

Using Field Experiments to Study Political Institutions

Guy Grossman*

Laura Paler†

Forthcoming in the Handbook of Comparative Political Institutions
eds. Jennifer Gandhi and Ruben Ruiz-Rufino, Routledge

Political scientists are increasingly using experiments to study the relationship between institutions and political and economic outcomes. Institutions are the “rules and procedures that structure social interaction by constraining and enabling actors’ behavior” (Helmke and Levitsky, 2006, 5). During more than two decades of renewed interest in institutions in political science, researchers have sought answers to broad questions, like: How do institutions affect outcomes such as growth and development, participation, accountability, and policy selection? Which institutions, and what elements of institutional design, matter for these outcomes? How do formal institutions interact with informal institutions? How can weak political institutions be strengthened? And, what are the causes of institutional change?

The interest in using experiments to address such questions reflects an enduring concern with causal inference in the institutions literature (Frye, 2012). Early scholarly work encountered limited success in dealing with the identification issues that arise because the causes and effects of institutions are highly endogenous. This motivated a large empirical literature that used instruments to exploit exogenous variation and isolate the causal effects of institutions (Acemoglu, Johnson and Robinson, 2001). Yet, it is notoriously difficult to find instruments that meet the requirements for unbiased inference (Harrison and List, 2004). It is also far from clear whether instrumental variables approaches identify a quantity or population of theoretical interest (Deaton, 2010). Moreover, many of these studies employed cross-national data and composite indices of institutional quality (such as Polity IV and Freedom House), which induce measurement problems and compromise the ability to identify the effects of any single institution (Pande and Udry, 2006).

Randomized experiments offer one of the most promising approaches available to addressing the causal inference problem in research on institutions. One defining feature of experiments—whether field, lab, or survey—is that the researcher randomly assigns units in the target population to a ‘treatment’ group, which receives a particular intervention, and a ‘control’ group that does not, or that receives a different version of the intervention.¹

*University of Pennsylvania

†University of Pittsburgh. We are grateful to Alex Garlick for his excellent research assistance.

¹Randomized experiments differ from natural experiments in which social and political processes are randomly, or as good as randomly, assigned by something outside of the researchers’ control.

With enough units, random assignment achieves, in expectation, balance across the groups on all pre-treatment covariates, making the control group a suitable counterfactual for the treatment group. Thus, differences in outcomes across the groups can be attributed solely to the effect of the treatment itself—eliminating concerns about reverse causality and spurious causation (Gerber and Green, 2012).

A second defining feature of a randomized experiment is that the researcher has some degree of control over the treatment itself. This makes it possible to study the effects of introducing new institutions or altering specific aspects of institutional design at the micro-level. Precisely because they offer a high-degree of control, lab experiments have proven a valuable means for investigating how institutions affect political behavior and social preferences (Palfrey, 2009; Levitt and List, 2007). Only recently have researchers started to experiment with institutions in field settings, inducing random variation into the nature of democratic institutions, decision-making rules, whether marginalized groups get reserved seats, and whether and how office holders are monitored.

The main goal of this chapter is to assess how field experiments contribute to the study of political institutions in comparative politics, with special emphasis on experiments conducted in developing countries.² In this chapter we primarily consider experiments related to those political institutions that determine the extent to which citizens can control their leaders and the political process.³ We focus on field experiments as an increasingly popular research approach that has exciting potential yet notable limitations. While much has already been written about the strengths and weaknesses of field experiments in general (Harrison and List, 2004; Humphreys and Weinstein, 2009), we begin by discussing how these apply to research on institutions in particular. We then review a number of experiments loosely grouped into three categories: experiments that study the effects of introducing new democratic institutions; that investigate the design of democratic institutions; and that aim to make weak democratic institutions work better.

We show that field experiments are becoming increasingly central to research on institutions, contributing meaningfully to both theory-testing and theory-building. They are particularly promising for knowledge accumulation because they make it possible to vary a number of different features of institutional design in different real-world contexts. Yet, it is also true that many field experiments on institutions to date have taken place at the local level. This raises important concerns about the extent to which they can answer questions about the national institutions—such as democracy, decentralization, the rule-of-law—that

²The focus on developing countries reflects both the authors' expertise and the fact that field experiments are typically conducted in developing countries for both theoretical and practical reasons.

³We do not address several interesting experiments on institutions in the political economy literature for lack of space. These include experiments that alter the incentives of bureaucrats to improve government performance, including in the areas of tax collection (Khan, Kwaja and Olken, N.d.), police performance (Banerjee et al., 2012), and the public sector more generally (Dal Bo, Finan and Rossi, 2013). Other notable experiments have studied the effects of strengthening informal (Blattman, Hartman and Blair, 2014) and formal (Sandefur and Siddiqi, 2013) institutions of dispute resolution.

political scientists often care about. The fact remains that, practically speaking, it is often more challenging using field experiments to study institutions than to study individual political behavior such as turnout and vote choice (Gerber and Green, 2000). Obtaining the necessary statistical power and implementation capacity to randomize institutions often requires overcoming substantial hurdles to obtain resources and build durable relations with donor or governmental partners. We argue that expanding the use of field experiments to study institutions in a way that can address challenges of external validity and scale will likely require developing new models for organizing and conducting research.

Why Use Field Experiments to Study Institutions?

Field experiments address many of the limitations in observational research with their focus on identifying the causal effects of specific institutions, or comparing different institutions or aspects of institutional design. A field experiment involves the implementation of a randomized intervention in a naturalistic way, meaning: in a natural environment, with naturally affected subjects, and where outcomes in the study resemble those in the real world. As such, field experiments strive not only for a high degree of internal validity (unbiased causal inference within the context of the experiment itself) but also ecological validity (the experiment closely approximates a real-world causal process (Morton and Williams, 2010)).

In emphasizing ecological validity, field experiments differ from lab experiments, which have also been used widely to study the causes and effects of institutions. Lab experiments—in which subjects (oftentimes students) are recruited and brought to a central location—are appealing because they give researchers a high degree of control over the treatment and underlying parameters. They have been used to test how the incentives created by institutions affect outcomes like turnout (Levine and Palfrey, 2007), voting (Duffy and Tavits, 2008), and cooperative behavior (Ostrom, Walker and Gardner, 1992). As Morton and Williams (2010) note, lab experiments are particularly useful for testing sharp theoretical predictions from formal theories or the effects of new or rare institutions like innovative voting rules (for example Casella, Gelman and Palfrey (2006)).

Dal Bo, Foster and Putterman (2010) provide one important example of the contribution lab experiments can make to knowledge on institutions. The authors tested in the lab whether the effect of a policy on cooperation varies depending on whether that policy is selected democratically or imposed externally. Subjects participated in several rounds of a prisoners' dilemma game in which they elected whether to enact a fine for defection. In a random set of cases, the researchers overrode the majority vote of subjects and imposed the fine. The authors find that, selecting the fine democratically caused a 40 percent increase in cooperation, demonstrating that the effect of the fine on behavior depended on how that policy was instituted. Moreover, the experimental design enabled the authors to parse out the impact of how the policy was instituted from selection effects (the probability that those

who support a policy will be differentially affected by it).

While lab experiments have generated valuable insights, they are often criticized for not predicting behavior in the ‘real world’ because they lack naturalism in the setting, stakes, and subject pool (Harrison and List, 2004). Lab-in-the-field experiments respond to this critique by selecting subjects that better represent the population of interest and bringing the lab to their natural environment. For example, Grossman and Baldassarri (2012) conducted a lab-in-the-field experiment among villagers in Uganda to shed further light on how electoral institutions mediate cooperation. The authors used a series of public goods games in which subjects were randomly assigned to conditions in which a central authority capable of punishing defection was either elected by subjects or imposed exogenously. The results show that, holding incentives constant, subjects contributed more when they elected their authority. Like Dal Bo, Foster and Putterman (2010), Grossman and Baldassarri (2012) help demonstrating that the political process by which institutions are put in place have an independent effect on behavior. As such, these experiments offer an important insight for the ‘new institutionalist’ literature, which has primarily focused on the effects of the incentives and constraints that institutions generate (Levitsky and Murillo, 2009).

Yet, lab-in-the-field experiments are subject to much of the same criticism that accompanies lab experiments. Lab-in-the-field experiments can also unnaturally influence behavior, for instance by introducing lab-specific ethical considerations or a high level of scrutiny of one’s actions (Levitt and List, 2009). This points to naturalism as the main appeal of field experiments. Yet, field experiments also have a number of limitations that raise questions about the extent to which they can answer important questions about institutions.

One limitation of field vis-a-vis lab experiments is that the researcher cannot generally achieve the same degree of control over the treatment, not to mention underlying parameters. This, in turn, can affect the extent to which the experiment addresses questions of theoretical import as well as the ability to adjudicate between similar experiments conducted in different locations that produce contrasting results. In implementing a field experiment, the researcher typically collaborates with a governmental or nongovernmental partner. While such collaborations are essential and beneficial—and also mean that the experimental results have a ready policy audience—it is not always feasible or desirable to test specific institutional variations.⁴ In a field experiment, the researcher also often has less control over implementation, which can compromise implementation (Chong et al., 2011) or lead to unintended spillover between treatment and control units (Ichino and Schündeln, 2012).

Because randomizing institutions is complex and cumbersome, and because the need for statistical power makes it difficult to randomize institutions at high levels of government, many field experiments on institutions have taken place at the local (e.g. village) level. This raises important external validity questions about the extent to which results generalize to

⁴This would be the case, for instance, if a treatment were considered too politically sensitive or if the partner specifically wanted to test a bundle of activities together, which is often the case in policy evaluations.

the big, high-level institutions (e.g. democracy, decentralization, rule-of-law) that political scientists care about. On one hand, field experiments can play an important role in uncovering the micro-foundations of more macro theories. Moreover, as Banerjee and Duflo (2011) argue, high level institutions (what they call ‘INSTITUTIONS’) consist of myriad lower-level institutions; identifying how to improve institutional design or performance on the margins is thus highly relevant. Critics, on the other hand, argue that more work is needed to explicate and test the conditions under which the results from these studies will hold.

Overall, issues of control and external validity are important to determining the extent to which field experiments are answering important questions in the institutions literature. In the next section we review several field experiments on democratic institutions to evaluate their contribution in light of the strengths and weaknesses described here.

The Growing Experimental Literature on Political Institutions

There is oftentimes a gap between how institutions work on paper and in practice (Helmke and Levitsky, 2006). This is particularly obvious in the functioning of democratic institutions in developing countries, where the quality of governance—according to metrics such as corruption and public service provision—is often low. Bad governance is widely attributed to weak accountability, or citizens’ lack of will or ability to select good leaders or exercise control over bad ones (Besley, 2006). Political failure can arise from a number of causes including ingrained power structures, information asymmetries between citizens and representatives, clientelistic voting, or outright electoral malfeasance. Others have argued that institutions are difficult to change in the short-term and through outside intervention (Easterly, 2006). In light of these obstacles to good governance, researchers are using experiments to grapple with questions like: Do new (externally-imposed) democratic institutions affect governance? How does the design of democratic institutions affect such outcomes? And, how can weak democratic institutions be strengthened? This section reviews recent field experiments that address each of these broad questions.

Introducing New Democratic Institutions

To what extent does introducing new democratic institutions facilitate development, political engagement, and accountability? A growing body of research has aimed to shed light on this question through field experiments conducted in the context of community-driven-development/ reconstruction (CDD/R) programs. While specifics vary across projects, CDD/R interventions typically involve a donor giving block aid grants to villages while requiring recipient communities to establish participatory and transparent institutions that enable all community members to have a voice in the decision over how the funds will be allocated for development. By coupling development aid with democratic institutions, CDD/R programs typically aim to improve economic welfare, governance and social cohesion at the

community level. In the past decade, development organizations have increasingly adopted CDD/R projects as a model for delivering assistance at the local level ⁵

The CDD/R model is based on two key assumptions. The first is that new participatory institutions can increase leader responsiveness to community members and reduce elite capture (Platteau and Abraham, 2002). A corollary assumption is that introducing new institutions is preferable to working through existing institutions if there are entrenched local power structures (Mansuri and Rao, 2013). Secondly, it is assumed that introducing new institutions in one development project will have an impact on village governance more broadly by demonstrating the benefits to community members of more inclusive decision-making. Yet, the political science literature mentioned above challenges these assumptions in arguing that it might be ineffectual to impose institutions externally, particularly in contexts with strong informal institutions or uneven power structures. The widespread adoption of the CDD model thus offers a unique opportunity to contribute to the study of political institutions since the model is predicated on core debates in the institutions literature.

Two recent field experiments involving CDD/R programs in Sierra Leone (Casey, Glennerster and Miguel, 2012) and the Democratic Republic of Congo (DRC) (Humphreys, de la Sierra and van der Windt, 2013) shed light on this debate.⁶ In the GoBifo project in Sierra Leone, researchers worked with the Ministry of Internal Affairs to assign randomly 236 communities to a treatment group that participated in a standard CDD/R program and a control group that received no aid. In the Tuungane program in the DRC, 1250 conflict-affected villages were similarly grouped into 280 treatment and control communities.

Interestingly, neither study finds that the CDD/R intervention caused higher levels of participation, accountability, or social cohesion in treatment communities. In Sierra Leone, the authors observe that the CDD/R program improved welfare by some metrics but had no effect on governance and social cohesion. For instance, women were no more likely to attend or voice an opinion at community meetings, despite their involvement in various CDD activities. Similarly, in the DRC, both survey data and behavioral measures reveal little effect of the Tuungane program. Overall, both of these experiments were very well done but neither finds evidence that the program reshaped village institutions, empowered marginalized populations, or improved collective action capacity.

These results not only have important implications for the design of future development programs but also contribute to the academic literature on institutions. Both studies suggest that exogenously introducing new participatory institutions has little effect on local governance and social cohesion in the short-run. This lends support to the idea that local governance institutions are resilient to attempts at alteration by external actors. Interestingly, however, the evidence also suggests that these results might reflect the fact that levels of general participation, public information, and equity in decision-making within communities

⁵The World Bank alone currently spends about \$2 billion per year on various CDD/R programs (Mansuri and Rao, 2013).

⁶A similar CDD/R experiment has been conducted in Liberia by Fearon, Humphreys and Weinstein (2011).

were relatively high to begin with.

Future CDD/R field experiments could advance the institutions literature through a number of promising research directions. These include studying the effects on development of institutions at different levels of government, which could be addressed by varying the level at which the CDD/R intervention is implemented (e.g. village or district). Another option is to investigate the effects of paying taxes (or making financial contributions) on engagement and accountability by varying whether recipient communities are required to make financial or in-kind contributions in order to receive aid. Alternatively, future research could isolate the effect of introducing new institutions by implementing the aid and institutional components of the treatment separately, as Blimpo and Evans (2011) did in a study on education administration in Gambia.

The Design of Democratic Institutions

How do differences in the design of democratic institutions affect policy outcomes as well as satisfaction with the overall political process? This section reviews experiments that help to answer this question by testing how direct versus representative democracy affect the allocation of resources and the impact of reserving seats for marginalized groups on various attitudinal outcomes.

Direct versus representative democracy

The political science literature has long been concerned with the effects of direct versus representative democracy on policy outcomes and satisfaction with the political process. As Olken (2010, 243) summarizes, there are two possible advantages of direct democracy. First, it may allow voters to circumvent unresponsive institutions. Second, inclusionary and participatory decision-making rules enhance the legitimacy of political decisions, regardless of the decisions themselves (Grossman and Baldassarri, 2012). In observational settings, the rules that govern policy or project selection are chosen endogenously, complicating the testing of these predictions. Field experiments that randomly assign decision rules are thus uniquely positioned to evaluate the impact of such political institutions.

Recent field experiments in Afghanistan and Indonesia were designed explicitly to test the above propositions, again in the context of CDD/R interventions. Beath, Christia and Enikolopov (2013a) conducted a field experiment in 250 Afghan villages participating in the National Solidarity Programme. Half of the study villages were randomly assigned to select development projects at village meetings convened by councils, mirroring traditional decision-making. The remaining villages were randomly assigned to select projects by secret-ballot referenda, giving villagers the opportunity to vote directly for their preferred project. The authors used surveys to elicit *ex-ante* preferences to isolate the effects of direct democracy on allocation and, specifically, the ability of elites to realize their preferences. The authors find

that elites' preferences only matter when selection occurs through the village council. This indicates that direct democracy limited elite influence over resource allocation.

Similarly, Olken (2010) conducted a field experiment in 49 Indonesian villages to study how different political institutions govern the selection of local public goods in the context of a national CDD program. Each village followed a political process that resulted in two infrastructure proposals, one "general project" proposed by the village at large and one "women's project" proposed exclusively by women in the village. As in Beath, Christia and Enikolopov (2013*a*), half of the study villages were randomly assigned to select projects through local assemblies while the remaining villages selected projects via plebiscites. Olken (2010) finds relatively little impact of the plebiscite treatment on the general project. Yet it had a notable impact on the women's project, resulting in the location of projects to poorer areas of the village.

Broadening representation for marginalized groups

Another recent set of experiments focuses on the effects of ensuring that marginalized groups, such as women and minorities, are represented in decision-making bodies by introducing institutional reform in the form of quotas.

Importantly, legal restrictions are usually not the cause of wide gender disparities in political representation. Instead, in both rich and poor countries, women's access to public office at local and national levels is restricted by voter bias in favor of male politicians (Inglehart and Norris, 2003). On one hand, scholars argue that the bias against electing women for public office is difficult to change in the short term (Inglehart and Norris, 2003) and that mandated quotas are likely to spark a backlash if they are seen as violating voters' sense of identity and restricting their vote choice (Rudman and Fairchild, 2004). On the other hand, supporters argue that quotas can improve the perceived effectiveness of leaders of marginalized groups through demonstration effects. For instance, electing women to power can weaken stereotypes about gender roles in the public and domestic spheres and eliminated the negative bias in the perceived effectiveness of female leaders among men and women alike.

The field experiment in Afghanistan conducted by Beath, Christia and Enikolopov (2013*b*) also sought to identify the causal effects of reforming the rules of representation by introducing gender quotas at the community level. In half of 500 villages, the research team introduced institutional rules that promote gender equality by establishing a gender-balanced village development council; requiring equal participation of men and women in elections and project selection; and mandating that at least one selected project was prioritized by women. Using survey data from 13,000 respondents, the study finds that these new rules had little effect on intra-family decision-making. Gender quotas, however, increased female participation in village governance and community life while also increasing support for female participation in village decision-making. These results mirror evidence from a recent policy experiment that exploits the random assignment of gender quotas across Indian village councils. Here

Beaman et al. (2012) find that the likelihood that a woman speaks in a village meeting increases by 25% when the local political leader position has been reserved for a woman. Also examining the Indian case, Bhavani (2009) finds that even after quotas have been removed female candidates are more likely to run and win elected local political positions.

All in all, experimental research on gender quotas suggests that, though social norms are resilient, institutional reforms can induce meaningful attitudinal and behavioral change (see also Paluck and Shepherd (2012)). Specifically, formal (quota) institutions that increase women's representation are able to get women *constituents* to participate more meaningfully in local governance. This is a particularly interesting in light of the null results reported by CDD/R experiments testing the effect of introducing new democratic institutions at the village level. Existing evidence thus suggests that giving women a formal designated role in village decision-making as in Beath, Christia and Enikolopov (2013*b*) produces a stronger effect than encouraging broader participation more generally as in Casey, Glennerster and Miguel (2012) and Fearon, Humphreys and Weinstein (2011). We suspect that this may be because gender quotas, which are designed to affect prior perceptions of women as leaders, are less threatening to local elites than democratic reforms that more directly challenge deep-rooted informal power structures.⁷ Finally, gender quotas experiments also contribute to the literature on informal institutions by showing that, at least in the case of informal institutions of female discrimination in leadership positions, significant change can occur in short order and does not have to be slow or incremental (Lauth, 2000; Knight, 1992).

Making Weak Democratic Institutions Work Better

Whereas the studies reviewed above assess the impact of introducing new democratic institutions or altering institutional design, a number of field experiments focused on information and monitoring have examined ways to make weak democratic institutions work better without directly altering the institutions themselves.

Information campaigns

There is good reason to believe that electoral institutions fail in part because voters lack the information they need to hold government accountable. A number of recent experiments have therefore investigated whether providing citizens with more information will have a significant effect on turnout and vote choice. One broad set of experiments explores the effect of civic education on political behavior. These experiments derive from the core insight in Almond and Verba (1963) that democratic institutions work better when the political culture is imbued with knowledge of the rights and duties of both citizens and government. Yet, such a democratic culture is rarely present in new democracies where citizens are unfamiliar with

⁷Existing evidence suggests that across the developing world, women who win seats via quotas are usually related to the male local leaders.

challenging authority, tend to vote along ethnic or religious lines, or are susceptible to vote buying and clientelistic exchange.

One example of such an experiment is a voter mobilization campaign aimed at informing women in Pakistan of the benefits of voting and of the secret ballot (Gine and Mansuri, 2012). Geographic clusters within villages were assigned to treatment and control groups; a further subset of households within treatment villages were assigned to receive the door-to-door campaign. The authors find that treated women, compared to controls, were 12 percent more likely to vote in the 2008 national elections. They also find similar turnout rates for *untreated* women within treatment clusters, demonstrating the potentially powerful spillover effects from civic education campaigns. Importantly, the authors show that the increase in female turnout caused a decrease in support for the winning party, which suggests that information could have a significant effect on the policy agenda.

A number of other studies have demonstrated that civic education campaigns can also have an effect on undesirable political behavior. For instance, Banerjee et al. (2013) show that messages encouraging citizens not to vote along ethnic lines caused a 10 percent reduction in ethnic voting and reduced support for criminal politicians in 2007 elections in India. Collier and Vicente (2014) demonstrate that an anti-violence pre-elections campaign in Nigeria reduced the intensity of violence and increased voter turnout. In another study, Vicente (2014) finds that an education campaign criticizing vote-buying in Sao Tome and Principe reduced the influence of money in elections. Interestingly, the campaign also decreased electoral participation and increased the vote-share of the incumbent. This experiment thus provides the insight that vote buying can increase turnout and that challengers might be more likely to use vote buying as a strategy for counteracting incumbency advantage. Overall, these experiments provide growing evidence that giving voters more information that aims to instill a ‘civic culture’ can indeed have a positive effect on political behavior.

Another set of experiments arises from the notion that citizens lack information on politician performance and therefore are ill-positioned to hold incumbents accountable by rewarding or sanctioning them at re-election time. Several recent experiments aimed to provide voters with better information on incumbent performance to assess the extent to which this affects voter behavior and, ultimately, accountability.

In an experiment implemented prior to state legislature elections in Delhi, India, Banerjee et al. (2011) examine whether the exogenous provision of political information through the media in 200 treatment villages influences voter turnout, incumbent vote share, and the use of vote-buying as an electoral strategy. Treatment villages received newspapers containing information on the performance of the incumbent legislator along a number of dimensions, on legislator responsibilities, and on the qualifications of major candidates. The authors find that the information caused higher turnout, reduced vote-buying, and higher vote share for better performing and more qualified incumbents. Chong et al. (2011) conducted an experiment in Mexico in which treatment households received different information on corruption

and public spending in the lead up to 2009 mayoral elections in Mexico. They find, in contrast to Banerjee et al. (2011), that information on corruption *decreased* voter turnout and increased apathy. Similarly de Figueiredo, Hidalgo and Kasahara (2013) find that, under certain conditions, corruption information can suppress voter turnout in a field experiment in Brazil.

Two recent information field experiments directly involve politicians rather than simply providing unmediated information to voters. To study the effects of information on the performance of Members of Parliament (MPs) in Uganda, Humphreys and Weinstein (2012) informed a random sample of MPs that their constituents would be given detailed information on their performance in the lead up to 2011 elections. While survey data shows that Ugandan voters wanted information about their MP's performance, the dissemination campaigns had no impact on reelection probabilities. Finally, Bidwell, Casey and Glennerster (2013) focus on an intervention that provides voters with information on campaign platforms through videos of candidate debates and find moderate to strong evidence for shifts in voting behavior.

All in all, the information experiments help to demonstrate that providing voters with better information affects political behavior and accountability. They provide some support for the notion that institutions—in this case, electoral institutions—can be improved without making changes to the rules by which those institutions operate. Yet, they also suggest that information is not always welfare enhancing and, under certain conditions, it may have adverse effect on turnout. Importantly, the mixed results are motivating a second generation of field studies that aim to explain how voting behavior varies depending on voters' priors (Paler, 2013); perceived importance of candidate identity (Conroy-Krutz, 2013); or expectations about what government should be doing in the first place (Gottlieb, 2012).

Independent Monitoring

Electoral institutions are unlikely to induce political accountability if informal power structures shape how those institutions operate. Recent research suggests that introducing independent monitoring institutions can compel democratic institutions to work better (Olken, 2007). Monitoring institutions constrain the behavior of politicians insofar as they come with a credible punishment for wrongdoing. A number of recent experiments have aimed to test the hypothesis that election-monitoring institutions will undermine politicians' ability to engage in electoral malfeasance. This is an important hypothesis given the concern that monitoring might simply cause candidates or parties to shift illegal activities to unmonitored locations or administration levels.

To test these hypotheses, Ichino and Schündeln (2012) implemented a field experiment examining the causal effect of *domestic* election observation on irregularities in voter registration in the lead up to 2007 elections in Ghana. Using a two-level design, the authors find evidence that observers displace a substantial portion of irregularities to nearby unobserved registration efforts following a pattern of communication among political party agents.

This sort of coordination among agents is also the focus of the recent paper by Callen and Long (forthcoming) who evaluate an innovative election monitoring technology that was introduced during the 2010 parliamentary elections in Afghanistan. The technology works by photographically recording differences between post-election polling center level counts and the corresponding numbers in the certified national aggregate. The authors find that such monitoring reduced both the incidence of theft or damage to election materials at polling centers and the number of votes cast for powerful candidates. The authors also find evidence of elite coordination in that election officials are more willing to alter results for candidates that could shield them from possible retribution. All in all, these studies contribute to knowledge of how monitoring affects institutional performance and sheds light on the strategic response of politicians to efforts to detect and deter misconduct.

Conclusions and Future Directions

This chapter has assessed the extent to which field experiments are contributing to the study of political institutions, particularly democratic institutions. This review has aimed to show that recent experiments designed to investigate the causal effects of introducing new democratic institutions, varying institutional design, and strengthening weak institutions are contributing both empirically and theoretically to the institutions literature.

With respect to research on the effects of introducing new democratic institutions, the recent set of experiments conducted in the context of CDD/R programs does not find evidence that introducing new democratic institutions at the local level results in significant improvements to local governance or collective action capacity. In contrast, experiments on different aspects of institutional design are showing that direct (versus representative) democracy and mandated gender quotas can ameliorate ingrained norms and have meaningful effects on policy outcomes. Another set of experiments is providing robust evidence that information and monitoring have an effect on citizen political behavior and the ways in which weak democratic institutions work.

While these studies are shedding light on long-standing debates in the institutions literature that cannot easily be resolved using observational data, it is important to take stock of the experiments literature. The number of experiments on institutions to date is relatively small, making it inappropriate to draw definitive conclusions from them. For instance, the fact that recent CDD/R experiments provide little evidence that introducing new institutions has an affect on governance and cohesion cannot be taken as conclusive proof that they will never have an effect. Instead, these ‘first-generation’ studies are primarily highlighting the need for more empirical and theoretical work on the conditions under which introducing new institutions might be effective. This speaks to the potential for future research to address concerns of external validity, namely by intentionally designing experiments to investigate how results vary in different contexts or if aspects of the treatment are altered. The idea is to

replicate experiments in different locations, but only after explicitly theorizing how specific contextual variables are likely to affect the relationship between the manipulated institution and the outcome of interest. We believe that better theorizing and testing of the relationship between distinct experimental manipulations and the disparate social and political contexts into which they are introduced is necessary for extracting generalizable knowledge that researchers and policy makers can put to use in varied settings.

There is a growing need for a large number of experiments (in various contexts using different permutations of treatment design) on a single topic if we are to use experimental methods to achieve meaningful knowledge accumulation. Such a research agenda goes beyond the important call for more replication studies to address external validity concerns (Humphreys and Weinstein, 2009). Instead, achieving knowledge accumulation likely entails rethinking the way that experimental research is organized in the first place. In addition to the current decentralized, entrepreneurial approach to research in the social sciences, a new model that fosters greater coordination among experimental research teams around topics of theoretical importance is needed. Such coordination would explicitly use site selection to increase generalizability as well as strive to achieve standardization around such things as research design, outcomes, sample selection, measurement, replication materials and implementation protocols. Such standardization is necessarily for conducting meta-analysis, the workhorse of knowledge accumulation.⁸

Future experimental work on political institutions must also address the thorny issue of level of analysis. As this chapter demonstrates, the vast majority of field experiments in developing countries take place at the community (village or town) level. For some research questions (for example, election monitoring), the community or neighborhood is the natural unit of analysis. In such cases field experiments at the local level can help uncover general or universal micro-foundational behavior. This seems to be the case, for example, with experiments testing the impact of mandated gender quotas where discrimination against women in leadership positions is more a matter of prejudice than about voter coordination.

However, for other outcomes—for example, those testing the resilience of informal institutions in the face of new formal institutions—it may be difficult to draw conclusions at higher levels from studies that take place at the village level. On one hand, village leaders perform several roles in a community (e.g., access point to higher levels of government and providers of services such as loans), which may reduce the willingness of community members to challenge their authority. On the other hand, the dense networks characterizing villages allow using tools, such as naming and shaming, to constrain elite behavior that are not available to constituents in large geographical units. The experimental literature needs to explicate more clearly the conditions under which findings at the local level might be generalizable to other levels of government and other contexts. Moreover, a greater awareness of the limitations of

⁸See, for example, the new granting initiative by the Experiments in Governance and Politics (www.e-gap.org), which prioritizes a high degree of coordination across funded research teams.

research on institutions at the local level is motivating researchers to conduct field experiments on political institutions at higher levels (Humphreys and Weinstein, 2012; Grossman, Humphreys and Sacramone-Lutz, forthcoming; Banerjee et al., 2012). This is an important development if field experiments are to continue making meaningful contributions the study of the causes and effects of political institutions.

References

- Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review* 91(5):1369–1401.
- Almond, G. and S Verba. 1963. *The Civic Culture*. Princeton: Princeton University Press.
- Banerjee, Abhijit, Donald P. Green, Jennifer Green and Rohini Pande. 2013. "Can Voters be Primed to Choose Better Legislators? Experimental Evidence from Rural India." *Mimeo* .
- Banerjee, Abhijit, 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." *NBER Working Paper* 17912.
- Banerjee, Abhijit V and Esther Duflo. 2011. *Poor economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs.
- Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su. 2011. "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India." *Mimeo* .
- Beaman, Lori, Esther Duflo, Rohini Pande and Petia Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science* 335(6068):582–586.
- Beath, Andrew, Fotini Christia and Ruben Enikolopov. 2013a. "Direct Democracy and Resource Allocation: Experimental Evidence from Afghanistan." *MIT Political Science Department Research Paper* No. 2011-6.
- Beath, Andrew, Fotini Christia and Ruben Enikolopov. 2013b. "Empowering Women: Evidence from a Field Experiment in Afghanistan." *American Political Science Review* 107(3):540–557.
- Besley, Timothy. 2006. *Principled Agents? The Political Economy of Good Government*. New York, NY: Oxford University Press.
- Bhavani, Rikhil R. 2009. "Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India." *The American Political Science Review* 103(1):23–35.
- Bidwell, Kelly, Katherine Casey and Rachel Glennerster. 2013. "The Impact of Voter Knowledge Initiatives in Sierra Leone." *Innovation for Poverty Action and J-PAL* .
- Blattman, Christopher, Alexandra Hartman and Robert Blair. 2014. "How to Promote Order and Property Rights under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior through Community Education." *American Political Science Review* 108(1):100–120.
- Blimpo, Moussa and David Evans. 2011. "School-Based Management and Educational Outcomes: Lessons from a Randomized Experiment." *Working Paper* .
- Callen, Michael and James D. Long. forthcoming. "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review* .

- Casella, Alessandra, Andrew Gelman and Thomas R. Palfrey. 2006. "An Experimental Study of Storable Votes." *Games and Economic Behavior* 57(1):123–154.
- Casey, Katherine, Rachel Glennerster and Edward Miguel. 2012. "Reshaping Institutions: Evidence on External Aid and Local Collective Action." *Quarterly Journal of Economics* 127(4):1755–1812.
- Chong, Alberto, Ana L. De La O, Dean Karlan and Leonard Wantchekon. 2011. "Looking Beyond the Incumbent: The Effects of Exposing Corruption on Electoral Outcomes." *NBER Working paper* no. 17679.
- Collier, Paul and Pedro C. Vicente. 2014. "Votes and Violence: Evidence from a Field Experiment in Nigeria." *the Economic Journal* 124(574):F327–F355.
- Conroy-Krutz, Jeffrey. 2013. "Information and Ethnic Politics in Africa." *British Journal of Political Science* 43(2):345–373.
- Dal Bo, Ernesto, Frederico Finan and Martin Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128(3):1169–1218.
- Dal Bo, Pedro, Andrew Foster and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review* 100(5):2205–2229.
- de Figueiredo, Miguel F. P., Daniel F. Hidalgo and Yuri Kasahara. 2013. "When Do Voters Punish Corrupt Politicians? Experimental Evidence from Brazil." *Mimeo* .
- Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48(2):424–55.
- Duffy, John and Margit Tavits. 2008. "Beliefs and Voting Decisions: A Test of the Pivotal Voter Model." *American Journal of Political Science* 52(3):603–618.
- Easterly, William. 2006. *The white man's burden: why the West's efforts to aid the rest have done so much ill and so little good*. New York: Penguin Press.
- Fearon, James D., Macartan Humphreys and Jeremy M. Weinstein. 2011. Democratic Institutions and Collective Action Capacity: Results from a Field Experiment in Post-Conflict Liberia. In *Annual Meetings of the American Political Science Association, Seattle, September 1-4*.
- Frye, Timothy. 2012. "In From the Cold: Institutions and Causal Inference in Postcommunist Studies." *Annual Review of Political Science* 15:245–263.
- Gerber, Alan S. and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–663.
- Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. Norton & Company.
- Gine, Xavier and Ghazala Mansuri. 2012. "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan." *Mimeo* .

- Gottlieb, Jessica. 2012. "Does raising voter expectations improve accountability? A field experiment in Mali." *Mimeo* .
- Grossman, Guy and Delia Baldassarri. 2012. "The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda." *American Journal of Political Science* 56(4):964–985.
- Grossman, Guy, Macartan Humphreys and Gabriella Sacramone-Lutz. forthcoming. "Does Information Technology Flatten Interest Articulation? Evidence from Uganda." *American Political Science Review* .
- Harrison, Glenn W. and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42(4):1009–1055.
- Helmke, Gretchen and Steven Levitsky. 2006. Introduction. In *Informal Institutions and Democracy: Lessons from Latin America*, ed. Gretchen Helmke and Steven Levitsky. Baltimore, MD: The John Hopkins University Press.
- Humphreys, Macartan and Jeremy M. Weinstein. 2009. "Field Experiments and the Political Economy of Development." *Annual Review of Political Science* 12:367–378.
- Humphreys, Macartan and Jeremy M. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." *Mimeo* .
- Humphreys, Macartan, Raul Sanchez de la Sierra and Peter van der Windt. 2013. "Exporting Institutions: A Randomized Experiment in Eastern Congo." *Working Paper* .
- 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(01):292–307.
- Inglehart, Ronald and Pippa Norris. 2003. *Rising tide: gender equality and cultural change around the world*. Cambridge, UK: Cambridge University Press.
- Khan, Adnan, Asim Kwaja and Ben Olken. N.d. "Property Tax Experiment in Punjab Pakistan: Testing the Role of Wages, Incentives and Audit on Tax Inspectors' Behavior." **URL:** <http://www.povertyactionlab.org/evaluation/property-tax-experiment-punjab-pakistan-testing-role-wages-incentives-and-audit-tax-inspe>
- Knight, Jack. 1992. *Institutions and social conflict*. New York, NY: Cambridge University Press.
- Lauth, Hans-Joachim. 2000. "Informal Institutions and Democracy." *Democratization* 7(4):21–50.
- Levine, David K. and Thomas R. Palfrey. 2007. "The Paradox of Voter Participation: A Laboratory Study." *American Political Science Review* 101(1):143–58.
- Levitsky, Steven and M. Victoria Murillo. 2009. "Variation in Institutional Strength: Causes and Implications." *Annual Review of Political Science* 12:115–133.
- Levitt, Steven D. and John A. List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?" *The Journal of Economic Perspectives* 21(2):153–174.

- Levitt, Steven D. and John A. List. 2009. "Field experiments in economics: The past, the present, and the future." *European Economic Review* 53(1):1–18.
- Mansuri, Ghazala and Vijayendra Rao. 2013. *Localizing development: does participation work?* Washington, D.C.: World Bank.
- Morton, Rebecca B. and Kenneth C Williams. 2010. *Experimental political science and the study of causality: from nature to the lab*. Cambridge, MA: Cambridge University Press.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2):200–249.
- Olken, Benjamin A. 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review* 104(02):243–267.
- Ostrom, Elinor, James Walker and Roy Gardner. 1992. "Covenants With and Without a Sword: Self-Governance is Possible." *American Political Science Review* 86(2):404–17.
- Paler, Laura. 2013. "Keeping the Public Purse: An Experiment in Windfalls, Taxes, and the Incentives to Restrain Government." *American Political Science Review* 104(7):706–725.
- Palfrey, Thomas. 2009. "Laboratory Experiments in Political Economy." *Annual Review of Political Science* 12(379-388).
- Paluck, Elizabeth Levy and Hana Shepherd. 2012. "The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network." *Journal of Personality and Social Psychology* 103(899–915).
- Pande, Rohini and Christopher Udry. 2006. Institutions and Development: A View from Below. In *Advances in Economics and Econometrics: Theory and Applications*, ed. R. Blundell, Newey W. and T. Persson. Vol. 2 Cambridge University Press pp. 349–412.
- Platteau, Jean-Philippe. and Anita Abraham. 2002. "Participatory Development in the Presence of Endogenous Community Imperfections." *Journal of Development Studies* 39(2):104–136.
- Rudman, Laurie A. and Kimberly Fairchild. 2004. "Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance." *Journal of Personality and Social Psychology* 87(2):157–176.
- Sandefur, Justin and Bilal Siddiqi. 2013. "Delivering justice to the poor: Theory and experimental evidence from Liberia." *Mimeo* .
- Vicente, Pedro C. 2014. "Is Vote-Buying Effective? Evidence from a Field Experiment in West Africa." *The Economic Journal* 124(574):F356—F387.