside venture philanthropy. *Society*, 40(7), 7–15.
- Gates, B. (2011) Small adjustments to aid programs can yield big results. The Gates notes blog. Retrieved May 2, 2017, <https://www.gatesnotes.com/Books/Poor-Economics>. Accessed 17 Oct 2018.
- Gates, B. (2014) A cautionary tale from Africa. The Gates notes blog. Retrieved May 2, 2017, <https://www.gatesnotes.com/Books/The-Idealist-A-Cautionary-Tale-From-Africa>. Accessed 17 Oct 2018.
- Gates Foundation. (2016) Grants database” and “annual letter 2013. Retrieved October 12, 2016, <http://www.gatesfoundation.org>.
- GiveWell. (2017) Top charities” and “DMI. GiveWell Org. Website. Retrieved October 15, 2017, <https://www.givewell.org/charities/DMI>. Accessed 17 Oct 2018.
- Glennerster, R. (2015) So you want to do an RCT with a government: Things you should know. Running Randomized Evaluations Blog. Retrieved June 1, 2017, <http://runningres.com/blog/2015/12/9/so-you-want-to-do-an-rct-with-a-government-things-you-should-know>. Accessed 17 Oct 2018.
- Glennerster, R., & Takavarasha, K. (2013). *Running randomized evaluations: A practical guide*. Princeton: Princeton University Press.
- Guala, F. (2007) How to do things with experimental economics in Do economists make markets?, edited by Donald MacKenzie, Fabian Muniesa, and Lucia Siu. Princeton University Press.
- Gueron, J. (2017). The politics and practice of social experiments: Seeds of a revolution. In A. Banerjee & E. Duflo (Eds.), *Handbook of field experiments*. Oxford: Elsevier.
- Haydu, J. (1998). Making use of the past: Time periods as cases to compare and as sequences of problem solving. *American Journal of Sociology*, 104(2), 339–371.
- Heckman, J., Hohmann, N., Smith, J., & Khoo, M. (2000). Substitution and dropout Bias in social experiments: A study of an influential social experiment. *The Quarterly Journal of Economics*, 115(2), 651–694.
- Heukelom, F. (2012). Sense of Mission: The Alfred P. Sloan and Russell sage foundations' behavioral economics program, 1984–1992. *Science in Context*, 25(2), 263–286.

- Heydemann, S., & Kinsey, R. (2010). The state and international philanthropy: The contribution of American foundations 1919–1991. In H. Anheier & D. Hammack (Eds.), *American foundations: Roles and contributions*. Washington: Brookings Institute.
- Hornick, R., Ingle, H., Mayo, J., Mcanany, E., & Schramm, W. (1973) *Final report: Television and educational reform in El Salvador*.
- Humpreys, M. (2015) What has been learned from the deworming replications: A nonpartisan view. Unpublished manuscript. Retrieved February 15, 2016, <http://www.macartan.nyc/comments/worms2/>. Accessed 17 Oct 2018.
- IADB. (2017) Production, use, and influence of IADB's impact evaluations. *Approach Paper Series*, Inter-American Development Bank publication.
- JPAL. (2016) The Abdul Latif Jameel poverty action lab website. Retrieved January 13, 2016, <https://www.povertyactionlab.org/evaluations>. Accessed 17 Oct 2018.
- Karlan, D., & Appel, J. (2011). *More than good intentions: How a new economics is helping to solve global poverty*. New York: Dutton Press.
- Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2014). Getting to the top of mind: How reminders increase saving. *Management Science*, 62(12), 3393–3411.
- Kim, J. (2017) Rethinking development finance. Lecture delivered at the London School of Economics (LSE) on April 11, 2017. Retrieved October 15, 2017, <http://www.lse.ac.uk/website-archive/newsAndMedia/videoAndAudio/channels/publicLecturesAndEvents/Home.aspx>. Accessed 17 Oct 2018.
- Krause, M. (2014). *The good project: Humanitarian relief NGOs and the fragmentation of reason*. Chicago: Chicago University Press.
- Krueger, A. (1995). Policy lessons from development experience since the second world war. In J. Behrman & T. N. Srinivasan (Eds.), *Handbook of development economics*, Volume III. Oxford: Elsevier Science Books.
- Krueger, A., Michalopoulos, C., & Ruttan, V. (1989). *Aid and development*. Baltimore: Johns Hopkins University Press.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge: Harvard University Press.
- Levitt, S., & List, J. (2008) Field experiments in economics: The past, the present and the future. *NBER Working Paper* 14356.
- Marks, H. (1997). *The Progress of experiment: Science and therapeutic reform in the United States, 1900–1990*. Cambridge: Cambridge University Press.
- Medevitz, T. (2012). *Think thanks in America*. Chicago: The University of Chicago Press.
- Miguel, E., & Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1), 159–217.
- Moyo, D. (2009). *Dead aid: Why aid is not working and how there is a better way for Africa*. New York: Fahar, Straus and Giroux.
- Mudge, S., & Vauchez, A. (2012). Building Europe on a weak field: Law, economics, and scholarly avatars in transnational politics. *American Journal of Sociology*, 118(2), 449–492.
- Mullainathan, S., & Thaler, R. (2000) Behavioral economics. *NBER Working Papers* 7948.
- Murray, F. (2010). The Oncomouse that roared: Hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions. *American Journal of Sociology*, 116(2), 341–388.
- OECD. (2011) Measuring aid: 50 years of DAC statistics: 1961–2011. OECD publications. Retrieved June 1, 2017, <https://www.oecd.org/dac/stats/documentupload/MeasuringAid50yearsDACStats.pdf>. Accessed 17 Oct 2018.
- Ogden, T. (2016). *Experimental conversations: Perspectives on randomized trials in development economics*. Cambridge: MIT Press.
- Panofsky, A. (2011). Generating sociability to drive science: Patient advocacy organizations and genetics research. *Social Studies of Science*, 41(1), 31–57.
- Parker, I. (2010) The Poverty Lab. Published in *The New Yorker*, on May 17<sup>th</sup>, 2010. Retrieved May 2, 2017, <http://www.newyorker.com/magazine/2010/05/17/the-poverty-lab>. Accessed 17 Oct 2018.
- Pinch, T., & Bijker, W. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, 14(3), 399–441.
- Population Council. (1986) *An experimental study of the efficiency and effectiveness of an IUD insertion and back-up component* (English summary of first six-month report, PCPES86). Lima, Peru: Population Council.

- Porter, T. (1995). *Trust in Numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Ravallion, M. (2009). Evaluation in the practice of development. *World Bank Research Observer*, 24(1), 29–53.
- Rayzberg, M. (2019). Fairness in the field: The ethics of resource allocation in randomized controlled field experiments. *Science, Technology, & Human Values*, 44(3), 371–398.
- Reckhow, S. (2013). *Follow the money—How foundations dollars change public schools politics*. New York: Oxford University Press.
- Riecken, H., & Boruch, R. (1975). *Social experimentation: A method for planning and evaluating social intervention*. New York: Academic Press.
- Rodrik, D. (2006). Goodbye Washington consensus, hello Washington confusion? A review of the World Bank's *Economic Growth in the 1990s: Learning from a Decade of Reform*. *Journal of Economic Literature*, 44(4), 973–987.
- Rodrik, D. (2008). The new development economics: We shall experiment, but how shall we learn? *HKS Faculty Research Working Paper Series*, RWP08–RW055.
- Rotemberg, M. (2009). Why academic involvement in RCTs is important? IPA blog. Retrieved June 1, 2017, <http://www.poverty-action.org/node/2156>. Accessed 17 Oct 2018.
- Sachs, J. (2005). *The end of poverty: How we can make it happen in our lifetime*. New York: Penguin USA.
- Santos, A. C. (2011). Behavioural and experimental economics: Are they really transforming economics? *Cambridge Journal of Economics*, 35, 705–728.
- Searle, B. (1985) *Evaluation in World Bank education projects: Lessons from three case studies*. World Bank Discussion Paper EDT5.
- Sommer, J. (1977). *Beyond charity: U.S. voluntary aid for a changing Third world*. Washington: Overseas Development Council.
- Stampnitzky, L. (2013). *Disciplining terror: How experts invented “terrorism”*. Cambridge: Cambridge University Press.
- Swidler, A., & Watkins, S. (2017). *A fraught embrace: The romance and reality of AIDS altruism in Africa*. Princeton: Princeton University Press.
- Teele, D. (2014). *Field experiments and their critics—Essays on the uses and abuses of experimentation in the social sciences*. New Haven: Yale University Press.
- Thaler, R., & Sunstein, C. (2009). *Nudge: Improving decisions about health, wealth and happiness*. London: Penguin Books.
- USAID. (2009). *Trends in development evaluation theory, policies and practices*. Washington: United States Agency for International Development Publications.
- Wacquant, L. (1992). "Toward a Social Praxeology: The Structure and Logic of Bourdieu's Sociology," in Wacquant, L. and Bourdieu, P. *An Invitation to Reflexive Sociology*. Chicago: Chicago University Press.
- Watkins, S., Swidler, A., & Hannan, T. (2012). Outsourcing social transformation: Development NGOs as organizations. *Annual Review of Sociology*, 38, 285–315.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Luciana de Souza Leão** is a PhD Candidate in Sociology and a Zuckerman Fellow at Columbia University. Luciana is a political and comparative sociologist with broad interests in knowledge-making processes, social inequalities, and the state. Her work has appeared in the *American Economic Journal*, *Revista Brasileira de Ciências Sociais*, *Ethnic and Racial Studies*, *Historical Social Research*, and *International Journal of Sociology*. Most recently, she contributed the chapter “A Double-Edged Sword: The Institutional Foundations of the Brazilian Developmental State” in *The Rise and Fall of the Developmental State in Latin America (1930–1990)*, edited by Miguel Centeno and Agustin Ferraro (Cambridge University Press). She is currently working on a book project, *Experimenting on the Poor: The Politics of Social Policy Evaluations in Brazil and Mexico*, which examines how politics, measurement practices, and expertise shape anti-poverty programs in Latin America.

**Gil Eyal** is Professor of Sociology at Columbia University. He is the author of, most recently, *The Crisis of Expertise*. (Polity, forthcoming in 2019). His other books include (with Berndan Hart, Emine Onculer, Neta Oren, and Natasha Rossi) *The Autism Matrix* (Polity, 2010); *The Disenchantment of the Orient* (Stanford

University Press, 2006); *The Origins of Post-Communist Elites* (Minnesota University Press, 2003); and (with Ivan Szelenyi and Eleanor Townsley) *Making Capitalism without Capitalists*. (Verso, 1998). His work examines the roles that experts and expertise play in diverse social spheres and in the production of political legitimacy. Recent papers in this vein are “Trans-science as a Vocation,” forthcoming in *The Journal of Classical Sociology* in 2019; and (with Rachel Adams, Paul S. Appelbaum, Matthew Jones, Alondra Nelson, Kevin Ochsner, John Rowe, Maya Sabatelo, Deborah Stiles, Kathryn Tabb, and Kristen Underhill) “The Physician-Patient Relationship in the Age of Precision Medicine,” *Genetics in Medicine* 14 September 2018, <https://doi.org/10.1038/s41436-018-0286-z>.

Reproduced with permission of copyright owner.  
Further reproduction prohibited without permission.