

Open Access

Rachel M. Gisselquist and Miguel Niño-Zarazúa*

What Can Experiments Tell Us About How to Improve Government Performance?

Abstract: In recent years, experimental methods have been both highly celebrated, and roundly criticized, as a means of addressing core questions in the social sciences. They have received particular attention in the analysis of development interventions. This paper focuses on two key questions: (1) what have been the main contributions of RCTs to the study of government performance? and (2) what could be the contributions, and relatedly the limits? It draws inter alia on a new systematic review of experimental and quasi-experimental studies on governance to consider both the contributions and limits of RCTs in the extant literature. A final section introduces the studies included in this symposium in light of this discussion. Collectively, the studies push beyond polarized debates over experimental methods towards a new middle ground, considering both how experimental work can better address identified weaknesses and how experimental and non-experimental techniques can be combined most fruitfully.

Keywords: development; governance; randomized controlled trials.

JEL classification: C93; D72; D73; H41.

DOI 10.1515/jgd-2014-0011

1 Introduction

Experimental studies using randomized controlled trials (RCTs) have long been a staple of medical research. In recent years, these methods have become increasingly popular in the social sciences. In development economics in particular,

***Corresponding author: Miguel Niño-Zarazúa**, United Nations University – World Institute for Development Economics Research (UNU-WIDER), Katajanokanlaituri 6 B, Helsinki 00160, Finland, e-mail: miguel@wider.unu.edu

Rachel M. Gisselquist: United Nations University – World Institute for Development Economics Research (UNU-WIDER), Helsinki, Finland

their use has attracted considerable debate with some scholars promoting them as the best means of identifying “what works” in terms of development interventions (Banerjee 2007; Glennerster and Kremer 2011), while others voice strong concerns and challenge their growing hegemony in the field (see e.g., Deaton 2009; Ravallion 2009).

This article, which also serves as the introduction to this symposium, focuses on the use of RCTs in identifying what works for one of the major topics in contemporary development policy: government performance. It asks two questions: (1) what have been the main contributions of RCTs to the study of government performance? and (2) what could be the contributions, and relatedly, the limits? Despite large separate literatures on government performance and on experimental methods, very little work has directly considered both together in this way. This article draws on review of both literatures, including a new systematic review of experimental and quasi-experimental studies on governance and government performance conducted by Gisselquist et al. (2013).

Broadly, we argue that RCTs have been and can be useful in studying the effects of some policy interventions to improve government performance, but that their use in testing hypotheses about the causes of change in government performance is limited in significant ways by the nature of the factors that we expect to matter most in this area. Theories of government and the state suggest that what might be most important in explaining variation in government performance are broad macro-structural shifts, national-level variation in institutions, and other not-easily-manipulable factors such as leadership. By contrast, RCTs have been used effectively and are best for studying targeted interventions to improve government performance (particularly in areas of public goods provision, voting behavior, and specific measures to address corruption and improve accountability), which are expected to have rapid results. In other words, the types of factors highlighted in some major theories of government performance are not amenable to study with RCTs, suggesting the continuing relevance of quasi-experimental and other techniques to hypothesis testing and policy design in this area, as well as the value of mixing experimental and non-experimental approaches.

As summarized in the final section of this article, the rest of this symposium builds from this discussion to consider both how experimental work might better address identified weaknesses and how experimental and non-experimental techniques might be combined most fruitfully, with particular attention to the study of government performance. Collectively, the studies in this symposium speak to debates over experimental methods in development studies more generally, pushing beyond what have become well-rehearsed and polarized positions towards a new middle ground. Contributors include scholars who regularly use experimental methods in their own work and those who have been more critical or reliant on other methods.

Discussion of experimental research, particularly in economics and political science, sometimes treats laboratory-type experiments, natural experiments, and RCTs or field experiments together, irrespective of their design features. Much of the discussion in this article is applicable to various experimental methods, but the focus is on RCTs, which imply a slightly different approach than the others to the testing of causal hypotheses: in the simplest experimental designs, causal effects are assessed by comparing measures “with” and “without” an intervention. This is most straightforward in a laboratory setting where other key variables can be held constant and measures can be taken before and after an intervention. In this setting, causal inference is relatively clear: the intervention causes the difference.

Many of the phenomena that we care about, however, are not amendable to this method. Outside of the laboratory setting, field experiments and RCTs study such phenomena using similar principles; because it is not always possible to hold constant all factors, in prospective experimental designs, with baseline and endline data, the identification of the counterfactual is achieved via random assignment to treatment, where measures from randomly selected “control” and “treatment” groups are taken before and after interventions, and the effect of the intervention then is the difference between “before” and “after” measures in the treatment as compared to the control groups. This basic and elegant logic underlies hypothesis testing and impact evaluation using RCTs, by ensuring that – in principle – any difference between the treatment and control is not systematic at the outset of the randomized experiment.

Quasi-experimental approaches by contrast rely on observational data that lack random treatment assignment. Relying on similar logic to experiments, they draw on quantitative techniques such as instrumental variables, regression discontinuity, or propensity score matching to estimate the impact of interventions. Our review of the literature in this study does not focus on other non-experimental methods, such as qualitative case studies, although the middle ground we explore in this symposium could also draw on these methods.

This article has three main sections following this introduction. The first briefly discusses theories of governance, highlighting several major ideas from the literature about the factors that cause variation in government performance. The second focuses on how RCTs have been used in the study of related topics, highlighting major findings from RCTs with respect to the provision of health, education, and other public goods; improvements in the performance of civil servants; and representation, participation, and deliberative democracy. The third part of the article brings these two sections together, exploring the limits and potential contributions of RCTs in the study of government performance and introducing the contributions of the studies in this symposium. A final section concludes.

2 Explaining Government Performance

Despite a wealth of literature on governance and government performance, even definitions remain contested (see Keefer 2009; Gisselquist 2012; Bratton 2013). Without delving too much into these debates, we adopt a basic definition of *governance* building on theories of government and the state. This work points to two major roles for public institutions in (1) providing public goods such as education, health care, water and sanitation facilities, and social protection, and (2) aggregating interests with respect inter alia to how and which public goods are provided. The later role can be achieved both through electoral and non-electoral forms of participation and is closely linked to discussion of accountability. As Putnam (1993: pp. 8–9) notes in *Making Democracy Work*, “public institutions are devices for achieving purposes, not just for achieving agreement. We want government to do things, not just decide things – to educate children, pay pensioners, stop crime, create jobs, hold down prices, encourage family values, and so on.” Individuals and groups within a polity have varying preferences about the type and manner of public goods provision and other collective issues, and a second key role of government is in “aggregating” and representing such interests to make collective decisions. In short, as Levi (2006: p. 5) summarizes, “Good governments are those that are (1) representative and accountable to the population they are meant to serve, and (2) effective – that is, capable of protecting the population from violence, ensuring security of property rights, and supplying other public goods that the populace needs and desires”. By extension, the quality of governance varies in the degree to which governments fulfil these two related roles.

A standard definition of *government performance* is the “capacity of governing authorities to provide public goods and services” (Bratton 2013: p. 2). As Bratton (2013: p. 2) elaborates:

Public goods and services are benefits intended for shared consumption from which nobody can be legally or effectively excluded. These benefits may be general political guarantees (such as national defense, law and order, or civil rights) or particular socio-economic commodities (such as transport infrastructure, electricity supply, or health and education services). Either way, government performance concerns the track record of state agencies at addressing collective social needs.

In studying government performance, researchers refer to best practices in the delivery of public goods and services, which are usually measured in terms of the outputs, outcomes or impacts of public programmes, including citizen assessments thereof.

In these terms, the study of government performance is understood as an integral part of the broader literature on governance. It highlights effectiveness in public goods provision, but representativeness and accountability are also key because

effectiveness may be judged in terms of the upholding of political guarantees that make possible representativeness and accountability (as well as the provision of socio-economic commodities), and because citizen assessment is relevant to the evaluation of how well collective social needs are addressed. Much more than in the literature on governance, the literature on government performance highlights performance management – “the planning, programming, implementation, measurement, monitoring, evaluation, and reporting of program results” (Bratton 2013: p. 2). A clear focus has been the range of ways in which better public sector management – including strategic management, financial management, human resource management, performance measurement, quality management, process management, and so on – influence government performance outcomes (see Bovaird and Löffler 2009). Work on “governance” by multilateral development banks, in particular the World Bank, can be understood in these terms as broadly indistinguishable from government performance. The World Bank (2012: p. 9) explicitly treats governance as “the processes by which authority is exercised in the management of a country’s economic and social resources” and “the capacity of governments to design, formulate, and implement policy and deliver goods and services.”

Theories of government and the state, however, push us beyond this focus on public sector management to suggest a number of additional explanations for variation in government performance, both across polities and over time, tied to a range of structural, institutional, and cultural factors, as well as individual agency. In general, this work deals with the two components of governance separately, offering explanations either for better representation and accountability (often framed in terms of the emergence of liberal democracy versus other forms of government), or for more effective public goods provision. Much work also addresses disaggregated government performance outcomes, such as the provision of effective policing, secure property rights, universal health care, or high quality state-funded education.

Far from having a single model of change in government performance, major theoretical traditions offer diverse and often contradictory explanations for key outcomes. For instance, what constitutes good government performance in terms of providing the institutional environment most conducive to economic growth? Friedman (1962) suggests that a “good” government serves as a “rule maker” and “umpire” to create and enforce minimal rules, such as property rights and a monetary framework. Disciples of Keynes (1964), by contrast, see a more extensive role for governments in fiscal and monetary policy. Similarly, major debates concern whether more or less regulation is most conducive to private sector development, and the form that regulation should take (see Kirkpatrick 2014).

One of the major structural factors highlighted in classical explanations of variation in the quality of governance is “modernization” or the level of development. Max Weber, for instance, suggested that modernization leads to fundamental

changes in the nature of authority, from traditional and charismatic towards the sort of rational-legal approach embodied in contemporary public sector management principles (Weber 2009). Modernization theory of the 1950s and 1960s posited that economic growth led to fundamental structural changes in the economy and society, which would in turn lead to greater popular participation in government and the foundations of democracy (Lipset 1981). Other major structural arguments highlight factors ranging from the class structure (e.g., Moore 1966; Luebbert 1991) and ethnic structure (e.g., Horowitz 1985) to the role of the state and state-society relations (e.g., Skocpol 1979; Migdal 1988) and geography (e.g., Herbst 2000).

Institutional explanations for governance outcomes are among the most diverse, focusing on a range of forms and mechanisms. Indeed, social scientists often define institutions so broadly – as formal and informal rules, norms, and organizations – that some of the structural explanations noted above and the cultural explanations noted below are treated in this camp (see Steinmo 2008). Since the 1990s, new institutionalist economics inspired by North (1990) and others has been particularly important in the thinking on governance underlying work by the World Bank and other multilateral development banks, which has focused on how the “rules of the game” shape economic development (Grindle 2010). The World Bank, for instance, has focused on its role in helping countries to “put in place institutions and systems that can become the foundations of sustainable growth” (World Bank 2012). It points to the need both to strengthen the capacity of institutions to enforce regulations, provide public services, and manage resources effectively, and to adopt the “right” institutions (e.g., regulations favorable to private sector development).

Institutional explanations often focus our attention on the long-running impact of “lock in” or path dependence that makes changing institutions difficult and costly, even when they are inefficient. North argues, for instance, that it was rather haphazard institutional choices that put England on a path toward efficient market economy, with relatively strong property rights, an impartial judicial system, and a fiscal system with expenditures tied to tax revenues, where other countries adopted different (and ultimately less effective) institutions that placed them on different paths. Likewise, from a sociological perspective, Immergut (1992) argues that the structure of political institutions in Sweden, France, and Switzerland influenced whether each country developed comprehensive national health care or more fragmented insurance programs. Political institutions and procedures, rather than the demands of social groups, set the terms of political negotiations, leading to divergent outcomes.

Institutionalists have also studied constitutional engineering and electoral system reform as means of improving representation, accountability, and governance more generally, particularly in divided societies (see Sartori 1997; Reilly 2001). Consociational theory, for instance, proposes that governance in a state

divided along ethnic, religious, or communal lines can be stable if democracy has four key institutional characteristics: a grand coalition formed by the political leaders of various factions; a mutual veto, necessitating consensus among groups for political decisions; proportionality, in that each group occupies a share of government posts proportional to its share of the population; and segmental autonomy, allowing autonomous rule for different groups (Lijphart 1977). Many favor for similar reasons the reorganization of the state along more decentralized or federal lines in order to defuse social divisions, increase public accountability and political participation, and improve public service delivery (see Seabright 1996; Bardhan 2005; Brancati 2006; Robinson 2007).

Cultural factors are also relevant to theories of variation in government performance. Tocqueville's (2003) classic exploration of the role of political culture in explaining democracy in America is one example. In the Tocquevillian tradition, Putnam (1993) explored government performance across Italian regions, taking advantage of a unique situation in which 15 new regional governments were established simultaneously in 1970 with similar constitutional structures and mandates, but some performed better than others. In explaining this variation, he pointed to the role of social capital – rooted in Italy in early medieval history. The role of social capital in government performance and the factors that influence its variation likewise have been emphasized in multiple contexts (e.g., Boix and Posner 1998; Varshney 2001; Putnam 2007).

Finally, a significant body of work points to the decisive role of individuals, namely decision-makers and political leaders, in affecting government performance. Kingdon's (1995) model of policy-making, for instance, posits three "streams of processes": problems, policies, and politics. Policy windows, which may arise predictably (such as during a vote on legislation) or suddenly (when problems arise), are periods during which the three streams are combined and issues may rise on the policy agency. Policy entrepreneurs take advantage of policy windows to push their agendas and particular policy solutions. Grindle and Thomas (1991) similarly suggest a critical role played by individuals in agenda setting, decision making, and the implementation of public reforms that affect government performance. Leaders are often seen to play an especially decisive role during periods of institutional change and uncertainty (Samuels 2003). Scholarly attention has also focused on the role of particular individuals (e.g., Mukunda 2012) and leadership styles (e.g., Jackson and Rosberg 1982).

In summary, the study of how and how well governments perform is central to the study of politics, and work in the field – just a very few examples of which are cited here – suggests a number of structural, institutional, cultural, and agent-focused explanations for variations in government performance. This brief overview provides a quick glimpse into some of the major approaches in the literature.

3 Findings From Experimental Work

In one of the earliest reviews on the use of field experiments to study governance and government performance, Humphreys and Weinstein (2009) identify four major questions on which researchers have focused: (1) what is the role of political institutions in the process of decision-making and policy implementation? (2) How do social norms and informal institutions affect individual and collective action? (3) What is the impact of information and incentives on political behaviour, notably accountability? And (4) how can violence and conflict be prevented? The authors cover a limited number of studies and acknowledge that “there has not yet been a significant accumulation of knowledge from the use of field experiments in the political economy of development [...] For this reason, we focus more on the promise of the field than on its achievements” (Humphreys and Weinstein 2009: p. 370).

In a subsequent review of experimental research on democratic governance, Moehler (2010: p. 30) addresses the related question of whether field experiments can “be productively employed to study the impact of development assistance on democracy and governance outcomes?” She highlights several key weaknesses of field experiments, but is generally sanguine about the possibilities: “The enterprise of DG field experiments,” she notes, “will be constrained more by mundane challenges to successful research design and implementation than by the inherent limitations of field experiments” (Moehler 2010: p. 42). Her review identifies 41 randomized field experiments of interest in the developing world, including eleven dealing with elections, ten with community-driven development, nine with government performance in public service delivery, three with the use of quotas, and seven with other topics. The majority of the reported studies (22) were conducted in Africa, and nine in India.

More recently, Olken and Pande (2011b) conducted a narrative, non-systematic review of the literature on governance, following a principal-agent approach. They include in the review 16 studies that adopt rigorous experimental and non-experimental methods to establish causality in the analysis of policies that aim to improve governance in developing countries, dividing the literature into two areas: (1) participation and participatory institutions to exercise greater control over politicians, and (2) the roots of corruption and the incentives and institutional features that can prevent rent-seeking behaviour and leakages.¹

Recent studies also review findings from RCTs with respect to development interventions more generally. Banerjee and Duflo (2012), for instance, draw largely

¹ For a review on the specific topic of corruption, see Olken and Pande (2011a).

on the results of their work at the MIT Poverty Action Lab to propose new solutions to global poverty, highlighting the role of “ideology, ignorance, and inertia” in explaining why aid is not always effective. Many of their solutions point to how the poor lack critical information and hold incorrect beliefs (e.g., about the benefits to education) that help to perpetuate their poverty. They highlight findings from RCTs dealing with hunger, health, education, family planning, risk management, microfinance, and entrepreneurship. Karlan and Appel (2012) build a similar argument about solutions to global poverty, also drawing heavily on findings from RCTs. They highlight “seven ideas that work:” microsavings, reminders to save, prepaid fertilizer sales, deworming, remedial education in small groups, chlorine dispensers for clean water, and commitment devices (Karlan and Appel 2012: pp. 272–275).

Building on this earlier work and in order to address potential threats of publication bias, a *systematic* review of published and unpublished papers using experimental and quasi-experimental methods to study government performance was conducted as part of the research for the broader project of which this symposium is a part.² From an initial sample of over 3000 papers on governance-related topics released between 1990 and 2012, the review identified and analyzed 139 papers judged relevant based on topic and methods employed. Building on the approach outlined above, the review includes both studies focused on the provision of public services and on aggregating interests in various ways. As summarized in Figure 1, studies are further classified based on the type of policy intervention, that is, whether it primarily aims to (1) improve supply-side capabilities of governments, and the social and political institutions that facilitate that process; (2) change individual behaviour through various devices, notably incentives, and/or (3) improve informational asymmetries. Interventions of the first type focus on the supply-side dimensions of policies, affecting how public institutions themselves provide public goods and services. Improving government performance in this context may involve changes both to what is provided (e.g., books, classrooms) and their quality (e.g., better books, better teachers). The second and third types of interventions directly influence the demand-side of government performance, i.e., how the population (usually individuals, households, and communities) interact with public institutions. Demand-side interventions

² For a detailed description of the systematic review protocol, see Gisselquist et al. (2013). One issue worth highlighting is that some papers may be classified in multiple ways. For the purposes of our discussion here, we chose a single classification based on our assessment of the most relevant category.

Government performance cluster	Type of policy intervention	Policy area
Aggregating interests (74)	Improve supply-side capabilities (29)	Institutional building (17) Non-electoral forms of participation (3) Voting behaviour (3) Ethnicity (2) Accountability and corruption (2) Democracy (1) Leadership (1)
	Change behaviour via incentives(20)	Institutional building (5) Accountability and corruption (5) Non-electoral forms of participation (5) Voting behaviour (2) Leadership (1) Migration (1) Social cohesion (1)
	Improve information asymmetries (25)	Voting behaviour (19) Institutional building (4) Non-electoral forms of participation (1) Accountability and corruption (1)
Provision of public goods (66)	Improve supply-side capabilities (18)	Health and/or education (11) Water and/or sanitation (3) Employment (2) Democratic governance (1) Social protection (1)
	Change behaviour via incentives (35)	Health and/or education (23) Employment (5) Social protection (4) Rural development (1) Housing (1) Water and/or sanitation (1)
	Improve information asymmetries (12)	Health and/or education (7) Water and/or sanitation (3) Employment (1) Non-electoral forms of participation (1)

Figure 1: Experimental and Quasi-experimental Studies to Study Government Performance. Number of reviewed studies in brackets. Several studies are classified in multiple categories in the Appendix but are counted here with only one.

Source: Authors.

are found either to provide incentives (often in cash) or to provide better information about the provision of goods and services, both with the objective of changing individual behavior. Finally, studies are grouped into commonly-referenced policy areas (e.g., social protection, health and education, democracy, accountability and corruption).

As shown in Figure 1, under the “provision of public goods” cluster, the largest number of studies (41) focuses on health care and education policies, with fewer addressing issues of employment, water and sanitation, housing, and other topics. Under “aggregating interests” particular attention is paid to institution building (26), particularly in the context of improving supply-side capabilities (17), and to electoral participation and voting behavior (24), particularly in the context of improving information asymmetries (19). The largest number of studies in our sample was conducted in the USA and India (see Table A1 in the Appendix)

It also is worth highlighting the challenge of drawing sharp distinctions in this body of work between experimental and quasi-experimental studies. In particular, our review of the literature reveals that a significant number of studies that adopted experimental designs had to resort to quasi-experimental regression techniques such as propensity score matching and instrumental variables techniques to address errors in implementation that caused endogeneity problems, spillovers, and sample contamination (see Table A1 in the Appendix).

A number of findings emerge from these studies that are relevant to explaining variations in government performance. As above, many of these relate to the ways in which governments (or donors) can improve the provision of basic public goods, particularly in the areas of health care and education and deal with the impact of projects providing specific goods or services. In their study of the Primary School Deworming Project in Kenya, for instance, Miguel and Kremer (2004) find that the program not only improved students’ health in both treatment schools and neighboring schools, but also reduced school absenteeism by a quarter (although there was no evidence of an effect on academic test scores). In demonstrating the impact of expanded insurance coverage on improved health outcomes among children, Quimbo et al. (2011) draw on the Quality Improvement Demonstration Study in the Philippines to show that zero co-payments and increased enrolment were associated after release from the hospital with reduced likelihood of wasting and of having an infection (9–12 and 4–9%, respectively). Kremer et al. (2009) evaluate the impact of a merit scholarship program in Kenya in which girls who scored well on exams had school fees paid and received a grant, finding that the program had an effect not only on improved student test scores, but also on teacher attendance.³

³ For recent systematic reviews of the literature on education policies, see Conn (2014), Krishnaratne et al. (2013), Masino and Niño-Zarazúa (2015), McEwan (in press) and Petrosino et al. (2012).

A number of studies explore the impact of public information campaigns on public goods provision. Pandey et al. (2009), for instance, evaluate the impact of a community-based information campaign across three Indian states consisting of eight or nine public meetings to disseminate information to communities about its state-mandated roles and responsibilities in school management. They find the largest impacts on teacher effort, and more modest improvements on student learning and the delivery of benefits to students (stipends, uniforms, and mid-day meal). Also in India, Pattanayak et al. (2009) explore the impact of the intensified “information, education, and communication” campaign carried out in Orissa as part of the nationwide Total Sanitation Campaign to change rural household attitudes about the use of latrines. The study found that latrine ownership rose significantly in treatment villages and remained the same in control villages. Pattanayak et al. (2009) further address the question of whether social and emotional costs (“shaming”) or financial incentives (“subsidies”) better influence behaviour. They find that although latrine ownership rose most among households below the poverty line and eligible for a government subsidy (5–36%), it also rose among wealthier households not eligible for the subsidy (7–26%), suggesting that shaming, even in the absence of subsidies, can work to change behaviour.

Conditional cash transfers as a strategy have received particular attention and been evaluated in several different contexts. A number of studies focus on Mexico’s Progresa/Oportunidades program (e.g., Stecklov et al. 2007; De La O 2013). Leroy et al. (2008) for instance, find the program to be associated with better growth in infants below 6 months of age (but to have no impact for babies 6–24 months). Other studies explore the impact of conditional cash penalty programs. One example is Dee’s (2011) study of the effects in ten counties of the state of Wisconsin’s Learnfare program, which sanctions a family’s welfare grant when teenagers in the family do not meet school attendance targets. Data suggest evidence in nine counties that Learnfare increased school enrolment by 3.5% and attendance by 4.5%.

Another set of studies focus on interventions to improve the performance of public sector employees such as teachers and nurses. Multiple studies highlight the impact of financial incentives. Duflo and Hanna (2005), for instance, find that a financial incentive program immediately reduced teacher absenteeism in rural India, which was also associated with an improvement in student test scores and achievement 1 year after the start of the program. Basinga et al. (2011) find in Rwanda that adoption of performance-based payment of health-care providers (“P4P”) was related to improvements in the use and quality of child and maternal care services, including a 23% increase in the number of institutional deliveries and increases in the number of preventive care visits by children (56% for those

23 months and younger, and 132% for those 24–59 months), and improvements in prenatal quality as measured by compliance with Rwandan prenatal care clinical practice guidelines.⁴ Other studies explore the impact of relatively minor administrative reforms: Banerjee et al. (2012), for instance, test the impact of four low-cost reforms across police stations in eleven districts in Rajasthan. Results suggest that two of these reforms – freezing staff transfers between police stations and providing in-service training in investigation skills and “soft” skills like communication and leadership – were effective in improving police effectiveness and public satisfaction, while the other two reforms – placing community observers in police stations and a weekly duty rotation – were not effective.

A growing body of experimental and quasi-experimental work also studies issues related to aggregating interests through electoral politics in new and emerging democracies. Fujiwara and Wantchekon (2013), for instance, explores whether public deliberation – in the form of town meetings – can overcome clientelism in Benin. The experimental data show a positive effect on perceived knowledge about policies and candidates and on voter turnout, as well as increased electoral support for the candidates participating in the intervention. Collier and Vicente (2008) evaluate the effect of a campaign against political violence run by an NGO in Nigeria, involving town meetings, popular theatres, and door-to-door distribution of material. They find that this intervention served to reduce the intensity of election-related violence. Hyde (2010a) shows that the presence of election observers had an effect on election quality in the 2004 Indonesian presidential elections, measured in terms of votes cast for the incumbent. Ichino and Schündeln (2012) study the effect of domestic observers on voter registration in Ghana in 2008. They find that because parties operate over large areas, observers in one registration center may displace irregularities to others, which suggests the need for some revisions to how such observers are deployed in many countries.

Finally, a number of studies explore topics at the intersection of representation and public service provision, with particular attention to the impact of community-based monitoring initiatives. Björkman and Svensson (2009), for instance, find in Uganda that holding meetings among community members and health workers to discuss health services and how to improve them, to compare citizen and health worker views of service provision, and to collectively discuss patient rights and provider responsibilities, led to improved health outcomes (reduced child mortality and increased child weight), as well as more

⁴ For a systematic review on nonclinical interventions against childhood diseases, see Seguin and Niño Zarazúa (2015).

community monitoring of health care a year after the intervention. Olken (2010) explores the relationship between direct democracy and local public goods provision in rural Indonesia, studying plebiscites introduced in some villages to replace a meeting-based process presumably dominated by elites. Plebiscites were associated with higher public satisfaction and perceived benefits from the project, greater willingness to contribute, and increased knowledge about the project. On the other hand, Olken's (2007) study of "top down" versus grassroots participation in corruption monitoring in Indonesia suggests that government-led approaches may be the more effective on this issue. Increasing government audits had a significant effect on reducing corruption in term of reducing missing expenditures and discrepancies between official project costs and independent estimates of costs, while increasing grassroots participation had little impact.

In summary, findings using experimental data highlight an array of strategies that governments can adopt to improve public service provision and representation and accountability in particular areas. Innovations that have been explored in multiple contexts include public information campaigns, conditional cash transfers, financial incentives to improve the performance of public sector employees, community-based monitoring, and public deliberation at the local level.

4 The Limits of Experimental Methods in the Study of Government Performance

The elegance of RCT findings arguably has a tendency to promote method-driven, rather than theory-driven, research: it tends to encourage work that asks questions that can be addressed with experiments, rather than work that begins with questions that are seen as important to answer and then proposes hypotheses and assesses the methods most appropriate for testing them.

One key criticism levelled at experimental work is that it does not address "big" questions and "big" theories (Hyde 2010b). Indeed, if we compare the factors explored in the experimental studies reviewed in Section 3 with those identified in major theories of government performance, there is clearly a disconnect. Key factors in major theories of governance, such as modernization, social structure, and national institutions, for instance, are largely absent as an object of study in experimental evaluations.

Proponents of RCTs, however, make a compelling argument that their avoidance of grand theory can be a strength as well. Banerjee and Duflo (2012), for one,

explicitly advocate for an incremental, micro approach. Their solutions posit that government functioning can be improved with small policy reforms that at the margin can lead to desirable improvements in policy, without major changes to social and political structures. Karlan and Appel (2012: p. 37) contend that development “needs to be on the ground,” “up in the realm of high-minded concepts... the air is thin and there are no poor people to be found.”

Still, despite such explicit rejection of grand theory, experimental approaches are not absent of theoretical underpinnings. Analysis often falls clearly within the rigor of behavioral economics, drawing implicitly on its theories of individual behaviour, (ir)rational choice, and information. Nevertheless, this micro focus exacerbates a second of the key weaknesses of experimental work: the low external validity of its findings. If findings from RCTs are to be used to identify generalizable impacts – to help predict the impact of similar interventions in other situations – experimental work must be able to say something about the broader context. Precisely because experimental researchers tend to adopt a micro approach to research enquiry, and eschew more high level theorizing about what within particular contexts might be unique or have influenced results, experimental studies tend to provide weak leverage on the question of whether similar outcomes might be expected in other contexts (see Pritchett and Sandefur 2013).

One strategy for improving external validity in experimental research is to speak more directly to broader theoretical propositions. Martel Garcia and Wantchekon (2015) in this Issue draw on structural theory to advance our understanding of external validity and generalization of causal problems. They show that external validity is a function of theoretical generalization, i.e., the ability of researchers to explain and predict outcomes across variations in treatments, outcomes, and settings. However, even when adopting such approaches, a degree of uncertainty remains with regard to the underlying mechanisms that explain, under a theoretical framework, the distribution of policy outcomes for a particular group (treatment and control) vis-à-vis the distribution for the entire population. This constraint forces us to look beyond experimental methods alone in the study of government performance.

Another strategy to improve the external validity in experimental research involves replicating the same type of intervention in different sets of conditions. For example, the Consultative Group to Assist the Poor (CGAP) and the Ford Foundation have funded the replication of Bangladesh’s BRAC’s Challenging the Frontiers of Poverty Reduction/Targeting the Ultra Poor program in India, Pakistan, Honduras, Peru, Ethiopia, Yemen, and Ghana, using experimental designs for their evaluations. The “Metaketa” projects, such as the new initiative on political information and electoral choices carried out in Benin, Brazil, Burkina Faso,

India, Mexico, and Uganda are another example.⁵ Further, with the rapid accumulation of experimental and quasi-experimental evidence, systematic reviews and meta-regression analyses have been increasingly used to produce generalizations about specific policy interventions and their effects under different socio-economic settings. This process has been supported by a number of organizations, including the Cochrane Collaboration, Campbell Collaboration, and International Initiative for Impact Evaluation (3ie).

A third limit to RCTs in the study of government performance is in the type of causal factors that they can reasonably study. This constraint follows partly from the need for large numbers of units to be studied in order to gain precise estimates, which encourages researchers to focus on low level factors, rather than on factors held by higher level units, such as national institutions (Moehler 2010). Some traction on such factors can be gained by “scaling up” findings from low level factors. For instance, some findings from studies of public administrative reforms carried out at the municipal level may inform reform at the national level. However, municipal versus national politics are so different in other ways that this sort of scaling up clearly provides only suggestive evidence.

Experimental evaluation protocols have nonetheless proved to be relevant in contexts where incumbent governments and political elites need to be persuaded to scale up social policies. Barrientos and Villa (2015) in this Issue analyze a new database on antipoverty social protection programs in Latin America and sub-Saharan Africa, documenting the influence of political competition, as well as the rise of “evidence-based” development policy, in explaining the adoption of such evaluations. They find that while in Latin America impact evaluation of antipoverty programs has been driven by agency competition, design features of programs, and political factors, in sub-Saharan Africa, the interactions between donors and domestic elites have undermined the demand for evaluations.

The limits to the causal factors that RCTs can study also follow from the simple inability of researchers to manipulate some key variables identified in the literature, such as the level of development, national institutions, culture, or the quality of national leadership. Putnam’s (1993) study of new regional governments in Italy serendipitously gave him a natural experiment to exploit, but such situations of “natural” random assignment may be both rare and problematic to analyze (Sekhon and Titiunki 2012).

In other cases, ethical considerations may impede the study of particular factors and outcomes. For instance, interventions expected to foment electoral

5 The first Metaketa project started in 2013. For more information, see Dunning et al. (2015).

violence or to influence which candidate wins an election are not undertaken for obvious ethical reasons. As Humphreys (2015) in this Issue notes, more work remains to be done in this area by researchers in the social sciences as the main principles of research ethics currently employed were developed by and for medical research. Humphreys provides a detailed discussion and critique of ethics as currently understood in work in this area and of approaches to consent that can inform the further development of ethical guidelines.

A fourth issue that limits the utility of RCTs in the social sciences is their relatively short-term window of analysis. This is particularly problematic in the study of government performance because so many of our major theories focus on “non-linear” processes that evolve over decades or generations, while RCTs rarely look at impacts beyond the “linear” trajectory between two points in time, usually a few years. Take, for example, the hypothetical case of a J-shaped curve derived from the long-term relationship between economic liberalization and political stability: in the short-term, economic liberalization may lead to a sudden rupture between economic and political actors that cause an increase in political instability. An RCT may conclude that economic liberalization is bad for political stability. However, if theory’s predictions are correct, once markets and institutions are developed further, political stability would actually improve (Gans-Morse and Nichter 2008). Although the time horizons of RCTs could be extended somewhat, they would still not be long enough to explore many of the major theories of governance.

One approach that may help to extend the time horizon is suggested by the use of what Baldwin and Bhavnani (2015) in this Issue call “ancillary experiments.” As more RCTs become available in the social sciences, opportunities arise to use existing experimental data to investigate new questions. Researchers can collect new data on populations assigned to treatment and control groups in previously executed experiments, and then rely on the initial randomization to identify new effects. Ancillary experiments also provide many of the advantages of RCTs but at lower cost, since the intervention has already been undertaken. A classic example of ancillary experiments is the study by Angrist (1990), who took advantage of the Vietnam draft lottery to examine the effects of military service on lifetime earnings.

A fifth key concern with RCTs is that they are similarly limited in terms of the unit of analysis upon which they can evaluate impacts, which is generally the individual. Some studies focus on other units of analysis, such as voting constituencies or local regions, but no studies of which we are aware conduct experiments with representative samples at the national level. This is simply due to the fact that the treatment effects arising from policy interventions are often small, and therefore large sample sizes are needed to conclude, with enough statistical power, that the differences between the treatment and control groups are unlikely to be due to chance. This connects to a final issue: the cost of RCTs.

Some scholars have questioned whether RCTs are worth the cost, which often involve million dollar budgets (Heckman and Smith 1995). Randomization by group or cluster is often used in medical science to lower the cost of RCTs via phased implementation. This approach significantly decreases the cost of running studies, particularly in contexts where the outcomes of interest are easily assessed; however, even if they could be adapted to address some key theories of government, it is not necessarily clear whether they would be more cost-effective in testing these theories than observational methods.

Observational studies, in contrast, use econometric methods each with its own assumptions, strengths, and weaknesses to estimate the impact of policy innovations and use data ranging from small surveys to nationally representative surveys and censuses. From the perspective of experimental design, observational studies have been subject to criticism on the basis of their internal validity. Dehejia (2015) provides in this Issue a survey of observational methods, including regression analysis, instrumental variables, difference-in-difference estimators, regression discontinuity, and matching, and assessed their internal and external validity relative both to each other and to RCTs. He concludes that both experimental and non-experimental methods face important challenges, but can sometimes be fruitfully combined.

Indeed, although experimental and observational methods are often presented as mutually exclusive, they are in fact complementary in two senses. First, the principles of experimental design can improve the internal validity of observational studies. Survey-based initiatives to study governance such as the Afrobarometer project for instance have begun to explore such strategies (Bratton 2013). Second, they can be combined using non-experimental adjustments to deal with implementation failures in RCTs. As our review of the literature in Section 3 suggests, experimental studies in the fields of development economics and political science – arguably more so than in medicine – can pose significant logistical and methodological challenges that result in reliance on quasi-experimental regression techniques to tackle problems such as confounding, selection bias, spillovers, and impact heterogeneity (Deaton 2009).

5 Conclusion

This study argues that RCTs have been and can be useful in studying the effects of some policy interventions and reforms, but that their use in the study of government performance is also limited in significant ways, particularly by the

nature of the factors that we expect to matter most. RCTs are best for studying targeted interventions (particularly in areas of public goods provision, voting behaviour, and specific measures to address corruption and improve accountability), which are expected to have rapid results. However, our review of theories of government and the state suggests that what might be most important in explaining variation in outcomes are broad, macro-structural shifts, national level variation in institutions and political culture, and the actions of individual policy decision-makers.

The focus in this article has been on the use of RCTs to test hypotheses about the causes of variation in government performance. One might argue that in fact a narrower focus is warranted, highlighting precisely the individual, household, or community-level factors that policy makers might leverage to improve government performance. In other words, RCTs may have particular resonance for a more policy-focused audience as the sort of interventions best studied with RCTs may be precisely those most readily manipulable by donors and domestic policy makers.

However, given the limits of RCTs explored in this study, it is clear that even within a narrower policy-focused frame, analysts must go beyond experimental methods to address the questions of “what works” and “what could work.” Small-scale interventions may be most readily manipulable by policy makers, but policy makers can also focus attention on influencing change in a number of larger-scale factors identified in the literature as potentially important to improving government performance, such as national institutions and social capital, that RCTs are ill-equipped to study. Thus, policy makers might look beyond experimental studies to inform their reform strategies in this sense. In addition, even with respect to reforms and interventions that have been explored through RCTs, the weak external validity of RCTs raises major questions about whether policy makers in other contexts should expect to see the same results, that is, whether they can be transferred successfully.

Finally, replication, systematic reviews, and meta-analysis may help in accreting knowledge using an array of data sources and methods. The discussion presented in Section 3 relies on a growing literature which employs both experimental and quasi-experimental methods. These studies have been examined following systematic review methodologies to provide a rigorous synthesis on the impact of various interventions and reforms to improve government performance. This synthesis itself speaks to a middle ground between experimental and non-experimental methods that forms the common theme of this symposium.

Acknowledgments: This UNU-WIDER symposium was developed under the project on “Experimental and Non-experimental Methods to Study Government Performance: Contributions and Limits,” which was supported under WIDER’s Research and Communication on Foreign Aid (ReCom) program (2011–2013). We gratefully acknowledge specific program contributions from the governments of Denmark (Ministry of Foreign Affairs, Danida) and Sweden (Swedish International Development Cooperation Agency – Sida) for ReCom, as well as core financial support to the UNU-WIDER work program from the governments of Denmark, Finland, Sweden, and the UK. We are grateful to Armando Barrientos, Kate Baldwin, Rikhil Bhavnani, Michael Bratton, Rajeev Dehejia, Macartan Humphreys, Fernando Martel Garcia, Javier Sajuría, Juan M. Villa, Leonard Wantchekon, and the Editors for helpful comments and suggestions on early versions of this article. The errors are ours alone.

Appendix

Table A1: Experimental and Quasi-experimental Studies to Study Government Performance.

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
Aggregating interests	Kroon et al.	Police Intervention in Riots: The Role of Accountability and Group Norms. A Field Experiment	1991	Journal article	Germany	Accountability	Improve Information asymmetries	Yes	MANOVA
	Duflo et al.	Efficiency and Rent Seeking in Local Government: Evidence from Randomized Policy Experiments in India	2005	Working paper	India	Accountability and corruption	Improve supply-side capabilities	Yes	OLS
	Janssen et al.	Coordination and Cooperation in Asymmetric Commons Dilemmas	2011	Journal article	USA	Cooperation and participation	Change behavior via incentives	Yes	OLS
	Asthana	Decentralization and Corruption Revisited: Evidence from a Natural Experiment	2012	Journal article	India	Corruption	Improve supply-side capabilities	No	DID, OLS
	Tyrefors	Do Merging Local Governments' Free Ride on their Counterparts when Facing Boundary Reform?	2009	Journal article	Sweden	Corruption	Change behavior via incentives	No	DID
	Bertrand et al.	Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption	2007	Journal article	India	Corruption	Change behavior via incentives	Yes	FE
	Fried et al.	Corruption and Inequality at the Crossroad: A Multi-method Study of Bribery and Discrimination in Latin America	2010	Journal article	Mexico	Corruption	Change behavior via incentives	Yes	OLS

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Olken	Monitoring Corruption: Evidence from a Field Experiment in Indonesia	2007	Journal article	Indonesia	Corruption	Change behavior via incentives (monitoring)	Yes	OLS, FE
	Azfar and Nelson	Transparency, Wages, and the Separation of Powers: An Experimental Analysis of Corruption	2007	Journal article	USA	Corruption	Change behavior via incentives (voters)	Yes	OLS, Probit, and RE
	Kalyitits and Vlachaki	When Does More Aid Imply Less Democracy? An Empirical Examination	2011	Journal article		Democracy	Improve supply-side capabilities	No	2SLS
	Glennerster et al.	Working Together: Collective Action in Diverse Sierra Leone Communities	2009	Working paper	Sierra Leone	Ethnic diversity	Improve supply-side capabilities	No	IV, OLS
	Habyarimana et al.	Why Does Ethnic Diversity Undermine Public Goods Provision?	2007	Journal article	Uganda	Ethnic diversity	Improve supply-side capabilities	Yes	OLS, FE
	Paul	Relief Assistance to 1998 Flood Victims: A Comparison of the Performance of the Government and NGOs	2003	Journal article	Bangladesh	Institutional building	Improve supply-side capabilities	No	ANOVA
	Roy	Governance and Development: The Challenges for Bangladesh	2005	Journal article	Bangladesh	Institutional building	Improve supply-side capabilities	No	PCA, 2SLS and OLS
	Beath et al.	Winning Hearts and Minds through Development: Evidence from a Field Experiment in Afghanistan	2012	Working paper	Afghanistan	Institutional building	Change behavior via incentives (grants and councils organization)	Yes	FE and PSM

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Scott	Assessing Determinants of Bureaucratic Discretion: An Experiment in Street level Decision Making	1997	Journal article	USA	Institutional building	Improve Information asymmetries	Yes	OLS
	Djankov et al.	Disclosure by Politicians	2010	Journal article		Institutional building	Improve supply-side capabilities	No	OLS
	Lokshin and Yemtsov	Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor?	2005	Journal article	Georgia	Institutional building	Improve supply-side capabilities	No	PSM, DID
	Chattopadhyay and Duflo	Women as Policy Makers: Evidence from a Randomized Policy Experiment in India	2004	Journal article	India	Institutional building + Participation	Change behavior via incentives (women quotas)	Yes	OLS
	Wallen et al.	Implementing Evidence-based Practice: Effectiveness of a Structured Multifaceted Mentorship Programme	2010	Journal article	USA	Institutional building	Change behavior via incentives	No	Correlational tests, and parametric stats
	Olson et al.	Governance and Growth: A Simple Hypothesis Explaining Cross-Country Differences in Productivity Growth	2000	Journal article		Institutional building + productivity	Improve supply-side capabilities	No	FE
	Lassen and Seiritzlew	Jurisdiction Size and Local Democracy: Evidence on Internal Political Efficacy from Large-scale Municipal Reform	2011	Journal article	Denmark	Institutional building	Improve supply-side capabilities	No	DID and PSM

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Wang et al.	Fiscal Reform and Public Education Spending: A Quasi-natural Experiment of Fiscal Decentralization in China	2012	Journal article	China	Institutional building	Improve supply-side capabilities	No	DiD, with RE and FE
	Bjørnskov et al.	The Bigger the Better? Evidence of the Effect of Government Size on Life Satisfaction around the World	2007	Journal article	Worldwide	Institutional building	Improve supply-side capabilities	No	OLS and 2LSL
	Pettersson-Lidbom	Does the Size of the Legislature Affect the Size of Government? Evidence from Two Natural Experiments	2011	Journal article	Finland and Sweden	Institutional building	Improve supply-side capabilities	No	RD
	Alence	Political Institutions and Developmental Governance in Sub-Saharan Africa	2004	Journal article	sub-Saharan Africa	Institutional building	Improve supply-side capabilities	No	WLS
	Corazzini et al.	A Prize To Give For: An Experiment on Public Good Funding Mechanisms	2010	Journal article		Institutional building	Improve supply-side capabilities	Yes	OLS
	Matkin and Frederickson	Metropolitan Governance: Institutional Roles and Inter-jurisdictional Cooperation	2009	Journal article	USA	Institutional building	Change behavior via incentives	Yes	Ordered Logit
	Thurmaier	Budgetary Decision-making in Central Budget Bureaus: An Experiment	1992	Journal article	USA	Institutional building	Improve Information asymmetries	Yes	OLS
	Drummond and Mansoor	Macroeconomic Management and the Devolution of Fiscal Powers	2003	Journal article	International	Institutional building	Improve supply-side capabilities	No	OLS, Cluster analysis

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Gibson et al.	Why Do People Accept Public Policies They Oppose? Testing Legitimacy Theory with a Survey-Based Experiment	2005	Journal article	USA	Institutional building	Improve supply-side capabilities	Yes	OLS
	Olken	Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia	2010	Journal article	Indonesia	Institutional building	Improve supply-side capabilities	Yes	Logit
	Yip and Eggleston	Addressing Government and Market Failures with Payment Change Behaviour via Incentives: Hospital Reimbursement Reform in Hainan, China	2004	Journal article	China	Institutional building	Change behavior via incentives (reform)	No	DID
	Batista and Vicente	Do Migrants Improve Governance at Home? Evidence from a Voting Experiment	2011	Journal article	Cape Verde	Institutional building + accountability	Improve Information asymmetries	No	IV, FE
	Grimmelikhuijsen	Linking Transparency, Knowledge and Citizen Trust in Government: an Experiment	2012	Journal article	Netherlands	Institutional building + accountability	Improve Information asymmetries	Yes	ANOVA, MANCOVA
	Korberg and Clarke	Beliefs About Democracy and Satisfaction with Democratic Government: The Canadian Case	1994	Journal article	Canada	Institutional building + democracy	Improve supply-side capabilities	No	Probit
	Cummings et al.	Tax Morale Affects Tax Compliance: Evidence from Surveys and an Art Factual Field Experiment	2009	Journal article	Botswana and South Africa	Institutional building (tax compliance)	Improve supply-side capabilities	Yes	Ordered Probit, Tobit, RE

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Banerjee et al.	Can Institutions be Reformed from Within? Evidence from a Randomized Experiment with the Rajasthan Police	2012	Working paper	India	Institutional building	Improve supply-side capabilities	Yes	FE, DID
	Humphreys et al.	The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe	2006	Journal article	São Tomé and Príncipe	Leadership	Improve supply-side capabilities	Yes	Leader FE
	Moxnes and van der Heijden	The Effect of Leadership in a Public Bad Experiment	2003	Journal article	Norway	Leadership	Change behavior via incentives	Yes	Pooled OLS
	Edmark	Migration Effects of Welfare Benefit Reform	2009	Journal article	Sweden	Migration	Change behavior via incentives	No	DID, FE
	Bowles and Gintis	Social Capital and Community Governance	2002	Journal article		Non-electoral forms of participation	Improve supply-side capabilities	Yes	OLS
	Travers et al.	Change Behaviour via Incentives for Cooperation: The Effects of Institutional Controls on Common Pool Resource Extraction in Cambodia	2011	Journal article	Cambodia	Non-electoral forms of participation	Improve supply-side capabilities	Yes	FE
	Cavalcanti et al.	Public Participation and Willingness to Cooperate in Common-pool Resource Management: A Field Experiment with Fishing Communities in Brazil	2009	Journal article	Brazil	Non-electoral forms of participation	Improve Information asymmetries	Yes	OLS, ordered Logit, ANOVA

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Yang and Pandey	Further Dissecting the Black Box of Citizen Participation: When Does Citizen Involvement Lead to Good Outcomes?	2011	Journal article	USA	Non-electoral forms of participation	Improve supply-side capabilities	No	OLS
	Dimitropoulos and Kontoleon	Assessing the Determinants of Local Acceptability of Wind-farm Investment: A Choice Experiment in the Greek Aegean Islands	2009	Journal article	Greece	Non-electoral forms of participation	Change behavior via incentives	Yes	RPL and MLM
	Slomczynski and Shabad	Can Support for Democracy and the Market Be Learned in School? A Natural Experiment in Post-Communist Poland	1998	Journal article	Poland	Non-electoral forms of participation	Change behavior via incentives	No	ANOVA
	Gugerty and Kremer	Outside Funding and the Dynamics of Participation in Community Associations	2008	Journal article	Kenya	Non-electoral forms of participation	Change behavior via incentives	No	OLS, Probit
	Beaman et al.	Powerful Women: Does Exposure Reduce Bias?	2009	Journal article	India	Non-electoral forms of participation	Change behavior via incentives (women quotas)	No	FE
	Fearon et al.	Can Development Aid Contribute to Social Cohesion after Civil War? Evidence from a Field Experiment in Post-Conflict Liberia	2009	Journal article	Liberia	Social Cohesion	Change behavior via incentives	Yes	PSM
	Hyde	The Observer Effect in International Politics: Evidence from a Natural Experiment	2007	Journal article	Armenia	Voting behavior	Improve supply-side capabilities	Yes	OLS

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Ichino and Schündeln	Detering or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana	2012	Journal article	Ghana	Voting behavior	Improve supply-side capabilities	Yes	OLS, FE, IV
	Hyde	Experimenting in Democracy Promotion: International Observers and the 2004 Presidential Elections in Indonesia	2010	Journal article	Indonesia	Voting behavior	Improve supply-side capabilities	Yes	OLS, FE
	De La O	Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico	2013	Journal article	Mexico	Voting behavior	Change behavior via incentives (benefits)	Yes	DID
	Bhavnani	Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India	2009	Journal article	India	Voting behavior	Change behavior via incentives (female quota)	No	Logit
	Humphreys and Weinstein	Policing Politicians: Citizen Empowerment and Political Accountability in Africa	2007	Conference (APSA 2007)	Uganda	Voting behavior	Improve Information asymmetries	No	Ordered Probit
	James	Performance Measures and Democracy: Improve Information asymmetries Effects on Citizens in Field and Laboratory Experiments	2011	Journal article	England	Voting behavior	Improve Information asymmetries	Yes	Probit
	Collier and Vicente	Votes and Violence: Evidence from a Field Experiment in Nigeria	2008	Working paper	Nigeria	Voting behavior	Improve Information asymmetries	Yes	DID, FE, Probit

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Esterling et al.	Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment with Members of Congress and Their Constituents	2011	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	Bayesian framework using MCMC with data augmentation
	Banerjee et al.	Do Informed Voters Make Better Choices? Experimental Evidence from Urban India	2011	Working paper	India	Voting behavior	Improve Information asymmetries	Yes	FE
	Ferraz and Finan	Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes	2008	Journal article	Brazil	Voting behavior	Improve Information asymmetries	Yes	FE, DiD, OLS, semi-parametric estimators
	Chong et al.	Improve Information asymmetries Dissemination and Local Governments' Electoral Returns, Evidence from a Field Experiment in Mexico	2010	Working paper	Mexico	Voting behavior	Improve Information asymmetries	Yes	FE, OLS
	Banerjee et al.	Can Voters be Primed to Choose Better Legislators? Evidence from Two Field Experiments in Rural India	2009	Working paper	India	Voting behavior	Improve Information asymmetries	Yes	IV, FE
	Harbridge and Malhotra	Electoral Change Behaviour via Incentives and Partisan Conflict in Congress: Evidence from Survey Experiments	2011	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	PSM

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Borges and Clarke	Cues in Context: Analyzing the Heuristics of Referendum Voting with an Internet Survey Experiment	2008	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	multinomial Logit
	Vicente and Wantchekon	Clientelism and Vote Buying: Lessons from Field Experiments in African Elections	2009	Journal article	West African	Voting behavior	Improve Information asymmetries	Yes	OLS
	Davenport	Public Accountability and Political Participation: Effects of a Face-to-Face Feedback Intervention on Voter Turnout of Public Housing Residents	2010	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	OLS, Probit and 2SLS
	Giné and Mansuri	Together We Will: Evidence from a Field Experiment on Female Voter Turnout in Pakistan	2011	Working paper	Pakistan	Voting behavior	Improve Information asymmetries	Yes	OLS and FE
	Gerber et al.	Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment	2010	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	OLS, 2SLS
	Gerber et al.	Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment	2008	Journal article	USA	Voting behavior	Improve Information asymmetries	Yes	OLS, FE
	Guan and Green	Noncoercive Mobilization in State-Controlled Elections	2006	Journal article	China	Voting behavior	Improve Information asymmetries	Yes	OLS
	Wantchekon	Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin	2003	Journal article	Benin	Voting behavior	Improve Information asymmetries	Yes	Probit

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Fujiwara and Wantchekon	Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin	2013	Journal article	Benin	Voting behavior	Improve Information asymmetries	Yes	Probit, FE
	Pande	Can Informed Voters Enforce Better Governance? Experiments in Low-Income Democracies	2011	Journal article		Voting behavior	Improve Information asymmetries	Yes	OLS
Provision of public goods	Buttenheim et al.	Impact Evaluation of School Feeding Programmes in Lao PDR	2011	Working paper	Lao	Education	Improve supply-side capabilities	No	DID, PSM and FE
	Koball	Living Arrangements and School Dropout Among Minor Mothers Following Welfare Reform	2007	Journal article	USA	Education	Change behavior via incentives	No	DID
	Singer and Stater	Going, Going, Gone: The Effects of Aid Policies on Graduation at Three Large Public Institutions	2006	Journal article	USA	Education	Change behavior via incentives	No	OLS
	Baulch	The Medium-Term Impact of the Primary Education Stipend in Rural Bangladesh	2010	Discussion Paper	Bangladesh	Education	Change behavior via incentives	No	PSM, DID
	Duflo and Hanna	Monitoring Works: Getting Teachers to Come to School	2005	Working paper	India	Education	Change behavior via incentives	Yes	OLS and 2SLS
	Kremer et al.	Change Behaviour via Incentives to Learn	2009	Journal article	Kenya	Education	Change behavior via incentives	Yes	OLS
	Kremer et al.	Decentralization: A Cautionary Tale	2003	Working paper	Kenya	Education	Change behavior via incentives	Yes	OLS, RE

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Kim and Joo	Did PRWORA's Mandatory School Attendance Policy Increase Attendance Among Targeted Teenage Girls?	2011	Journal article	USA	Education	Change behavior via incentives (benefits)	No	DID
	Reinikka and Svensson	The Power of Improve Information Asymmetries in Public Services: Evidence from Education in Uganda	2011	Journal article	Uganda	Education	Improve Information asymmetries	No	FE, IV
	Reinikka and Svensson	Local Capture: Evidence from a Central Government Transfer Programme in Uganda	2004	Journal article	Uganda	Education	Improve Information asymmetries	No	FE, RE
	Dee	Conditional Cash Penalties in Education: Evidence from the Learn Fare Experiment	2011	Journal article	USA	Education	Change behavior via incentives	Yes	FE
	Pandey et al.	Community Participation in Public Schools: Impact of Improve Information Asymmetries Campaigns in Three Indian states	2009	Journal article	India	Education + Accountability	Improve Information asymmetries	Yes	DID
	Patrinos and Sakellariou	Schooling and Labor Market Impacts of a Natural Policy Experiment	2005	Journal article	Venezuela	Education + Employment	Change behavior via incentives (policy)	No	IV
	Meghir et al.	Education, Health and Mortality: Evidence from a Social Experiment	2011	Working paper	Sweden	Education + health	Improve supply-side capabilities	No	CPHR, OLS
	Gao et al.	How Does Public Assistance Affect Family Expenditures? The Case of Urban China	2010	Journal article	China	Education + Health	Change behavior via incentives	No	PSM

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Monsiad et al.	Education and Fertility: Evidence from a Natural Experiment	2008	Journal article	Norway	Education + Health	Change behavior via incentives (reform)	No	2SLS
	Miguel and Kremer	Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities	2004	Journal article	Kenya	Education and Health	Improve supply-side capabilities	Yes	OLS
	Fan et al.	Propensity Score Techniques to Evaluate Returns of College Education	2009	Conference Paper	China	Employment	Improve supply-side capabilities	No	PSM
	Micklewright and Nagy	The Effect of Monitoring Unemployment Insurance Recipients on Unemployment Duration: Evidence from a Field Experiment	2009	Journal article	Hungary	Employment	Improve supply-side capabilities	Yes	CPHR, OLS
	Dufló and Saez	The Role of Improve Information Asymmetries and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment	2003	Journal article	USA	Employment	Change behavior via incentives	Yes	OLS, FE and IV
	van Ours and Vodopivec	How Shortening the Potential Duration of Unemployment Benefits Affects the Duration of Unemployment: Evidence from a Natural Experiment	2006	Journal article	Slovenia	Employment	Change behavior via incentives (benefits)	No	DID

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Campolietia and Riddell	Disability Policy and the Labor Market: Evidence from a Natural Experiment in Canada, 1998-2006	2011	Journal article	Canada	Employment	Change behavior via incentives (earning exceptions)	No	DiD
	Washbrook et al.	Public Policies, Women's Employment after Childbearing, and Child Well-Being	2011	Journal article	USA	Employment	Change behavior via incentives (policy)	No	DiD
	Ziebarth and Karlsson	A Natural Experiment on Sick Pay Cuts, Sickness Absence, and Labor Costs	2010	Journal article	Germany	Employment	Change behavior via incentives (reform)	No	DiD, PSM
	Boeri and Tabellini	Does Information Increase Political Support For Pension Reform?	2010	Journal article	Italy	Employment	Improve Information asymmetries	Yes	Probit, PSM
	Zhou et al.	Community Effectiveness of Stove and Health Education Interventions for Reducing Exposure to Indoor Air Pollution from Solid Fuels in Four Chinese Provinces	2006	Journal article	China	Health	Improve Information asymmetries	No	DiD
	Liu et al.	The Expansion of Public Health Insurance and the Demand for Private Health Insurance in Rural China	2010	Journal article	China	Health	Improve supply-side capabilities	No	DiD
	Cooper et al.	Does Competition Improve Public Hospitals' Efficiency? Evidence from a Quasi-Experiment in the English National Health Service	2012	Discussion Paper	England	Health	Improve supply-side capabilities	No	DiD, FE

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Jesmin et al.	Does Team-based Primary Health Care Improve Patients' Perception of outcomes? Evidence from the 2007-08 Canadian Survey of Experiences with Primary Health	2012	Journal article	Canada	Health	Improve supply-side capabilities	No	PSM
	Chadwick et al.	Effects of Downsizing Practices on the Performance of Hospitals	2004	Journal article	USA	Health	Improve supply-side capabilities	No	OLS, Logit
	Steinhardt et al.	Removing User Fees for Basic Health Services: A Pilot Study And National Roll-out in Afghanistan	2011	Journal article	Afghanistan	Health	Change behavior via incentives	No	DID
	Basinga et al.	Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: an Impact Evaluation	2011	Journal article	Rwanda	Health	Change behavior via incentives	No	DID, FE
	Nguyen et al.	Encouraging Maternal Health Service Utilization: An Evaluation of the Bangladesh Voucher Programme	2012	Journal article	Bangladesh	Health	Change behavior via incentives	No	DID, FE
	Thanh et al.	An Assessment of the Implementation of the Health Care Funds for the Poor Policy in Rural Vietnam	2010	Journal article	Vietnam	Health	Change behavior via incentives	No	DID, PSM
	Schreyögg and Grabka	Co-payments for Ambulatory Care in Germany: a Natural Experiment Using A Difference-in-Difference Approach	2010	Journal article	Germany	Health	Change behavior via incentives	No	Logit, DID

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Amtz and Thomsen	Crowding Out Informal Care? Evidence from a Field Experiment in Germany	2011	Journal article	Germany	Health	Change behavior via incentives	Yes	DiD, FE, Probit
	Thanh et al.	Does “the Injury Poverty Trap” Exist?: A Longitudinal Study in Bavi, Vietnam	2006	Journal article	Vietnam	Health	Change behavior via incentives (injury)	No	PSM
	Quimbo et al.	Evidence of a Causal Link Between Health Outcomes, Insurance Coverage, and a Policy to Expand Access: Experimental Data from Children in the Philippines	2011	Journal article	Philippines	Health	Change behavior via incentives (policy)	No	DiD
	Trujillo et al.	The Impact of Subsidized Health Insurance for the Poor: Evaluating the Colombian Experience Using Propensity Score Matching	2005	Journal article	Colombia	Health	Change behavior via incentives (policy)	No	PSM
	King et al.	A “Politically Robust” Experimental Design for Public Policy Evaluation, With Application to the Mexican Universal Health Insurance programme	2007	Journal article	Mexico	Health	Improve supply-side capabilities	Yes	PSM, OLS
	Propper et al.	Change Behavior Via Incentives and Targets in Hospital Care: Evidence From A Natural Experiment	2010	Journal article	UK	Health	Change behavior via incentives (policy)	No	DiD

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Rodriguez et al.	A Randomized Experiment of Issue Framing and Voter Support of Tax Increases for Health Insurance Expansion	2010	Journal article	USA	Health	Improve Information asymmetries	Yes	ANOVA, OLS
	Justesen	Democracy, Dictatorship, and Disease: Political Regimes and HIV/AIDS	2012	Journal article		Health + democracy	Improve supply-side capabilities	No	PSM
	Braakmann	The Causal Relationship Between Education, Health and Health Related Behaviour: Evidence from a Natural Experiment in England	2011	Journal article	England	Health + Education	Improve supply-side capabilities	No	OLS and IV
	Velamuri	Taxes, Health Insurance and Women's Self-employment	2012	Journal article	USA	Health + Employment	Change behavior via incentives (health insurance)	No	DID
	Behrman et al.	Evaluating Preschool Programs When Length of Exposure to the Programme Varies: A Nonparametric Approach	2004	Journal article	Bolivia	Health + Nutrition + Education	Change behavior via incentives (day-care, nutrition, education)	No	PSM
	Björkman and Svensson	Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda	2009	Journal article	Uganda	Health + participation	Improve Information asymmetries	Yes	FE, DID

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Poder and He	How can Infrastructures Reduce Child Malnutrition and Health Inequalities? Evidence from Guatemala	2011	Working paper	Guatemala	Health and nutrition	Improve supply-side capabilities	No	PSM
	Schwartz et al.	The External Effects of Place-based Subsidized Housing	2006	Journal article	USA	Housing	Change behavior via incentives (policy)	No	DID, FE
	Desai and Tarozzi	Microcredit, Family Planning Programmes, and Contraceptive Behavior: Evidence From a Field Experiment in Ethiopia	2011	Journal article	Ethiopia	Microcredit + health	Improve Information asymmetries	Yes	DID, IV
	Banerjee et al.	Pitfalls of Participatory Programmes: Evidence from a Randomized Evaluation in Education in India	2010	Journal article	India	Non-electoral forms of participation	Improve Information asymmetries	Yes	IV, Pooled OLS
	Galasso and Umapathi	Improving Nutritional Status through Behavioral Change: Lessons from Madagascar	2007	Working paper	Madagascar	Health + Nutrition	Improve Information asymmetries	No	DID, PSM, FE
	Domínguez-Torreiro and Soliño	Provided and Perceived Status Quo in Choice Experiments: Implications for Valuing the Outputs of Multifunctional Rural Areas	2011	Journal article	Spain	Rural development	Change behavior via incentives	Yes	OLS, RPL
	De Jaeger	Residual Household Waste: From Pay-Per-Bag To Pay-Per-Kilogram. An Evaluation Study for Flanders	2010	Journal article	Flanders	Sanitation	Improve supply-side capabilities	No	DID

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Gregg et al.	Family Expenditures Post-welfare reform in the UK: Are Low-Income Families Starting To Catch Up?	2006	Journal article	UK	Social Protection	Improve supply-side capabilities	No	DID
	Leroy et al.	The Oportunidades Programme Increases the Linear Growth of Children Enrolled at Young Ages in Urban Mexico	2008	Journal article	Mexico	Social Protection	Change behavior via incentives	No	DID, PSM
	Baird et al.	Cash or Condition? Evidence from a Cash Transfer Experiment	2010	Working paper	Malawi	Social Protection	Change behavior via incentives	Yes	OLS, FE
	de Hoop et al.	Do Cash Transfers Crowd Out Community Investment in Public Goods? Lessons from a Field Experiment on Health Education	2011	Journal article	Peru	Social Protection	Change behavior via incentives (funding health education)	Yes	OLS
	Stecklov et al.	Unintended Effects of Poverty Programmes on Childbearing in Less Developed Countries: Experimental Evidence from Latin America	2007	Journal article	Latin America	Social Protection	Change behavior via incentives	Yes	DID
	Pattanayak et al.	How Valuable Are Environmental Health Interventions? Evaluation of Water and Sanitation Programmes in India	2010	Journal article	India	Water and Sanitation	Improve supply-side capabilities	No	DID
	Hope	Evaluating Social Impacts of Watershed Development in India	2007	Journal article	India	Water and Sanitation	Improve supply-side capabilities	No	PSM

(Table A1: Continued)

Governance cluster	Author(s)	Title	Year	Type of publication	Country	Policy area	Aim of intervention	Experimental design?	Analytical methods*
	Veettil et al.	Complementarity Between Water Pricing, Water Rights And Local Water Governance: A Bayesian Analysis of Choice Behaviour of Farmers in the Krishna River Basin, India	2011	Journal article	India	Water and Sanitation	Change behavior via incentives (models)	Yes	OLS, Multinomial Probit
	Pattanayak et al.	Shame or Subsidy Revisited: Social Mobilization for Sanitation in Orissa, India	2009	Journal article	India	Water and Sanitation	Improve Information asymmetries	No	DID
	Kurz et al.	Utilizing a Social-Ecological Framework to Promote Water and Energy Conservation: A Field Experiment	2005	Journal article	Australia	Water and Sanitation	Improve Information asymmetries	Yes	ANOVA
	Watson et al.	An Opportunistic Field Experiment in Community Water Conservation	1999	Journal article	Australia	Water and Sanitation	Improve Information asymmetries	Yes	MANOVA, OLS
	Moehler	Democracy, Governance, and Randomized Development Assistance	2010	Journal article	World	Democratic governance	Improve supply-side capabilities	Yes	Review article

*Abbreviations stand as follows: PCA, principal component analysis; OLS, ordinary least squares; 2SLS, two-stage least squares; WLS, weighted-least-squares; FE, fixed effects; RE, random effects; PSM, propensity score matching; DID, difference-in-difference; ANOVA, analysis of variance; MANOVA, multivariate analysis of variance; RPL, random parameter logit; MLM, mixed logit model; RD, regression discontinuity; CPHR, cox proportional hazard regressions; RPL, random parameter logit.

Source: Adapted from Gisselquist et al. (2013).

References

- Angrist, Joshua D. (1990) "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *The American Economic Review*, 80(5):1284–1286.
- Arndt, Christiane and Charles Oman (2006) *Uses and Abuses of Governance Indicators*. Paris: OECD Development Centre.
- Baldwin, Kate and Rikhil R. Bhavnani (2015) "Ancillary Studies of Experiments: Opportunities and Challenges". *Journal of Globalization and Development*, 6(1):113–146.
- Banerjee, Abhijit and Esther Duflo (2012) *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.
- Banerjee, Abhijit V., Raghavendra Chattopadhyay, Esther Duflo, Daniel Keniston and Nina Singh (2012) *Can Institutions Be Reformed from Within? Evidence from a Randomized Experiment with the Rajasthan Police*. Working Paper 17912. Cambridge, MA: NBER.
- Banerjee, Abhijit Vinayak (2007) "Making Aid Work." In: (A. V. Banerjee, ed.) *Making Aid Work*. Cambridge, MA: MIT Press.
- Bardhan, Pranab (2005) *Scarcity, Conflicts and Cooperation: Essays in the Political Economy and Institutional Economics of Development*. Cambridge, MA: MIT Press.
- Barrientos, Armando and Juan M. Villa (2015) "Evaluating Antipoverty Transfer Programmes in Latin America and Sub-Saharan Africa. Better Policies? Better Politics?," *Journal of Globalization and Development*, 6(1):147–179.
- Basinga, Paulin, Paul J. Gertler, Agnes Binagwaho, Agnes L. B. Soucat, Jennifer Sturdy and Christel M. J. Vermeersch (2011) "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation," *The Lancet*, 377:1421–1428.
- Björkman, Martina and Jakob Svensson (2009) "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda," *The Quarterly Journal of Economics*, 124:735–769.
- Boix, C. and D. N. Posner (1998). "Social Capital: Explaining Its Origins and Effects on Government Performance," *British Journal of Political Science*, 28(4):686–693.
- Bovaird, Tony and Elke Löffler, eds. (2009) *Public Management and Governance*, 2nd ed. Abingdon: Routledge.
- Brancati, Dawn (2006) "Decentralization: Fuelling the Fire or Dampening the Flames of Ethnic Conflict and Secessionism?" *International Organization*, 60(3):651–685.
- Bratton, Michael (2013) *Measuring Government Performance in Public Opinion Surveys in Africa: Towards Experiments?* WIDER Working Paper. Helsinki: UNU-WIDER.
- Collier, Paul and Pedro C. Vicente (2008) *Votes and Violence: Evidence from a Field Experiment in Nigeria*. CSAE WPS/2008-16. Oxford: University of Oxford.
- Conn, K. (2014) *Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Rigorous Impact Evaluations*. Ph.D. dissertation, Columbia University.
- De La O. Ana (2013) "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico." *American Journal of Political Science*, 57(1):1–14.
- Deaton, Angus (2009) "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." Research Program in Development Studies, Center for Health and Wellbeing. Princeton, NJ: Princeton University.
- Dee, Thomas S. (2011) "Conditional Cash Penalties in Education: Evidence from the Learnfare Experiment," *Economics of Education Review*, 30:924–937.

- Dehejia, Rajeev (2015) “Experimental and Non-Experimental Methods in Development Economics: A Porous Dialectic,” *Journal of Globalization and Development*, 6(1):47–69.
- Duflo, Esther and Rema Hanna (2005) *Monitoring Works: Getting Teachers to Come to School*. Working Paper Series 11880. Cambridge, MA: NBER.
- Dunning, T., G. Grossman, M. Humphreys, S. Hyde, C. McIntosh, C. Adida, E. Arias, T. Boas, M. Buntaine, S. Bush, S. Chauchard, J. Gottlieb, F. D. Hidalgo, M. E. Holmlund, R. Jablonski, E. Kramon, H. Larreguy, M. Lierl, G. McClendon, J. Marshall, D. Nielson, M. Platas Izama, P. Querubin, P. Raffler and N. Sircar (2015). “Political Information and Electoral Choices: A Pre-meta-analysis Plan.” Available at: <http://egap.org/research/metaketa/>.
- Friedman, Milton (1962) *Capitalism and Freedom*. Chicago: University Of Chicago Press.
- Fujiwara, Thomas, and Leonard Wantchekon (2013) “Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin.” *American Economic Journal: Applied Economics*, 5(4): 241–55.
- Gans-Morse, Jordan and Simeon Nichter (2008) “Economic Reforms and Democracy Evidence of a J-Curve in Latin America.” *Comparative Political Studies*, 41:1398–1426.
- Gisselquist, Rachel M. (2012) *Good Governance as a Concept, and Why this Matters for Development Policy*. Working Paper 2012/30. Helsinki: UNU-WIDER.
- Gisselquist, Rachel M., Miguel Niño-Zarazúa and Javier Sajuria. (2013) “Improving Government Performance in Developing Countries. A Systematic Review.” Unpublished paper, UNU-WIDER, Helsinki.
- Glennerster, Rachel and Michael Kremer (2011) *Small Changes, Big Results: Behavioral Economics at Work in Poor Countries*. Boston Review: A Political and Literary Forum.
- Goldin, Kenneth D. (1977) “Equal Access vs. Selective Access: A Critique of Public Goods Theory,” *Public Choice*, 29:53–71.
- Grindle, Merilee S. (2010) *Good Governance: The Inflation of an Idea*. Faculty Research Working Paper Series RWP 10-023. Cambridge, MA: Harvard Kennedy School.
- Grindle, Merilee S. and John W. Thomas (1991) *Public Choices and Policy Change: The Political Economy of Reform in Developing Countries*. Baltimore: John Hopkins University Press.
- Herbst, Jeffrey (2000) *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.
- Heckman, James J. and Jeffrey A. Smith (1995) “Assessing the Case for Social Experiments,” *Journal of Economic Perspectives*, 9(2):85–110.
- Horowitz, Donald L. (1985) *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Humphreys, Macartan and Jeremy M. Weinstein (2009) “Field Experiments and the Political Economy of Development,” *The Annual Review of Political Science*, 12:367–378.
- Humphreys, Marcatan (2015) “Reflections on the Ethics of Social Experimentation,” *Journal of Globalization and Development*, 6(1):87–112.
- Hyde, Susan D. (2010a) “Experimenting in Democracy Promotion: International Observers and the 2004 Presidential Elections in Indonesia,” *Perspectives on Politics*, 8:511–527.
- Hyde, Susan D. (2010b) “The Future of Field Experiments in International Relations,” *The ANNALS of the American Academy of Political and Social Science*, 628:72–84.
- Ichino, Nahomi and Matthias Schündeln (2012) “Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana,” *The Journal of Politics*, 74:292–307.
- Immergut, Ellen M. (1992) *Health Politics: Interests and Institutions in Western Europe*. Cambridge: Cambridge University Press.

- Jackson, R. H. and C. G. Rosberg (1982) *Personal Rule in Black Africa*. Los Angeles: UCLA Press.
- Karlan, Dean and Jacob Appel (2012) *More Than Good Intentions: Improving the Ways the World's Poor Borrow, Save, Farm, Learn, and Stay Healthy*. New York: Plume.
- Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobaton (1999) *Governance Matters*. Policy Research Working Paper 2196. Washington, DC: World Bank.
- Keefer, Philip (2009) "Governance." In: (T. Landman and N. Robinson, eds.) *The Sage Handbook of Comparative Politics*. London: Sage.
- Keynes, John M. (1964) *The General Theory of Employment, Interest, and Money*. London: Harcourt Brace Jovanovich.
- Kingdon, John W. (1995) *Agendas, Alternatives, and Public Policies*, 2nd ed. New York: HarperCollins.
- Kirkpatrick, C. (2014) "Assessing the Impact of Regulatory Reform in Developing Countries," *Public Administration and Development*, 34(3):162–168.
- Kremer, Michael, Edward Miguel and Rebecca Thornton (2009) "Incentives to Learn," *Review of Economics and Statistics*, 91:437–456.
- Krishnaratne, S., H. White and E. Carpenter (2013) *Quality Education for all Children? What Works in Education in Developing Countries*. 3ie Working Paper. New Delhi: International Initiative for Impact Evaluation (3ie).
- Leroy, Jef L., Armando García-Guerra, Raquel García, Clara Dominguez, Juan Rivera and Lynnette M. Neufeld (2008) "The Oportunidades Program Increases the Linear Growth of Children Enrolled at Young Ages in Urban Mexico," *The Journal of Nutrition*, 138:793–798.
- Levi, Margaret (2006) "Why We Need A New Theory of Government," *Perspectives on Politics*, 4:5–19.
- Lijphart, Arend (1977) *Democracy in Plural Societies: A Comparative Exploration*. New Haven: Yale University Press.
- Lipset, S. M. (1981) *Political Man: The Social Bases of Politics*. Baltimore, MD: Johns Hopkins University Press.
- Luebbert, Gregory M. (1991) *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. Oxford: Oxford University Press.
- Masino, S. and M. Niño-Zarazúa (2015) *What Works to Improve the Quality of Student Learning in Developing Countries?* WIDER Working Paper 2015 (033).
- Martel García, Fernando and Leonard Wantchekon (2015) "A Graphical Approximation to Generalization: Definitions and Diagrams," *Journal of Globalization and Development*, 6(1):71–86.
- McEwan, P. J. (in press) "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments," *Review of Educational Research*.
- Migdal, Joel (1988) *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton: Princeton University Press.
- Miguel, Edward and Michael Kremer (2004) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72:159–217.
- Moehler, Devra C. (2010) "Democracy, Governance, and Randomized Development Assistance," *The ANNALS of the American Academy of Political and Social Science*, 628:30–46.
- Moore, Barrington (1966) *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. New York: Beacon Press.
- Mukunda, G. (2012) *Indispensable: When Leaders Really Matter*. Cambridge: Harvard Business Review Press.

- North, Douglass C. (1990) *Institutions, Institutional Change and Economic Performance*. Cambridge, MA: Cambridge University Press.
- Olken, Benjamin A. (2007) “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” *Journal of Political Economy*, 115:200–249.
- Olken, Benjamin A. (2010) “Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia,” *American Political Science Review*, 104:243–267.
- Olken, Benjamin A. and Rohini Pande (2011a) *Corruption in Developing Countries*. Working Paper 17398. Cambridge, MA: NBER.
- Olken, Benjamin A. and Rohini Pande (2011b) *Governance Review Paper*. J-PAL Governance Initiative: Abdul Latif Jameel Poverty Action Lab.
- Pandey, Priyanka, Sangeeta Goyal and Venkatesh Sundararaman (2009) “Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States,” *Education Economics*, 17:355–375.
- Pattanayak, Subhrendu K., Jui-Chen Yang, Katherine L. Dickinson, Christine Poulos, Sumeet R. Patil, Ranjan K. Mallick, Jonathan L. Blitstein and Purujit Praharaj (2009) “Shame or Subsidy Revisited: Social Mobilization for Sanitation in Orissa, India,” *Bulletin of the World Health Organization*, 87:580–587.
- Petrosino, A., C. Morgan, T. Fronius, E. Tanner-Smith and R. Boruch (2012) “Interventions in Developing Nations for Improving Primary and Secondary School Enrollment of Children: A Systematic Review,” *Campbell Systematic Reviews*, 19:1–191.
- Pritchett, Lant and Justin Sandefur (2013) *Context Matters for Size: Why External Validity Claims and Development Practice Don’t Mix*. Working Paper 336. Washington, DC: Center for Global Development.
- Przeworski, Adam and Fernando Limongi (1997) “Modernization: Theories and Facts,” *World Politics*, 49:155–183.
- Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub and Fernando Limongi (2000) *Democracy and Development: Political Institutions and Well-being in the World, 1950–1990*. New York: Cambridge University Press.
- Putnam, Robert (1993) *Making Democracy Work: Civic Traditions in Modern Italy*. New York: Princeton University Press.
- Putnam, R. D. (2007) “E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture,” *Scandinavian Political Studies*, 30(2):137–174.
- Quimbo, Stella A., John W. Peabody, Riti Shimkhada, Jhiedon Florentino and Orville Solon (2011) “Evidence of a Causal Link Between Health Outcomes, Insurance Coverage, and a Policy to Expand Access: Experimental Data from Children in the Philippines,” *Health Economics*, 20:620–630.
- Ravallion, Martin (2009) “Should the Randomistas Rule?” *Economists’ Voice*, 6:1–5.
- Reilly, Benjamin (2001) *Democracy in Divided Societies: Electoral Engineering for Conflict Management*. Cambridge: Cambridge University Press.
- Robinson, Mark (2007) “Does Decentralisation Improve Equity and Efficiency in Public Service Delivery Provision?,” *IDS Bulletin*, 38(1):7–17.
- Rothstein, B. O. and J. A. N. Teorell (2008) “What Is Quality of Government? A Theory of Impartial Government Institutions,” *Governance*, 21:165–190.
- Samuels, Richard J. (2003) *Machiavelli’s Children: Leaders and Their Legacies in Italy and Japan*. Ithaca: Cornell University Press.
- Sartori, Giovanni (1997) *Comparative Constitutional Engineering: An Inquiry into Structures, Incentives, and Outcomes*. New York: New York University Press.

- Seabright, Paul (1996) "Accountability and Decentralization in Government: An Incomplete Contracts Model," *European Economic Review*, 40(1):61–89.
- Seguin, M. and M. Niño Zarazúa (2015) "Non-clinical Interventions for Acute Respiratory Infections and Diarrhoeal Diseases Among Young Children in Developing Countries," *Tropical Medicine & International Health*, 20(2):146–169.
- Sekhon, J. S. and R. Titiunki (2012) "When Natural Experiments Are Neither Natural nor Experiments," *American Political Science Review*, 106(1):1–23.
- Skocpol, Theda (1979) *States and Social Revolutions: A Comparative Analysis of France, Russia and China*. Cambridge: Cambridge University Press.
- Stecklov, Guy, Paul Winters, Jessica Tood and Ferdinando Regalia (2007) "Unintended Effects of Poverty Programmes on Childbearing in Less Developed Countries: Experimental Evidence from Latin America," *Population Studies*, 61:125–140.
- Steinmo, Sven (2008) "Historical Institutionalism." In: (D. Della Porta and M. Keating, eds.) *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective*. Cambridge: Cambridge University Press.
- Thomas, M. A. (2009) "What Do the Worldwide Governance Indicators Measure?" *European Journal of Development Research*, 22:31–54.
- Tocqueville, Alexis de (2003) *Democracy in America and Two Essays on America*. Translated by Gerald E. Bevan. London: Penguin Books.
- Varshney, A. (2001) "Ethnic Conflict and Civil Society – India and Beyond," *World Politics*, 53:362–398.
- Weber, Max (2009) "Politics as a Vocation." In: (H. H. Gerth and C. Wright Mills, eds.) *From Max Weber: Essays in Sociology*. Abingdon: Routledge.
- World Bank (2012) *Strengthening Governance: Tackling Corruption. The World Bank Group's Updated Strategy and Implementation Plan*. Washington, DC: World Bank.