



UNITED NATIONS
UNIVERSITY

UNU-WIDER

World Institute for Development
Economics Research

WIDER Working Paper No. 2013/077

What can experiments tell us about how to improve governance?

Rachel M. Gisselquist and Miguel Niño-Zarazúa*

August 2013

Abstract

In recent years, randomized controlled trials have become increasingly popular in the social sciences. In development economics in particular, their use has attracted considerable debate in relation to the identification of ‘what works’ in development policy. This paper focuses on a core topic in development policy: governance. It aims to address two key questions: (1) ‘what have the main contributions of randomized controlled trials been to the study of governance?’ and (2) ‘what could be the contributions, and relatedly the limits of such methods?’. To address these questions, a systematic review of experimental and quasi-experimental methods to study government performance was conducted. It identified 139 relevant papers grouped into three major types of policy interventions that aim to: (1) improve supply-side capabilities of governments; (2) change individual behaviour through various devices, notably incentives, and (3) improve informational asymmetries. We find .../

Keywords: randomised control trials, governance, development

JEL classification: C93, D72, D73, H41

Copyright © UNU-WIDER 2013

*both authors UNU-WIDER, corresponding author email: miguel@wider.unu.edu

This study has been prepared within the UNU-WIDER project ‘ReCom—Research and Communication on Foreign Aid’, directed by Tony Addison and Finn Tarp.

UNU-WIDER gratefully acknowledges specific programme contributions from the governments of Denmark (Ministry of Foreign Affairs, Danida) and Sweden (Swedish International Development Cooperation Agency—Sida) for ReCom. UNU-WIDER also gratefully acknowledges core financial support to its work programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

ISSN 1798-7237

ISBN 978-92-9230-654-0



that randomized controlled trials can be useful in studying the effects of some policy interventions in the governance area, but they are limited in significant ways: they are ill-equipped to study broader governance issues associated with macro-structural shifts, national level variation in institutions and political culture, and leadership. Randomized controlled trials are best for studying targeted interventions, particularly in areas of public goods provision, voting behaviour, and specific measures to address corruption and improve accountability; however, they can provide little traction on whether the intervention is transferable and ‘could work’ (and why) in other contexts, and in the longer run.

Acknowledgements

We thank Danielle Resnick, Macartan Humphreys, and participants of the UNU-WIDER project on ‘Experimental and non-experimental methods in the study of government performance’ for valuable input and comments on earlier versions of this paper. The views presented here are those of the authors alone and do not represent the views or policies of UNU-WIDER.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy-making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

www.wider.unu.edu

publications@wider.unu.edu

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Typescript prepared by Lisa Winkler at UNU-WIDER.

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute or the United Nations University, nor by the programme/project sponsors, of any of the views expressed.

1 Introduction

Experimental studies using randomized controlled trials (RCTs) have long been a staple of medical research. In recent years, these methods have also become increasingly popular in the social sciences. In development economics in particular, their use has attracted considerable debate with some scholars promoting them as the best means of identifying ‘what works’ in development policy (Banerjee 2007; Glennerster and Kremer 2011), while others voice strong concerns about their growing hegemony in the field (see e.g., Deaton 2009; Ravallion 2009).

This paper focuses on the use of such methods in identifying ‘what works’ for one of the major topics in contemporary studies of development policy: governance. It asks two key questions: (1) what have the main contributions of RCTs been to the study of governance? and (2) what could be the contributions, and relatedly, the limits of such methods? Despite large separate literatures on governance and on experimental methods, very little work has directly considered both together in this way. This paper draws on reviews of both literatures, including a systematic review of experimental and quasi-experimental studies of government performance that is described further below.

Broadly, this paper argues that RCTs have some, but limited utility in the study of governance. It discusses some of the key contributions that RCTs have made in the study of targeted interventions with relatively rapid results; however it also shows that major hypotheses about, e.g., macro-structural and cultural shifts over long periods of time, national level policy changes, and changes in political leadership are not amenable to study using RCTs.. Such limitations suggest that researchers should expect to use other methods to address these important areas in the field of development studies.

This paper is divided into four parts. The first focuses on theories of governance, highlighting several major hypotheses from the literature about how the quality of governance changes. The second focuses on how RCTs have been used in the study of governance-related topics, highlighting some of the 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 paper brings these two sections together, exploring whether and how RCTs could be used to address major theories of governance. A final section concludes.

Discussion of experimental research, particularly in economics and political science, sometimes treat laboratory-type experiments, natural experiments, and RCTs or ‘field experiments’ together, irrespectively of their design features. Much of the discussion in this paper 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 amenable to this method. Outside of the laboratory setting, field experiments using 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 (‘difference in difference’). 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 experiment.

2 Explaining governance

Despite a wealth of literature on governance, even its definition remains contested (Gisselquist 2012; Keefer 2009). A large amount of the literature focuses on definition, conceptualization, and measurement (see, e.g., Arndt and Oman 2006; Kaufmann, Kraay, and Zoido-Lobaton 1999; Rothstein and Teorell 2008; Thomas 2009). Without delving too much into these debates, this paper adopts a basic definition 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 to the poor and vulnerable, 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.

Public goods are better provided collectively than individually for reasons of efficiency and necessity (Goldin 1977; Samuelson 1954). As Putnam (1993) notes in his classic *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’ (Putnam 1993: 8-9). 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 somehow ‘aggregating’ and representing such interests to make collective decisions. In short, as Levi (2006) 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’ (Levi 2006: 5). By extension, the quality of governance, as understood here, varies in the degree to which governments fulfil these two related roles.

Theories of government and the state suggest a number of explanations about why the quality of governance in this sense varies, both across polities and over time, highlighting a range of structural, institutional, and cultural factors, as well as individual agency. In general, this work deals with the two roles of government 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 focuses on explaining disaggregated governance 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, different theoretical traditions offer different and sometimes contradictory explanations for key governance outcomes. One example important both for theory and contemporary politics is what

constitutes good governance in terms of providing the institutional environment most conducive to economic growth. Friedman (1962), for instance, 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 ‘good’ governments in fiscal and monetary policy. Similarly, there are major debates over whether more or less regulation is most conducive to private sector development, and over the form that regulation should take (see Kirkpatrick 2012).

One of the major structural factors highlighted in explanations of variation in the quality of governance is ‘modernization’ or the level of development. Max Weber, for instance, suggests that modernization leads to fundamental changes in the nature of authority, from traditional and charismatic towards rational-legal (the rule of law) (see Weber 2009). Modernization theory of the 1950s and 1960s highlighted how economic growth led to fundamental structural changes in the economy and society, such as the growth of the middle class and middle class ideology, and the emergence of ‘cross pressures’, that led to greater popular participation in government and created the foundations for the emergence of democracy (see Lipset 1981). Later work has also highlighted modernization as a key factor in democratic governance, but challenged the specific mechanism proposed by the modernization theorists. Przeworski and Limongi (1997), for instance, argue that it is not that modernization leads to the emergence of democratic governance, but that democracy is more likely to endure (once it is born for other reasons) in countries at higher levels of economic development – in particular, they found, in countries with annual per capita incomes higher than US\$6,000 (in 1985 dollars) (see also Przeworski et al. 2000).

Other structural arguments highlight modernization and the class structure, positing different mechanisms. Moore (1966), for instance, argues that ‘the ways in which the landed upper classes and the peasants reacted to the challenge of commercial agriculture were decisive factors in determining the political outcome’ (Moore 1966: xxiii). As agrarian societies transformed, resulting bourgeois revolutions led to capitalist liberal democracy (e.g., England, France, and the United States), abortive bourgeois revolutions led to fascism (e.g., Japan), and peasant revolutions led to communism (e.g., Russia and China). Focusing also on France, Russia, and China, Skocpol (1979), by contrast, highlights the autonomous role of the state in relation both to domestic class and political forces and to other states. In all three cases, she argues, social revolutions led leaders to strengthen, centralize, and rationalize state organizations (public institutions). The different character of resulting political regimes is explained by variation in the socioeconomic legacies of the old regimes, international circumstances, and the ideology and process of state-building after the revolution. Another structural argument highlighting the same factors is proposed by Luebbert (1991), who explains four different regime types in interwar Europe as resulting from the path taken from preindustrial politics to the crises of the 1920s and 1930s and particular constellations of urban-rural coalitions: a centre-right coalition and the early inclusion of the working class led to liberal democratic governance (Britain, Switzerland, France), for instance, while an alliance of the urban working class and the middle peasantry led to social democracy or democratic corporatism (Norway, Sweden, Denmark, Czechoslovakia). Other structural arguments highlight still other factors, such as the ethnic structure of society (e.g., Horowitz 1985), geography (e.g., Herbst 2000), and the strength of society relative to the state (e.g., Migdal 1988), which may affect how difficult it is to govern a particular polity.

Institutional explanations for governance outcomes are among the most diverse, highlighting a range of institutions and mechanisms. Indeed, social scientists often define institutions so

broadly – as formal and informal rules, norms, and organizations – that some of the ‘structural’ explanations reviewed above and the cultural explanations reviewed below are sometimes 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’s 2012 strategy on governance, for instance, highlights its role in helping countries to ‘put in place institutions and systems that can become the foundations of sustainable growth’ (World Bank 2012). It highlights both the need 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 favourable to private sector development).

One of North’s (1990) key arguments that is echoed in many institutional explanations is the effect of institutional ‘lock in’ or path dependence that makes changing institutions 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.

Other institutionalist work adopts more historical or sociological perspectives. Focusing on governance with respect to health care policy, for instance, Immergut (1992) argues that the structure of political institutions in Sweden, France, and Switzerland influenced whether they developed comprehensive national health care or more fragmented insurance programmes. Political institutions and procedures, rather than the demands of social groups, set the terms of political negotiations, leading to divergent outcomes.

Institutionalists have also been vocal supporters of constitutional engineering and revision of electoral systems as a means of improving representation, accountability, and governance more generally, particularly in divided societies (see Sartori 1997, Reilly 2001). Many advocate for similar reasons for the reorganization of the state along more decentralized or federal lines. 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). The Netherlands, Switzerland, and Belgium offer positive examples, while critics highlight the failure of the consociational model in Lebanon. More recently, consociational recommendations were partially adopted in South Africa’s democratic transition and 1994 constitution, including a grand coalition government and proportional representation (Horowitz 1992).

Cultural factors are also highlighted in explanations for governance variation. Tocqueville’s classic exploration of the role of political culture in explaining democracy in America is one example (see Tocqueville 2003). One of the major works on democratic governance in recent years, Putnam (1993) also highlights the role of political culture in explaining variation in government performance across Italian regions, noting that ‘Tocqueville was right: Democratic government is strengthened, not weakened, when it faces a vigorous civil

society' (Putnam 1993: 182). Putnam takes 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 why, he argues that while socioeconomic factors certainly play a role (explaining, for instance, why northern regions on average performed better than southern regions), the level of social capital – patterns of civic engagement and social solidarity – is far more important in 'making democracy work'. Social capital works largely to facilitate resolution of collective action dilemmas, making it easier to carry out public projects and facilitating stronger public engagement and oversight. Although social capital can be built in the short-term, he argues, it is not easy. He argues that in Italy this variation in social capital has long roots stretching back to early medieval history, in which northern regions had stronger traditions of self-government and horizontal collaboration and relied less on vertical hierarchy than southern regions.

Finally, a significant body of work focuses on the role of individuals, and especially political leaders, in effecting governance outcomes. Although his name has come to be associated with a particular style of (ruthless) leadership, Machiavelli famously highlighted the potentially decisive role of leaders ('princes'). In *Machiavelli's Children*, Samuels (2003), for instance, explores the role of political and business leaders in 19th and 20th century state-building in Japan and Italy. Jackson and Rosberg (1982) map the roles and influences of different leadership styles across Sub-Saharan African countries, highlighting four broad models: prices and oligarchic rule (Senghor in Senegal, Kenyatta in Kenya, Tubman and Tolbert in Liberia, Selassie in Ethiopia, Sobhuza II in Swaziland, and Nimeiri in Sudan), autocrats and lordship (Houphouët-Boigny in Côte d'Ivoire, Ahidjo in Cameroon, Bongo in Gabon, Banda in Malawi, and Sese Seko in Zaire), prophets (Nkrumah in Ghana, Touré in Guinea, and Nyerere in Tanzania), and tyrants and abusive rulers (Macías in Equatorial Guinea and Amin in Uganda).

Major theories of public policy similarly identify the key role of individuals. 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 agenda. 'Policy entrepreneurs' take advantage of policy windows to push their agendas and particular policy solutions.

In summary, the study of how and how well governments govern is central to the study of politics, and the field offers a variety of structural, institutional, cultural, and other arguments to explain both variation in the quality of 'governance' broadly defined and particular aspects of it. The brief discussion here is by no means exhaustive, but intended to provide a broad introduction to major arguments in the literature.

3 Findings from experimental work

In one of the earliest reviews on the use of field experiments to study contemporary governance issues, Humphreys and Weinstein (2009) identify four major questions in which researchers have primarily focused, namely: (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: 370).

In a subsequent review of experimental research on governance, Moehler (2010a) focuses on the related question of whether field experiments can ‘be productively employed to study the impact of development assistance on democracy and governance outcomes?’ (Moehler 2010a: 30). 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 2010a: 42). Her review identifies 41 randomized field experiments of interest in the developing world, including 11 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 although not a systematic review of the literature, following a principal-agent approach to governance. 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. More specifically, they divide the literature into two broader 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

In order to address any potential threat of publication bias, a *systematic* review of published and unpublished papers using rigorous experimental and quasi-experimental methods to study governance was conducted as part of the research for the study of which this paper is a part (Gisselquist, Niño-Zarazúa, and Sajuría forthcoming). It identified 139 relevant papers. According to our classification, which is derived from the basic definition of governance adopted in this paper, we identify three major types of policy interventions that overall cluster around the provision of public goods and aggregating interests, and aim 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 (3) improve informational asymmetries.

Our typology varies from the ones described above in the sense that the first set of factors, focus on the ‘supply’-side dimensions of policies, affecting how public institutions themselves supply goods and social services. Improved governance in that context involves changes both to what is provided (e.g., books and classrooms) and the quality of services (e.g., via a reduction in absenteeism by teachers). The second and third set of factors directly influence the ‘demand’-side for government-provided goods and services, i.e., how the population (usually individuals, households, and occasionally communities) interact with public institutions. Demand-side interventions are found to either provide incentives (often cash), or better information about the provision of goods and services, both with the objective of changing demand-side behaviour.

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

As shown in Table 1, the largest number of papers identified under the governance cluster of ‘aggregating interests’ focus on institutions and corruption, while 27 studies address issues related to participation, including voting behaviour and eight address non-electoral forms of participation. Similarly, under the ‘provision of public goods’ cluster, we find that the largest number of studies (42) focus on health care and education policies, whereas other studies focus on issues related to employment, water and sanitation, and housing. The largest number of studies in our sample was conducted in the USA and India (see Table A1 in the Appendix)

Table 1: Experimental and quasi-experimental studies to study governance

| Governance cluster | Type of policy intervention | Policy area |
|---|--|---|
| <p>Aggregating interests (74)</p> <p>49 out of 74 studies adopted experimental research designs; however, more than half had to resort to quasi-experimental regression techniques such as propensity score matching and instrumental variables to address issues related to endogeneity, spillovers, and sample contamination</p> | <p>Improve supply-side capabilities (29)</p> | <p>Accountability and corruption (2)</p> <p>Democracy (1)</p> <p>Ethnicity (2)</p> <p>Institutional building (18)</p> <p>Non-electoral forms of participation (3)</p> <p>Voting behaviour (3)</p> |
| | <p>Change behaviour via incentives (20)</p> | <p>Accountability and corruption (5)</p> <p>Institutional building (6)</p> <p>Non-electoral forms of participation (4)</p> <p>Voting behaviour (5)</p> |
| | <p>Improve information asymmetries (25)</p> | <p>Accountability and corruption (1)</p> <p>Institutional building (4)</p> <p>Non-electoral forms of participation (1)</p> <p>Voting behaviour (19)</p> |

| Governance cluster | Type of policy intervention | Policy area |
|--|--|--|
| <p>Provision of public goods (65)</p> <p>22 out of 65 studies adopted experimental research designs; however, 11 of those had to resort to quasi-experimental regression techniques such as propensity score matching and instrumental variables to address issues related to endogeneity, spillovers, and sample contamination</p> | <p>Improve supply-side capabilities (18)</p> | <p>Health and/or education (12)</p> |
| | | <p>Water and/or sanitation (3)</p> |
| | | <p>Employment (2)</p> <p>Social protection (1)</p> |
| | <p>Change behaviour via incentives (35)</p> | <p>Health and/or education (23)</p> |
| | | <p>Water and/or sanitation (1)</p> |
| | | <p>Employment (5)</p> |
| | | <p>Social protection (4)</p> |
| | | <p>Rural development (1)</p> <p>Housing (1)</p> |
| | <p>Improve Information asymmetries (12)</p> | <p>Health and/or education (7)</p> |
| <p>Water and/or sanitation (3)</p> | | |
| <p>Employment (1)</p> | | |
| <p>Non-electoral forms of participation (1)</p> | | |

Note: number of reviewed studies in brackets.

Source: compiled by authors.

Several recent studies also review related findings from RCTs with respect to development more generally. Banerjee and Duflo (2012), for instance, draw largely 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. In particular, 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. The ‘seven ideas that work’ that they highlight are: 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: 272-275).

In summary, a number of findings emerge from review of the literature on RCTs that are relevant to explaining variations in the quality of governance. As Banerjee and Duflo's (2012) and Karlan and Appel's (2012) books suggest, 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, sanitation, and education. A number of these studies deal with the impact of projects on 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 programme not only improved students' health in both treatment schools and neighbouring 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 per cent respectively). Kremer, Miguel, and Thornton (2009) evaluate the impact of a merit scholarship programme in Kenya in which girls who scored well on exams had school fees paid and received a grant, finding that the programme had an effect not only on improved student test scores, but also on teacher attendance.

A number of studies explore the impact of public information campaigns on public goods provision. Pandey, Goyal, and Sundararaman (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 to 36 per cent), it also rose among wealthier households not eligible for the subsidy (7 to 26 per cent), 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 Progreso/Oportunidades programme (e.g., De La O 2008; Stecklov et al. 2007). Leroy et al. (2008) for instance, find the programme to be associated with better growth in infants below six months of age (but to have no impact for babies 6-24 months). Other studies explore the impact of conditional cash penalty programmes. One example is Dee's (2011) study of the effects in ten counties of the state of Wisconsin's Learnfare programme, 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 per cent and attendance by 4.5 per cent.

Another set of experimental 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 programme immediately reduced teacher absenteeism in rural India, which was also associated with an improvement in student test scores and achievement one year after the start of the programme. 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 per cent increase in the number of institutional deliveries and increases in the number of preventive care visits by children (56 per cent for those 23 months and younger, and 132 per cent for those 24-59 months), and improvements in prenatal quality as measured by compliance with Rwandan prenatal care clinical practice guidelines. Other studies explore the impact of relatively minor administrative reforms: Banerjee et al. (2012) 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 work also studies issues related to aggregating interests through the study of elections in new and emerging democracies. Wantchekon (2009), 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 centre 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 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 from RCTs highlight a range of strategies, projects, and other inventions that governments could adopt to improve public service provision and representation and accountability in particular areas. Inventions 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 governance

The elegance of RCT findings arguably has a tendency to promote method-driven, rather than theory-driven, research: in other words, it tends to encourage work that asks what questions can be addressed with RCTs, rather than work that begins with questions that are seen as most important to answer and then proposes hypotheses and assesses whether RCTs are an appropriate method for testing them. In Section 3, we summarized some key findings based on RCTs that are relevant to governance. This section considers this summary in light of the review of theories of governance presented above, exploring the extent to which RCTs have – or could – contribute to the testing and building of hypotheses that follow from this long-standing discussion.

One related criticism levelled at experimental work is that it does not address ‘big’ questions and ‘big’ theories (Hyde 2010b). If we compare the factors explored in the RCTs with those identified in the theories of governance reviewed above, there is certainly something to that criticism. One of the major questions in the literature on government, for instance, is about the factors leading to particular regime types, a question completely absent from experimental work. Similarly, major theories of governance highlight factors like social structure that are also largely absent as an object of study in experimental evaluations.

On the other hand, proponents of RCTs make a compelling argument that their avoidance of ‘grand theory’ could be a strength of the literature. Banerjee and Duflo (2012), for instance, advocate a decidedly incremental and ‘micro’ approach. Their solutions posit that governance can be improved with small policy reforms that at the ‘margin’ can lead to desirable improvements in policy, and without major changes to social and political structures. Karlan and Appel (2012) contend that ‘up in the realm of high-minded concepts...the air is thin and there are no poor people to be found...[development] needs to be on the ground’ (Appel 2012: 37). This is in stark contrast with prominent ‘grand’ theories of development that emphasize the role of political institutions in the process of economic development and democratic governance (Acemoglu and Robinson 2012).

However, despite their explicit rejection of ‘grand theory’, it should also be acknowledged that these approaches are not absent of theoretical underpinnings, and in the case of experimental research in particular, analysis often falls clearly within the tradition of behavioural economics, drawing on its theories of individual behaviour, (ir)rational choice, and information.

Compelling as it is in some contexts, this ‘micro’ focus exacerbates one of the key weaknesses highlighted about experimental work: the low external validity of its findings,

despite the fact that the key strength is the high internal validity of its findings. If findings from RCTs are to be used to identify generalizable impacts – i.e., in which specific experiments, conducted in particular situations, can help to 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 such ‘micro’ approaches 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 lend almost no empirically – or theoretically – grounded leverage on the question of whether similar outcomes might be expected in other contexts.

One strategy for improving external validity in experimental research involves precisely speaking to broader theoretical propositions, including drawing on structural theory (Martel Garcia and Wantchekon 2010). But 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. And this constraint inevitably forces us to look beyond experimental methods alone in the study of governance.

Of particular importance is in that context the fact that unlike RCTs that were undertaken for medical research, experimental designs in the field of development and governance pose significant logistical and methodological challenges that often result in the implementation of quasi-experimental regression techniques to tackle the problems of confounding, selection bias, spillovers, and impact heterogeneity that RCTs aim to avoid in the first place (Deaton 2009). Our review of the literature reveals that more than half of the studies that adopted experimental research designs had to resort to quasi-experimental regression techniques such as propensity score matching and instrumental variables techniques to address issues related to endogeneity, spillovers, and sample contamination (see Table 1).

Furthermore, because experiments are relatively rarely replicated across multiple contexts, empirical data that help experimental researchers to address the external validity challenge are limited and often lacking in many developing country contexts. Thus, adopting more of this sort of replication is a strategy that has been recently explored and supported by private and public organizations promoting impact evaluations of development interventions, and which could in principle complement a broader theoretical view. However, as important as experimental design replication may be, researchers are faced with the ultimate challenge of overcoming external validity by the simple fact that parameter heterogeneity is commonly driven by economy-wide and institutional factors that cannot be easily controlled. This often leads to the implicit assumption of constant treatment effects across contexts in systematic reviews and meta-regression analyses.

A third limit to RCTs in the study of governance 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 encourage researchers to focus on low level factors, rather than on factors held by higher level units, such as national institutions (Moehler 2010b). Some traction on such factors can be gained by ‘scaling up’ findings from low level factors. For instance, studies of deliberative democracy at the village level may be used to study deliberative democracy at the national level. However, village versus national politics are so different in other ways that this sort of scaling up clearly provides only suggestive evidence of how deliberative democracy might function.

The limits on 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, for instance, serendipitously gave him a natural experiment to exploit with RCT inspired tools, but most experimental researchers are not so lucky.

In other cases, ethical considerations may impede the study of particular factors. Wantchekon (2003), for instance, explores the impact on voting in Benin of electoral platforms highlighting either public goods provision or 'clientelist' promises to ones' ethno-region. Although he was able to work with political parties to place varied electoral platforms, he was only able to run the experiment in 'safe' districts where it was not expected to matter to election results. Furthermore, it would have been intellectually interesting if his study had more directly tested whether divisive ethnic appeals garnered more or fewer votes, but designing interventions in that way could have exacerbated ethnic tensions in a way that would simply have been unethical.

A fourth issue that limits the utility of RCTs in the study of governance is their relatively short-term window of analysis. Indeed, many theories of governance (and development) focus on 'non-linear' processes that evolve over decades, 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 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.

Fifth, RCTs 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 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 the final issue: the cost of RCTs. 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 governance, it is not necessarily clear whether they would be more cost-effective in testing these theories than regression methods.

5 Conclusion

This paper argues that RCTs have been and can be useful in studying the effects of some policy interventions in the governance area, but that their use in the study of governance 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), where it is expected to have rapid results, but theories of government and the state suggest that what might be most important in explaining variation in governance outcomes are broad, macro-structural shifts, national level variation in institutions and political culture, and leadership.

The focus in this paper has been on the use of RCTs to test hypotheses about why the quality of governance varies. If our focus is narrower on precisely what individual, household, or community factors policy makers might affect that would matter to governance, RCTs appear to be a bit more promising. Policy makers also cannot rapidly change the macrostructure of their country; they govern within the constraints of that structure. Thus, the sorts of interventions studied by RCTs may be precisely of the sort that are most relevant to many policy makers.

However, even adopting this more narrow focus, RCTs have significant weaknesses. For one, policy makers can concentrate their efforts on changing causal factors identified in the literature, such as national institutions and social capital that RCTs cannot. RCTs provide little insight into whether such efforts would be worthwhile. In addition, the inherently weak external validity of RCTs raises major questions about whether policy makers in other contexts and constituencies should expect to see the same results from an RCT-tested intervention. In short, RCTs can tell us ‘what works’ in one specific context, but they can provide little traction on whether the intervention is transferable and ‘could work’ (and why) in other contexts.

References

- Acemoglu, D. and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Business.
- Arndt, Christiane and Charles Oman. 2006. ‘Uses and Abuses of Governance Indicators’. Paris: OECD Development Centre.
- 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.
- 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.
- Collier, Paul and Pedro C. Vicente. 2008. ‘Votes and Violence: Evidence from a Field Experiment in Nigeria’. CSAE WPS/2008-16. Oxford: Univesity of Oxford.

- De La O, Ana. 2008. 'Do Poverty Relief Funds Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico'. *Field Experiments in Comparative Politics and Policy*. Manchester: University of Manchester.
- 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.
- Duflo, Esther and Rema Hanna. 2005. 'Monitoring Works: Getting Teachers to Come to School'. Working Paper Series 11880. Cambridge, MA: NBER.
- Friedman, Milton. 1962. *Capitalism and Freedom*. Chicago: University Of Chicago Press.
- 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, Miguel Niño-Zarazúa, and Javier Sajuria (forthcoming) 'Improving Government Performance in Developing Countries. A Systematic Review. WIDER Working Paper Series.
- 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.
- Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- Horowitz, Donald L. 1992. *A Democratic South Africa? Constitutional Engineering in a Divided Society*. 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.
- Hyde, Susan D. 2010a. 'Experimenting in Democracy Promotion: International Observers and the 2004 Presidential Elections in Indonesia'. *Perspectives on Politics* 8:511-527.
- . 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 Edition. New York: HarperCollins.
- Kirkpatrick, Colin. 2012. 'Economic Governance: Improving the Economic and Regulatory Environment for Supporting Private Sector Activity'. Working Paper 2012/108. Helsinki: UNU-WIDER.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. 'Incentives to Learn'. *Review of Economics and Statistics*, 91: 437-456.
- 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.
- Martel Garcia, Fernando and Leonard Wantchekon. 2010. 'Theory, External Validity, and Experimental Inference: Some Conjectures'. *The ANNALS of the American Academy of Political and Social Science*, 628: 132-147.
- 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. 2010a. 'Democracy, Governance, and Randomized Development Assistance'. *The ANNALS of the American Academy of Political and Social Science*, 628: 30-46.
- . 2010b. '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.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge, MA: Cambridge University Press.
- 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.
- 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. 2007. 'Monitoring Corruption: Evidence from a Field Experiment in Indonesia'. *Journal of Political Economy*, 115: 200-249.
- 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.
- 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
- 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.
- Reilly, Benjamin. 2001. *Democracy in Divided Societies: Electoral Engineering for Conflict Management*. Cambridge: Cambridge University Press.
- 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.
- Samuelson, Paul A. 1954. 'The Pure Theory of Public Expenditure'. *The Review of Economics and Statistics*, 36: 387-389.
- Sartori, Giovanni. 1997. *Comparative Constitutional Engineering: An Inquiry into Structures, Incentives, and Outcomes*. New York: New York University Press.

- 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.
- Wantchekon, Leonard. 2003. 'Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin'. *World Politics*, 55: 399-422.
- . 2009. 'Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin'. New York.
- 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'. World Bank, Washington, DC.

Appendix

Table A1: Experimental and quasi-experimental studies to study governance

| GOVERNANCE CLUSTER | Author(s) | Title | Year | Type of publication | Country | Policy area | Aim of intervention | Experimental design? | Analytical methods 1/ |
|-----------------------|-----------------|---|------|---------------------|---------|-------------------------------|----------------------------------|----------------------|-----------------------|
| 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 behaviour 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 behaviour 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 behaviour 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 behaviour via incentives | Yes | OLS |

| | | | | | | | | |
|------------------------|---|------|-----------------|--------------|------------------------|--|-----|---------------------|
| Olken | Monitoring Corruption: Evidence from a Field Experiment in Indonesia | 2007 | Journal article | Indonesia | Corruption | Change behaviour 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 behaviour via incentives (voters) | Yes | OLS, Probit, and RE |
| Kalyvitis and Vlachaki | When Does More Aid Imply Less Democracy? An Empirical Examination | 2011 | Journal article | | Democracy | Improve supply-side capabilities | No | 2SLS |
| Glennester 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 behaviour via incentives (grants and councils organization) | Yes | FE and PSM |
| 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 behaviour 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 behaviour 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 Serritzlew | 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 |
| 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 | Metropolitan Governance: | 2009 | Journal | USA | Institutional | | Change | Yes | Ordered Logit |

| | | | | | | | | | |
|----------------------|--|------|-----------------|---------------------------|---|--|-----|---------------------------|--|
| Frederickson | Institutional Roles and Inter-jurisdictional Cooperation | | article | | building | behaviour via incentives | | | |
| 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 | |
| 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 behaviour 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 | |

| | | | | | | | | |
|-----------------------------|---|------|-----------------|-----------------------|--------------------------------------|--|-----|---------------------------|
| 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 behaviour via incentives (leadership) | Yes | Pooled OLS |
| Edmark | Migration Effects of Welfare Benefit Reform | 2009 | Journal article | Sweden | Migration | Change behaviour via incentives (benefits) | 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 |
| 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 behaviour 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 behaviour 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 behaviour via incentives | No | OLS, Probit |
| Beaman et al. | Powerful Women: Does Exposure Reduce Bias? | 2009 | Journal article | India | Non-electoral forms of participation | Change behaviour 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 behaviour via incentives | Yes | PSM |
| Hyde | The Observer Effect in International Politics: Evidence from a Natural Experiment | 2007 | Journal article | Armenia | Voting behaviour | Improve supply-side capabilities | Yes | OLS |
| Ichino and Schündeln | Deterring or Displacing Electoral Irregularities? Spill over Effects of Observers in a Randomized Field Experiment in Ghana | 2012 | Journal article | Ghana | Voting behaviour | 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 behaviour | Improve supply-side capabilities | Yes | OLS, FE |
| De La O | Do Poverty Relief Funds Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico | 2008 | Conference | Mexico | Voting behaviour | Change behaviour 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 behaviour | Change behaviour via incentives (female quota) | No | Logit |
| Humphreys and Weinstein | Policing Politicians: Citizen Empowerment and Political | 2007 | Conference (APSA) | Uganda | Voting behaviour | Improve Information | No | Ordered Probit |

| | | | | | | | | |
|------------------------|---|------|-----------------|---------|------------------|---------------------------------|-----|--|
| | Accountability in Africa | | 2007) | | | asymmetries | | |
| James | Performance Measures and Democracy: Improve Information asymmetries Effects on Citizens in Field and Laboratory Experiments | 2011 | Journal article | England | Voting behaviour | Improve Information asymmetries | Yes | Probit |
| Collier and Vicente | Votes and Violence: Evidence from a Field Experiment in Nigeria Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment with Members of Congress and Their Constituents | 2008 | Working paper | Nigeria | Voting behaviour | Improve Information asymmetries | Yes | DiD, FE, Probit |
| Esterling et al. | Do Informed Voters Make Better Choices? Experimental Evidence from Urban India | 2011 | Journal article | USA | Voting behaviour | Improve Information asymmetries | Yes | Bayesian framework using MCMC with data augmentation |
| Banerjee et al. | Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes | 2008 | Journal article | Brazil | Voting behaviour | Improve Information asymmetries | Yes | FE, DiD, OLS, semi-parametric estimators |
| Ferraz and Finin | Improve Information asymmetries Dissemination and Local Governments' Electoral Returns, Evidence from a Field Experiment in Mexico | 2010 | Working paper | Mexico | Voting behaviour | Improve Information asymmetries | Yes | FE, OLS |
| Chong et al. | Can Voters be Primed to Choose Better Legislators? Evidence from Two Field Experiments in Rural India | 2009 | Working paper | India | Voting behaviour | Improve Information asymmetries | Yes | IV, FE |
| Banerjee et al. | Electoral Change Behaviour via Incentives and Partisan Conflict in Congress: Evidence from Survey Experiments | 2011 | Journal article | USA | Voting behaviour | Improve Information asymmetries | Yes | PSM |
| Harbridge and Malhotra | Cues in Context: Analyzing the Heuristics of Referendum Voting with an Internet Survey Experiment | 2008 | Journal article | USA | Voting behaviour | Improve Information asymmetries | Yes | multinomial Logit |
| Borges and Clarke | | | | | | | | |

| | | | | | | | | |
|------------------------|---|------|-----------------|--------------|------------------|---------------------------------|-----|----------------------|
| Vicente and Wantchekon | Clientelism and Vote Buying: Lessons from Field Experiments in African Elections | 2009 | Journal article | West African | Voting behaviour | 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 behaviour | 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 behaviour | Improve Information asymmetries | Yes | OLS and FE |
| Gerber et al. | Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment | 2010 | Journal article | USA | Voting behaviour | 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 behaviour | Improve Information asymmetries | Yes | OLS, FE |
| Guan and Green | Noncoercive Mobilization in State-Controlled Elections | 2006 | Journal article | China | Voting behaviour | Improve Information asymmetries | Yes | OLS |
| Wantchekon | Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin | 2003 | Journal article | Benin | Voting behaviour | Improve Information asymmetries | Yes | Probit |
| Wantchekon | Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin | 2009 | Working paper | Benin | Voting behaviour | Improve Information asymmetries | Yes | Probit, FE |
| Pande | Can Informed Voters Enforce Better Governance? Experiments in Low-Income Democracies | 2011 | Journal article | | Voting behaviour | 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 behaviour 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 behaviour via incentives | No | OLS |
| Baulch | The Medium-Term Impact of the Primary Education Stipend in Rural Bangladesh | 2010 | Discussion Paper | Bangladesh | Education | Change behaviour via incentives | No | PSM, DiD |
| Duflo and Hanna | Monitoring Works: Getting Teachers to Come to School | 2005 | Working paper | India | Education | Change behaviour via incentives | Yes | OLS and 2SLS |
| Kremer et al. | Change Behaviour via Incentives to Learn | 2009 | Journal article | Kenya | Education | Change behaviour via incentives | Yes | OLS |
| Kremer et al. | Decentralization: A Cautionary Tale | 2003 | Working paper | Kenya | Education | Change behaviour via incentives | Yes | OLS, RE |
| Kim and Joo | Did PRWORA's Mandatory School Attendance Policy Increase Attendance Among Targeted Teenage Girls? | 2011 | Journal article | USA | Education | Change behaviour 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 behaviour via incentives | Yes | FE |
| Pande et al. | Community Participation in Public Schools: Impact of Improve Information Asymmetries Campaigns in | 2009 | Journal article | India | Education + Accountability | Improve Information asymmetries | Yes | DiD |

Three Indian states

| | | | | | | | | | |
|--------------------------|---|------|------------------|-----------|----------------------|---|--|-----|----------------|
| Patrinos and Sakellariou | Schooling and Labor Market Impacts of a Natural Policy Experiment | 2005 | Journal article | Venezuela | Education Employment | + | Change behaviour 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 behaviour via incentives | No | PSM |
| Monstad et al. | Education and Fertility: Evidence from a Natural Experiment | 2008 | Journal article | Norway | Education Health | + | Change behaviour 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 |
| Duflo 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 behaviour 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 behaviour via incentives (benefits) | No | DiD |

| | | | | | | | | |
|-------------------------|--|------|------------------|-------------|------------|--|-----|-------------|
| Campolietia and Riddell | Disability Policy and the Labor Market: Evidence from a Natural Experiment in Canada, 1998-2006 | 2011 | Journal article | Canada | Employment | Change behaviour 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 behaviour 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 behaviour 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 | * | 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 |
| 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 | 2011 | Journal | Afghanistan | Health | Change | No | DiD |

| | | | | | | | | | |
|----------------------|--|------|-----------------|-------------|--------|--|--|-----|-----------------|
| | Basic Health Services: A Pilot Study And National Roll-out in Afghanistan | | article | | | | behaviour via incentives | | |
| 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 behaviour 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 behaviour 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 behaviour 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 behaviour via incentives | No | Logit, DiD |
| Arntz and Thomsen | Crowding Out Informal Care? Evidence from a Field Experiment in Germany | 2011 | Journal article | Germany | Health | | Change behaviour 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 behaviour 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 behaviour 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 behaviour 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 Behaviour Via Incentives and Targets in Hospital Care: Evidence From A Natural Experiment | 2010 | Journal article | United Kingdom | Health | | Change behaviour via incentives (policy) | No | DiD |
| 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 behaviour 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 behaviour 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 |

| | | | | | | | | | |
|-------------------------------|--|------|-----------------|----------------|-----------------------------------|-----|--|-----|----------------|
| Poder and He | How can Infrastructures Reduce Child Malnutrition and Health Inequalities? Evidence from Guatemala | 2011 | Working paper | Guatemala | Health nutrition | and | Improve supply-side capabilities | No | PSM |
| Schwartz et al. | The External Effects of Place-based Subsidized Housing | 2006 | Journal article | USA | Housing | | Change behaviour 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 participation | of | Improve Information asymmetries | Yes | IV, Pooled OLS |
| Galasso and Umapathi | Improving Nutritional Status through Behavioral Change: Lessons from Madagascar | 2007 | Working paper | Madagascar | 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 behaviour 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 |
| Gregg et al. | Family Expenditures Post-welfare reform in the UK: Are Low-Income Families Starting To Catch Up? | 2006 | Journal article | United Kingdom | 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 behaviour via incentives | No | DiD, PSM |
| Baird et al. | Cash or Condition? Evidence from a Cash Transfer Experiment | 2010 | Working paper | Malawi | Social Protection | | Change behaviour 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 behaviour 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 behaviour 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 Sanitation | and | Improve supply-side capabilities | No | DiD |
| Hope | Evaluating Social Impacts of Watershed Development in India | 2007 | Journal article | India | Water Sanitation | and | Improve supply-side capabilities | No | PSM |
| Veetil 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 Sanitation | and | Change behaviour 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 Sanitation | and | 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 Sanitation | and | Improve Information asymmetries | Yes | ANOVA |
| Watson et al. | An Opportunistic Field Experiment in Community Water Conservation | 1999 | Journal article | Australia | Water Sanitation | and | Improve Information asymmetries | Yes | MANOVA, OLS |
| Moehler | Democracy, Governance, and Randomized Development Assistance | 2010 | Journal article | World | | | Improve supply-side capabilities | Yes | Review article |

Notes: 1/ Abbreviations stand as follows: Principal Component Analysis (PCA); Ordinary Least Squares (OLS); Two-Stage Least Squares (2SLS), Weighted-least-squares (WLS), Fixed Effects (FE), Random Effects (RE), Propensity Score Matching (PSM), Difference-in-Difference (DiD), Analysis of variance (ANOVA), Multivariate analysis of variance (MANOVA), Random Parameter Logit (RPL), Mixed Logit Model (MLM), Regression Discontinuity (RD), Cox proportional hazard regressions (CPHR), Random Parameter Logit (RPL).

Source: Adapted from Gisselquist, Niño-Zarazúa and Sajuría (forthcoming).