

# Stanford SOCIAL INNOVATION<sup>Review</sup>

---

## *Features*

### **The Power of Letting Go**

By Dan Honig

Stanford Social Innovation Review  
Winter 2019

Copyright © 2018 by Leland Stanford Jr. University  
All Rights Reserved







➔ New research explores when top-down control works best in international development work, and when organizations should let employees in the field navigate challenges by using their own judgment.

# The Power of Letting Go

BY DAN HONIG

Illustration by  
Jakob Hinrichs

**F**ifteen years ago, I was riding on the back of a motorcycle down the side of a mountain in rural East Timor during a monsoon. At the handlebars was Vicente Brito, my colleague at our small NGO focused on youth agriculture. That morning, the road had been dry and passable. Now, we were driving down what looked more like a river, sliding perilously close to the edge.

We paused to discuss our options. I wanted to stop for the night and wait for the morning light to find our way down the mountain. Vicente disagreed. He insisted in Tetun, East Timor's lingua franca, that he "knew" the road. I argued that it was better to live than to risk death.

"You're the boss," he said. I looked at Vicente and thought about what he meant by "knowing" the road. So much of our NGO's work came down to what each of us "knew" in a vague way. Knowing when to push a group of young people to do more, and when to offer a sympathetic ear. Knowing which local leaders to have faith in, and which to keep at arm's length. And now Vicente—who had grown up just a few kilometers away from where we were—was telling me he knew the way home. "Okay," I said. "Let's go."

We slowly made our way down the mountain, the thin headlight of our motorcycle illuminating little beyond the sheets of rain in front of us. Vicente navigated by feel, by memory, guided first and foremost by his own informed judgment. When we reached the bottom, he lifted his visor and coolly turned to me, as if to say, "I told you I could do this." And then off we went, back to our office in Dili, East Timor's capital.

The tension between my perception of the impassible road and the clear path Vicente could see has become something of an object lesson. Organizations often must juggle fallible employee judgment and top-down control by those with less contextual knowledge. The reasons not to "navigate by judgment" are many. Agents may

not have the same goals as their managers. They may lack the skills, acumen, or ability to execute tasks properly, despite the best intentions. On the other hand, even the best-designed controls can stifle employees or grantees who have to follow them. This kind of strain raises the question: When does well-intentioned management control actually improve performance, and when does it have a net negative effect?

My research shows that when the terrain is unknown or rapidly changing, better outcomes result when those actually on the ground are in control of decision-making processes. Employees who are in the field have the geographic advantage that enables them to respond quickly when flexibility and adaptation are needed—and who, through their daily experience, can incorporate what numbers miss.

## THE TRADE-OFF

To figure out what's really going on in the field, aid delivery organizations must rely on their field staff. These employees have asymmetric information—access to knowledge about what's going on "on the ground" that their bosses lack. While this information is valuable to organizations, asymmetry also gives field staff the power to misrepresent their work or shirk their responsibilities. This produces a classic principal-agent problem, as economics and political science literature usually refer to it: The boss (principal) needs to rely on employees (agents) to get things done but doesn't fully know what they're doing. Agents may not share the principal's goals, or may act in ways that do not advance the principal's goals, despite these agents' best intentions. The principal can attempt to monitor and control the agents in a variety of ways to ensure that agents act in ways desired by the principal.

Just as too little control is a risk, so is too much. Monitoring may prompt agents to execute the tasks

that are being monitored to the exclusion of harder-to-evaluate elements of their jobs. Management control may also make organizations less flexible and responsive, causing agents to act based only on what they know their principals can also see and verify. Nobel laureate economist Jean Tirole, in collaboration with Philippe Aghion, has framed the tension between management control and agent action as a trade-off between principal control and agent initiative.

Top-down control has varying degrees of success in development projects. When progress can be verified from a distance and quantitative performance targets serve as reliable summary statistics for projects—as may be the case for delivering vaccines or building roads—quantitative targets and monitoring might be quite effective. For other projects, however, the best measures are still likely to be weak proxies for a project’s purpose. Attempts at judicial or civil service reform, for example, are far less tractable to quantification. Physical locations also affect how likely top-down control is to work. Contexts vary regarding how much the right actions depend on things that can be seen but not verified, and with regard to how rapidly things change and thus whether flexible responses are necessary. We might expect, for instance, more fragile states to be places where navigation by judgment is particularly important.

In my new book, *Navigation by Judgment*, I examine when organizations might be better served by putting greater control in the hands of field staff, and increased top-down management is more conducive to organizational success. I built a database of more than 14,000 projects from nine different bilateral and multilateral aid agencies across 180 recipient countries over 40 years, in order to investigate the relationship between management practices, country context, and project success. I complement this quantitative analysis with eight qualitative case studies examining US Agency for International Development (USAID) and United Kingdom Department for International Development (DFID) projects in Liberia and South Africa.

The data suggest that agencies that navigate by judgment are much more able to cope with unpredictable environments; their performance stays remarkably stable as recipient countries become less predictable. This pattern holds up both across countries and in single recipient countries over time. As a given country becomes more fragile and unpredictable (as rated by the State Fragility Index), agencies that place greater control in the field are more able to maintain their project performance. Tasks that are less tractable to measurement drive this overall country-level effect; it’s not when, for instance, projects focus on building roads, but rather when they concentrate on improving transportation-sector management that we clearly see the advantages of greater field control.

### THE VALUE OF SOFT INFORMATION

Field staff who have the freedom to navigate by judgment can make use of “soft information”—local, contextually bound information that is difficult to include in a formal report or in an e-mail back to headquarters. Soft information is useful in many contexts; often, an organization’s success depends on it. In retired US general Stanley McChrystal’s analysis of counterinsurgency operations in *Team of Teams*, he describes Iraq as a complex, unfamiliar, and opaque environment. In previous operations, he had managed his agents via top-down control. In this environment, however, that strategy felt inappropriate. Instead, McChrystal relied on “empowered execution

**DAN HONIG** is an assistant professor of International Development at Johns Hopkins SAIS and the author of *Navigation by Judgment: Why and When Top Down Management of Foreign Aid Doesn't Work* (2018).

by field agents” and prioritized agent initiative and soft information over principal control—to marked success.

“In the old model, subordinates provided information and leaders disseminated commands,” McChrystal writes. “We reversed it: we had our leaders provide information so that subordinates, armed with context, understanding, and connectivity, could take the initiative and make decisions.”<sup>1</sup> Reducing or eliminating the control mechanisms and approval processes that slowed things down put more control in the hands of officers in the field, enabling them to respond rapidly. The organization had a greater ability to react to changing circumstances; operations could better incorporate agents’ soft information.

Evidence from international development assistance shows that soft information plays an indispensable role in development work. But such information contributes to different levels of intervention success, depending on how it is incorporated. A comparison of two of the eight case studies in *Navigation by Judgment* demonstrates this point: In the mid-2000s, both USAID and DFID had projects aimed at improving the effectiveness of South African municipal governance. However, the ways in which USAID and DFID designed and implemented their interventions were quite different.

USAID’s municipal governance project operated by delivering trainings to municipalities. On a given day, a trainer would travel to a community to hold a session on a prearranged topic—say, debt management. Success indicators suggested that all of the staff should be trained in debt management practices. Following the trainings, agents would verify that the trainings had occurred and track how many people had been trained.

By contrast, DFID’s project worked primarily by embedding in local municipalities advisors who resided there for extended periods of time, building skills and systems on an ongoing basis. DFID advisors relied on their soft information to inform their own judgment. Project documents had specific reporting requirements, but they did not rely on quantifiable outputs. Rather, DFID asked that “resident advisor ISFs [integrated service facilitators] conduct an assessment of [the] status quo and prepare a report.” Essentially, DFID advisors set their own goals and then reported their own performance.

How did the projects compare? The USAID effort proved to be a disappointment, even though it met its targets. The “numbers didn’t tell about the impact,” said the head of USAID project implementation. The training numbers weren’t fabricated, but, as one USAID actor described them, all the organization counted were “bums on seats.” Municipalities were not interested in the trainings, and little was changing. In multiple cases, national South African government officials didn’t recall the advisory component of USAID’s project, and in one case a long-serving municipal manager whose municipality had received both USAID training and a USAID advisor had no memory of USAID’s existence.<sup>2</sup> As one staff person put it, the Local Government Support Program was “a real disappointment.”

DFID, however, had some success. Its reporting, according to one implementor, was “more content-rich, not a numbers game.” As full-

time residents, DFID's advisors were often able to find a way to positively influence municipal systems. By using soft information, they could make judgments about what reforms were appropriate and how to achieve them in ways they never could have formalized for a distant headquarters. Both beneficiaries and project staff reported that DFID advisors achieved some shifts in municipal practices. It would be an overstatement to say DFID accomplished all of what it set out to do in terms of direct municipal impact. However, the DFID project was successful enough for the South African national government to use it as a model when the government launched its own municipal support program, Project Consolidate.

These projects illustrate the Tirole-and-Aghion trade-off between agent initiative and principal control. USAID and DFID implemented programs with similar goals, but DFID's project exhibited far greater navigation by judgment than USAID's, which settled on an initial model that delivered measurable trainings. USAID's program was more rule-bound, and its tight principal control precluded soft information from being incorporated into organizational decisions.

DFID, by contrast, navigated substantially by judgment. The "price" of DFID's greater degree of agent initiative was a lesser degree of principal control. DFID intentionally designed the intervention so that field agents' judgments would determine the project's direction. Field agents were the primary drivers of what the project did and when, as well as when course corrections were necessary.

## OUR COUNTING OBSESSION

While USAID's project in South Africa may have seemed extreme in its reliance on quantifiable measures, the notion of setting performance targets or objective performance criteria will sound familiar to many in the nonprofit sector and the aid industry. Use of these measures as a tool of organization control is intuitively logical: Performance targets enable principals to guide interventions from afar, preventing agents who may not share the organization's best interests from distorting projects or simply failing to work toward its goals. Organizations can also tie compensation, success, and promotion to the accomplishment of targets the principal can observe.

A recent Organisation for Economic Co-operation and Development (OECD) review of USAID found that the agency uses about 200 standard indicators, and many more custom indicators, to monitor and evaluate projects. Performance targets can orient field staff and give principals a way of holding them accountable if they do not reach their targets. All of this works well when we can set clear targets—when we know what we want our agency to accomplish, how long it's likely to take, and whether we've gotten there.

But what about when we can't set targets? Appropriate targets that drive agents toward the ultimate impact of a program are often difficult to find. For example, the OECD's Development Assistance Committee published, in 2013, a member survey of 28 international development organizations' (IDOs') experience in "Managing and Measuring for Results."<sup>3</sup> All 28 IDOs reported that they sometimes, often, or always had difficulty in selecting appropriate indicators against which to measure. Three of the 28 IDOs said that achieving this goal was a problem *100 percent of the time*. They never felt as if it was easy or straightforward to choose targets.<sup>4</sup>

Part of the problem is what's known as Goodhart's Law: Measures must be reasonable; if messy, proxies for success may lose their use-

fulness when employed to control or guide field staff actions. Counting the number of people trained may well be a good way to tell if a completed project has reached its intended audience. But when the same measure is used as a control mechanism, it inevitably will focus field staff on meeting the target and will in turn undermine the measure's accuracy as a proxy for the broader success of the intervention.

This is not to say that quantitative measures can't be useful. When what we can measure is really what we want, we should absolutely focus our organizations on achieving it. One of the true classics of private-sector management scholarship is titled "On the Folly of Rewarding A, While Hoping for B."<sup>5</sup> The "folly" is not orienting agents toward something measurable (A), but rather pointing agents toward a target (A) when some other, broader thing (B) is what the organization desires. If a project is focused on relatively verifiable tasks, such as building a road or delivering a vaccine, targets can drive field workers (and their organizations) toward success. In these situations, I am a strong supporter of the shift toward payment for performance, particularly when the standards by which performance is evaluated emphasize outcomes over process.

But plenty of the work IDOs and NGOs do has no measurable "A" that is a reliable summary statistic for interventions. IDOs' efforts at policy and administrative tasks, for example, rarely have reliable standards by which to judge success. Neither do NGO efforts to strengthen civil society, raise awareness, or improve organizational capacity.

We should know better—and, in fact, we do. A 2013 review of NGO reporting in the humanitarian sector found that only 3 percent of indicators NGOs use focus on impact, in contrast with 38 percent that focus on outputs.<sup>6</sup> The same review quotes a 2012 ALNAP report, "State of the Humanitarian System," to assert that "outputs, while easier to measure, can be misleading as indicators." I suspect that many readers have often been in rooms where someone has said, "Well, it's not a perfect measure, but it's the best we've got." Why do we need to measure, even when we know it may well distort what gets done in projects and grants? Why do we keep on engaging in Kerr's "folly"?

One reason is that we often use quantitative measures not merely to drive performance but also to report on performance. This reporting function drives us to use measures even when we know they are not accurate. To "feed the beast," we need to keep producing the data that legislators, voters, or donors seem to value, regardless of whether we believe these numbers are meaningful. The more politically insecure an IDO—the more an organization feels the need to manage up to its authorizers and funders—the less likely it is to navigate by judgment.

This isn't the only reason for our obsession with counting, though. Numbers give us a sense of security; a sense that we have "objective" data on which to base our assessments. But the numbers that give us such a sense of security are often a facade. Numbers may reflect objectivity, but they are not necessarily any more indicative of broader truth than any "subjective" assessment. When we reduce our understanding of our own efforts to what we can count, we may well improve our organizations' perceived accountability in the eyes of funders. But at the expense of actual results, this seems like a Pyrrhic victory.

## RETHINKING ACCOUNTABILITY

Going beyond a world where our accountability technology is based on what we can quantify requires us to rethink what, precisely, it means to be accountable. Core to our modern use of "accountability"



is ensuring that money achieves as much impact as possible. So, too, is demonstrating impact to stakeholders, authorizers, and funders. But my research shows a real tension between the demonstration of impact and the impact itself. What should we do when the act of measurement itself distorts for the worse the thing that is being measured?

One less-than-satisfying option, usually taken as the default, is to accept that any demonstration of impact needs to be quantifiable, and then to try to do the best job possible while also focusing accountability efforts on measuring and reporting what is countable. But if this hinders the work whose impact it aims to demonstrate, perhaps we need to consider different forms of accountability.

If we empower agents, we must hold them accountable. But *accountability* and *countability* are not the same thing. Merriam-Webster's dictionary defines "accountable" first as "subject to giving an account; answerable," and second as "capable of being explained; explainable." One way forward for an organization attempting to implement a project that is difficult to manage using measurement of either outputs or outcomes is simple, if somewhat radical: stop using measures for the purpose of evaluating interventions or managing agents. Lant Pritchett, in reviewing *Navigating by Judgment*, frames this as the distinction between "accounting" and "account-based" accountability.<sup>7</sup> We can hold people accountable, and may do a better job of doing so, the less we focus on counting, or accounting for, the numbers.

A few months ago, I had the good fortune to chat with Ruth Levine, program director for global development and population at the William and Flora Hewlett Foundation, about how and when field judgment is compatible with accountability. We ended up talking about her interactions with her program officers who make grants on Hewlett's behalf. "The question when I talk to a new program officer about what she wants to support isn't whether I would make the same decision," Levine said. "It's about the quality of her reasoning—whether she's thinking through risks and possibilities."

Levine's accountability system for her staff is primarily account-based, rather than accounting-based. This does not mean Hewlett uses no metrics in its work; Hewlett has arguably led the field in its focus on results, including the measurable results of projects (though it does so while often giving grantees substantial long-term support, which encourages trusting relationships and gives grantees the flexibility and leeway to experiment, fail, learn, and improve). Quantified performance data does play a role in evaluating program officers for Levine as well; biennial surveys of grantees provide information, including numeric ratings of program officers, which Levine considers. But these data are not the primary tool that holds staff accountable; "the numbers" are inputs into, rather than answers to, an evaluation process. Quantitative data inform judgment, rather than substitutes for it.

An accountability system like Levine's requires trust—a manager's trust in his or her own judgment, and in his or her ability to trust the judgment of others. Agents, too, must trust their supervisors and organizations. Building trust takes time and requires effort from both agents and their supervisors. Much of Levine's orientation of new staff focuses on shared expectations regarding behavior (e.g., "share information; ask for permission, not forgiveness"); the full team also discusses these expectations at every annual retreat. This process helps establish and maintain a mutual understanding of what Levine expects of her staff, and what they can expect of her.

An accountability system that facilitates agent judgment also requires the organization to have agents (staff or grantees) who the organization believes are capable of good judgment—and thus implies a greater organizational focus on who these agents are and what motivates them, rather than the carrots and sticks to which these agents might respond. This concept echoes the ideas of Harvard University political scientist Jane Mansbridge, who has argued, in situations where the best monitoring fails us, that we need to move to more trust-based "selection" accountability, not simply maintain our traditional understanding of accountability, based around "sanctions."<sup>8</sup> Mansbridge argues that an accountability system oriented to the carrots and sticks of sanctions—rewards for good performance, or penalties for poor performance—"not only stems from distrust but also creates distrust." In situations where monitoring is incomplete, a sanctions-based system may undermine trust between management and agents. Organizations may do better by focusing on selecting and training agents, instead of by implementing tight top-down monitoring and performance-based sanctions.

An organization that rethinks quantitative performance data as its primary tool of accountability does not need to eliminate measurement; it just needs to use measures for different purposes. Measures can help an organization learn and improve, or can serve as an input into decision making.<sup>9</sup> But when measures are tied to performance expectations, or taken as *the* answer to whether performance is acceptable, they become counterproductive. If USAID agents had not felt pressured to meet output targets in its South African municipal governance project, the number of people trained might have been a useful measure that catalyzed management and understanding of the project. However, the pressure accompanying these measures distorted their meaning and usefulness and played an important role in preventing the project from achieving its intended broader impact.

Another tool that can hold employees accountable but still give them the room to navigate by judgment is an after-action peer review. For example, doctors and medical personnel often engage in institutional reviews, or peer reviews, to diagnose any issues that occurred during a surgery or medical procedure. Where soft information is needed for day-to-day decision making, such as determining a course of medical action, peer reviews can be a way to improve processes. Judgments from a jury of peers can hold agents accountable and still provide mechanisms to change behavior and future action.

Too much control, not just too little control, can cause poor performance. Essential to judgment-based accountability is ensuring that reactions to any mistaken judgment or poor performance are an opportunity for professional growth through the nurturing of employees' skills. Good judgment needs to be kindled and coaxed; it cannot be dictated, and it will be lost quickly if agents anticipate that they will lose their autonomy and ability to make judgments at the first error. Organizations and their agents need the space to fail and to learn from those failures. This does not imply an absolute tolerance of mistakes. While a young doctor is not barred from the profession for a single error, neither is a consistently errant physician in training set loose on the general public. The key is remembering that one of the best ways to educate judgment is to use it and to learn from error.

Redesigning accountability to encompass the expertise of those in the field, and the wisdom of those who have done similar work, can lead to greater, more sustained organizational success. Such

an “account-based” accountability system will generate data that are harder to quickly summarize in clean, well-formatted charts for inclusion in annual reports than will one centering on what can be quantified. Yet an account-based system may well generate greater organizational learning and performance improvement over time.

### NAVIGATION BY JUDGMENT

The tension between field autonomy and the need for accountability and fidelity to an organization’s plans is common in the sector. International, national, and local NGOs, big bilateral and multilateral donors, and private-sector operators in the developing world must all try to manage unpredictable environments. Much of what these organizations do is difficult to guide using top-down control and quantifiable performance. Navigation by judgment offers the advantage of incorporating agent knowledge to improve on-the-ground performance, but it comes at the cost of principal control. For all the gains that a focus on measurable results has brought the sector, a fixation on the measurable aspects of a project can undermine the desired results.

To be sure, resources are now available to help principals manage agents. A world of GPS devices, satellite images, and smartphones has made observation of field staff easier than ever before. A recent episode of the popular National Public Radio program *Planet Money* suggested that “the future of work looks like a UPS truck.”<sup>10</sup> Apparently, the search for efficiency has driven UPS, a parcel delivery service, to a high degree of technology-enabled process control. Drivers are instructed on how to load their truck, what order to deliver packages in, and where to stop their truck on a street to make multiple deliveries. To save valuable seconds when signing forms, left-handed drivers are required to keep their pens in their right front pocket, right-handed drivers in their left front pocket. “Technology means that no matter what kind of job you have, whether you’re alone on a truck on an empty road or sitting in a cubicle in front of your computer, your company can now monitor everything you do,” NPR reported.

But a management system that works well for UPS may not work for a banker, plumber, or school principal. For struggling US public school administrators, a focus on what can be quantified may ensure that students are in classrooms but not that they are learning. On a factory floor, management technology-enabled observation and quantification may ensure that production targets are met, but at the expense of quality.<sup>11</sup> Generally, soft information and informed judgment remain essential to good outcomes.

In a variety of settings and fields, agents historically have been afforded discretion not by design, but by default; principals lacked the ability to monitor tasks well enough to make eliminating discretion a viable option. Now, monitoring technology has made possible what was previously impossible. This means organizations must make conscious, intentional decisions about when the benefits of monitoring and measurements exceed their costs and when they don’t.

Sometimes organizations will be more effective with fewer controls, less measurement, and rethinking what and who will drive programs to the results they seek. They can seek to measure smarter, rather than simply crunching an ever-increasing volume of numbers. Measuring smarter will sometimes mean measuring different things—for example, shifting from outputs to outcomes where outcomes are verifiable. Sometimes choosing not to measure, even when it is possible to do so, is measuring smarter. Smart

measurement needs to include a consideration of possible distortions in agent behavior.

An organization seeking to increase navigation by judgment need not adopt it wholesale. It can experiment by giving some offices, sectors, or projects greater control on a pilot basis and committing to evaluating the performance after it can compare the long-term impact of projects with the results from a similar country, sector, or project. But such experimentation must go beyond merely tweaking the formal rules. “Ultimately, institutional reform and real change requires more than new architecture,” a recent Overseas Development Institute paper on reforming the World Bank concludes. “It requires a change in the plumbing too—the internal systems, processes, and behaviors within agencies.”<sup>12</sup> Piloting new navigation strategies must involve not just monitoring and evaluation practices, but also management practice and HR processes for hiring, promotion, and staff rewards, to succeed. Top-down instruction can no more mandate flexibility and initiative than can “better performance.” Only careful design and patient support, not fiat, can induce a shift to greater reliance on staff judgment.

Moving toward greater navigation by judgment has its challenges. To be sure, changing organizational management strategy involves risk for those leaders at NGOs, foundations, and IDOs. But we need to weigh these risks against the benefits of greater efficiency and performance. Cost is not an issue: Navigation by judgment enables organizations to attain better results without substantial infusion of capital or high-priced technology. To forsake improvement because of our comfort with a system built on what can be quantified is to condemn many foreign aid efforts to the mere facade of success built on meaningless numbers. ■

### NOTES

- 1 Stanley McChrystal, Tantum Collins, David Silverman, and Chris Fussell, *Team of Teams: New Rules of Engagement for a Complex World*, New York: Penguin, 2015.
- 2 For this and all case study data, as well as lists of interviewees, please see Dan Honig, *Navigation by Judgment: Why and When Top-Down Management of Foreign Aid Won't Work*, New York: Oxford University Press, 2018. Some material here overlaps with the book; used with permission.
- 3 OECD Development Assistance Committee, “Managing and Measuring for Results: Survey Highlights,” Paris: OECD, 2013. The title itself suggests the aid sector’s muddling of measurement as a management tool on the one hand and measurement as a means of rigorously evaluating impact on the other.
- 4 Fifteen of the 28 reported it was “often” a problem, 10 “sometimes.”
- 5 Steven Kerr, “On the Folly of Rewarding A, While Hoping for B,” *Academy of Management Journal*, vol. 9, no. 1, 1975.
- 6 Saul Guerrero, Sophie Woodhead, and Marieke Hounjet, *On the Right Track? A Brief Review of Monitoring and Evaluation in the Humanitarian Sector*, London: Action Against Hunger and the Consortium of British Humanitarian Agencies, 2013.
- 7 Lant Pritchett, “Account based accountability and Aid Effectiveness,” Harvard Building State Capability Program, <https://buildingstatecapability.com/2018/06/28/account-based-accountability-and-aid-effectiveness>.
- 8 Jane Mansbridge, “A Contingency Theory of Accountability,” in Mark Bovens, Robert E. Goodin, and Thomas Schillemans, eds., *The Oxford Handbook of Public Accountability*, New York: Oxford University Press, 2014.
- 9 For a fuller discussion of these issues, see Robert D. Behn, “Why Measure Performance? Different Purposes Require Different Measures,” *Public Administration Review*, vol. 63, no. 5, 2003.
- 10 *Planet Money*, “The Future of Work Looks Like a UPS Truck,” Episode 563, National Public Radio, 2015.
- 11 On this last, factory-floor example, see Ethan Bernstein, “The Transparency Paradox: A Role for Privacy in Organizational Learning and Operational Control,” *Administrative Science Quarterly*, vol. 57, no. 2, 2012.
- 12 Katherine Bain, David Booth, and Leni Wild, *Doing Development Differently at the World Bank*, London: Overseas Development Institute, 2016.