

**EXPERIMENTAL RESEARCH ON
DEMOCRACY AND DEVELOPMENT**

Ana L. De La O¹

Leonard Wantchekon²

Contribution to Cambridge Handbook of Experimental Political Science,
Cambridge University Press, Druckman J.N. et al. (eds)

Words: 7,685

¹ Assistant Professor of Political Science, Yale University, ana.delao@yale.edu

² Professor of Politics, Princeton University, lwantche@princeton.edu

Expectations about the role of democracy in development have changed considerably in recent years. In principle, the exercise of political rights sets democracies apart from other political regimes in that voters can pressure their representatives to respond to their needs. It has been argued that such pressure “helps voters constrain the confiscatory temptations of rulers and thereby secure property rights; increases political accountability, thus reduces corruption and waste; and improves the provision of public goods essential to development” (Boix and Stokes 2003, 538). Thus, the argument follows, democracy is development-enhancing. Yet deprivations, such as malnutrition, illiteracy, and inequalities in ethnic and gender relationships have proven to be resilient, even within the nearly two-thirds of the world's countries ranked as electoral democracies. The persistence of deprivations is a reminder that there is still a great deal to be learned about the relationship between democracy and development.

Not surprisingly, scholars have explored numerous ways in which democracy can be related to development, ranging from macropolitical examinations (e.g. are democracies better at producing development than are authoritarian regimes?), to microexplanations (under what circumstances can voters limit bureaucrats' rent-seeking behavior?). Yet the bulk of empirical evidence in this respect is inconclusive (Przeworski and Limongi 1997; Boix and Stokes 2003; Keefer 2007). Is democracy a requirement for development or is it the other way around? Are formal institutions the causes or the symptoms of different levels of development? Which should come first, property rights or political competition? Civil liberties or public service provision? Why are elections compatible with rampant corruption? As critical as these questions are to the discipline,

what we know thus far is plagued by problems of simultaneous causality, spurious correlations, and unobserved selection patterns.

Recently, experimental research on the political economy of development has blossomed. Despite its novelty, progress has been rapid and continues apace. As experiments in this field have evolved, several features distinguish them from earlier empirical contributions. First, scholars have started to address central debates in the field by mapping broad theoretical issues to more specific and tractable questions (Humphreys and Weinstein 2009). For example, instead of asking how different political regimes shape development, recent studies ask whether various methods of preference aggregation produce divergent provisions of public goods. Second, unlike previous macrostudies based on cross-country regressions, recent work has focused on the subnational level. Third, researchers are increasingly making use of field experiments to study how politics affects development and how development shapes politics in developing countries.

Throughout this chapter, as in the rest of this volume, when we speak of experiments we mean research projects where the subjects under study are randomly assigned to different values of potentially causal variables (i.e., different treatment and control groups). For example, a researcher might assign one group of households to receive cash transfers and assign another group of households to receive the same cash transfers but make the latter transfer conditional on parents investing in their children's education. In some, but not all, designs there is also a control group that does not receive any treatment. As Druckman et al. explain in the introduction to this volume, random

assignment means that each entity being studied has an equal chance to be in a particular treatment or control condition.

Experimentation in the field of political economy of development has taken several forms: the increasingly popular field experiments take place in a naturally occurring setting; laboratory experiments take place in a setting controlled by the researcher; laboratory experiments in the field resemble field experiments more generally, in that interventions take place in a naturally occurring setting, but researchers have more control over the setting and the treatment; survey experiments involve an intervention in the course of an opinion survey; and finally, there are some instances when interventions of theoretical interest have been randomly assigned not by researchers but by governments. We group studies that take advantage of this type of randomization in the category of natural experiments.

Because experimentation is still a novel research tool in the field, throughout this chapter we review some of the ongoing and published research projects that illustrate how random assignment is being used to tackle questions about the political economy of development. We begin Section 1 by considering examples of pioneering field experiments executed in collaboration with nongovernmental organizations (NGOs). Section 2 describes two unique field experiments done in partnership with political parties. Section 3 presents several studies that took advantage of natural experiments. Section 4 introduces the use of laboratory and laboratory in the field experiments. Section 5 discusses some of the challenges faced by researchers conducting experiments on development and democracy, such as internal and external validity, as well as ethical

issues. This section also presents practical solutions to some of these challenges drawing from recent experimental designs.

In section 6, we conclude that, despite the challenges, experiments are a promising research tool that have the potential to make substantial contributions to the study of democracy and development, not only by disentangling the causal order of different components of democracy and development, but also by providing evidence that other empirical strategies cannot produce. Moving forward, we argue that the best of the experimental work in the field of democracy and development should reflect well-chosen populations, a deep understanding of the interaction of the interventions with their contexts, and should test theoretical mechanisms such that scientific knowledge starts to accumulate.

1. Field experiments in collaboration with NGOs

Olken's (2010) study of two political mechanisms –plebiscites and meetings-- in Indonesia illustrates the use of field experiments to test a particular angle of the relationship between democracy and development. While most of the previous work on the topic takes institutions as a given and studies their effects (Shepsle 2006), Olken's study starts from the recognition that, in countless examples, institutions and the public policies that follow them are endogenous.

Olken, with support from the World Bank and UK's Department for International Development (DfID), conducted a field experiment in forty-eight Indonesian villages, each of which was preparing to petition for infrastructure projects as part of the Indonesian Kecamatan Development Program. All villages in the experiment followed

the same agenda-setting process to propose two infrastructure projects -- one general project determined by the village as a whole, and one women's project. The experiment randomly assigned villages to make the final decision regarding the projects either through a meeting or through a plebiscite. Olken examined the impact of meetings and plebiscites on elite capture along two main dimensions. First, he examined whether the types of projects chosen moved closer to the preferences of village elites. Second, he tested whether the location of projects moved toward wealthier parts of the villages.

The experiment's findings paint a mixed picture. Whether there was a meeting or a plebiscite had little impact on the general project, however the plebiscite did change the location of the women's project to the poorer areas of a village. The type of project chosen by women, however, was closer to the stated preferences of the village elites than to poor villagers' preferences. Olken explains that because the experiment left the agenda-setting process unchanged, the elite's influence over the decision-making process regarding the type of project remained unchallenged. The experiment thus confirms previous arguments on the relevance of political mechanisms to aggregate preferences. At the same time, it shows the resilience of political inequalities.

The persuasiveness of the results comes from the research design, which guaranteed that plebiscites and meetings were allocated to villages regardless of their social and political configuration or any other observed or unobserved characteristic. Therefore, differences in the type and location of projects can be adjudicated with certainty to the political mechanism in place.

Olken's experiment is an example of a growing trend in political science and development economics where researchers collaborate with NGOs in order to implement

an intervention and evaluate its effects. This type of partnership has proven fruitful for the study of a vast array of topics central to our understanding of the relationship between democracy and development. For example, Humphreys, Masters, and Sandbu (2006) explore the role of leaders in democratic deliberations in São Tomé and Príncipe; Bertrand et al. (2007) collaborate with the International Finance Corporation to study corruption in the allocation of driver's licenses in India; Blattman, Fiala, and Martinez (2008) study the reintegration of ex-combatants in Northern Uganda; Collier and Vicente (2008) test the effectiveness of an antiviolenence intervention in Nigeria; Moehler (2008) investigates the role of private media in the strengthening of accountability; Levy Paluck and Green (2009) examine how media broadcasts affect interethnic relations in a post-conflict context; Fearon, Humphreys, and Weinstein (2009) collaborate with the International Rescue Committee to evaluate the impact of a community-driven reconstruction program in Liberia.³ All of these studies were made possible in large part through collaboration with local and international NGOs.

Interventions led by NGOs can shed much light on social phenomena in contexts where the involvement of independent actors comes naturally, such as in the experiments described previously. There are cases, however, where one must give special consideration to the effect that an NGO's involvement may itself have on the social phenomena at hand. Ravallion (2008) writes:

the very nature of the intervention may change when it is implemented by a government rather than an NGO. This may happen because of unavoidable differences in (inter alia) the quality of supervision, the incentives facing service providers, and administrative capacity (17).

³ For two excellent summaries of these studies, see Humphreys and Weinstein (2009) and Moehler (2010).

Moreover, there are social contexts where an NGO's involvement is not easily justified. In such cases, researchers have two options. First, they can undertake the enterprise of forging alliances with the relevant actors, such as government officials or politicians, required to randomize an intervention of substantive interest. Second, they can take advantage of the growing number of cases where natural experiments are already in place due to policymakers' decisions to randomize an intervention of interest.

2. Field experiments in collaboration with politicians

Wantchekon's (2003) study of clientelism and its electoral effectiveness in Benin is an example of a unique collaboration between researchers and politicians to implement a treatment. Wantchekon worked directly with presidential candidates to embed a field experiment in the context of the first round of the March 2001 presidential elections. Together with the candidates, Wantchekon randomly selected villages to be exposed to purely clientelist or purely public policy platforms.

Prior to this study, scholars had given little attention to the effectiveness of clientelist and programmatic mobilization strategies. Stokes (2007) notes that "most students and casual observers of clientelism assume that it works as an electoral strategy - that, all else equal, a party that disburses clientelist benefits will win more votes than it would have had it not pursued this strategy. In general we do not expect parties to pursue strategies that are ineffective. And yet we have some theoretical reasons for believing that conditions are not always ripe for clientelism" (622). The challenge of estimating the effectiveness of clientelism, patronage, and pork-barrel as mobilization strategies rests in the possibility that electoral performance can shape spending decisions (Stokes 2007).

The Benin experiment empirically validates the argument that clientelist appeals are a winning electoral strategy, whereas public policy appeals produce mixed results. Beyond confirming these arguments, the Benin experiment presents a wide range of new results that are counterintuitive and could not likely have been derived from any other form of empirical research because in Benin we almost never observe a candidate campaigning on public policy. The experiment shows for instance that 1) clientelist appeals reinforce ethnic voting (not the other way around), 2) voters' preference for clientelist or public goods messages depends in large part on political factors, such as incumbency, and on demographic factors, such as gender, and 3) the lack of support for programmatic platforms is not due to opposing preferences among groups, level of education, or poverty, but instead to the fact that public policy platforms lack credibility, presumably because they tend to be vague.

In a follow-up experiment implemented in the context of the 2006 presidential elections, Wantchekon (2009) finds that broad-based platforms can be effective in generating electoral support when they are specific and communicated to voters through town hall meetings. As a result of these experiments, discussions of how to promote broad-based electoral politics in Benin now have empirical basis.

3. Natural Experiments

While experiments like Wantchekon's (2003) are still rare, scattered throughout the literature on development are examples of randomized interventions where assignment of treatment is outside of researchers' control. Chattopadhyay and Duflo's (2004) study of the quota system for women's political participation and the provision of

public goods in India is such an example. The natural experiment was facilitated by the 73rd Amendment, which required that one-third of Village Council head positions be randomly reserved for women. Chattopadhyay and Duflo's evidence confirms that correcting unequal access to positions of representation leads to a decrease in unequal access to public goods. To begin with, the quota system was effective. In the two districts studied (West Bengal and Rajasthan), all positions of chief in local village councils (Gram Panchayats, henceforth GPs) reserved for women were, in fact, occupied by females. In turn, having a woman chief increased the involvement of women in GPs' affairs in West Bengal, but had no effect on women's participation in GPs in Rajasthan. Moreover, the increase in women's nominal representation translated into substantive representation.

The study of the quota system shows that women invest more in goods that are relevant to the needs of local women: water and roads in West Bengal and water in Rajasthan. Conversely, they invest less in goods that are less relevant to the needs of women: nonformal education centers in West Bengal and roads in Rajasthan. The evidence from this study confirms that some classic claims of representative democracy, such as the relevance of rules and the identity of representatives, hold true. Subsequent studies, however, show that despite institutional innovations, political inequalities and prejudice continue to bias the representation system against minority and disadvantaged groups. In particular, once the GPs' chief position was no longer reserved for women, none of the chief women were reelected, even though villages reserved for women leaders have more public goods and the measured quality of these goods is at least as high as in nonreserved villages (Duflo and Topalova 2004).

In Latin America, Ferraz and Finan (2008) make use of a natural experiment to study the effects of the disclosure of local government corruption practices on incumbents' electoral outcomes in Brazil's municipal elections. The research design takes advantage of the fact that Brazil had initiated an anti-corruption program whereby the federal government began to randomly select municipal governments to be audited for their use of federal funds. To promote transparency, the outcomes of these audits were then disseminated publicly to the municipality, federal prosecutors, and the general media. Ferraz and Finan compare the electoral outcomes of mayors eligible for reelection between municipalities audited before and after the 2004 municipal elections.

Ferraz and Finan find that, conditional on the level of corruption exposed by the audit, incumbents audited before the election did worse than incumbents audited after the election. Furthermore, in those municipalities with local radio stations, the effect of disclosing corruption on the incumbent's likelihood of reelection was more severe. This finding is in line with earlier contributions that show how access to information affects the responsiveness of governments. Moreover, it also corroborates findings that the media is important to diffuse information and discipline incumbents for poor performance (Besley and Burgess 2002; Stromberg 2004).

De La O's (2008) study of the electoral effects of the Mexican conditional cash transfer program (Progresa) is a third example of the use of a natural experiment. Finding the electoral effectiveness of programmatic spending presents similar challenges to the ones previously discussed. In order to evaluate the causal effect of spending, one needs to find exogenous variation on it. De La O empirically examines whether Progresa influenced recipients' voting behavior by taking advantage of the fact that the first rounds

of the program included a randomized component. Five hundred and five villages were enrolled in the program twenty-one and six months before the 2000 presidential election. De La O finds that the program increased turnout in 2000 by five percentage points and increased the incumbent's vote share by four percentage points.

4. Lab and lab-in-the-field experiments

Research opportunities such as the ones described in previous sections are becoming more common as governments, NGOs and sponsors around the world are giving priority to the systematic evaluation of interventions. There are, however, other questions central to the field of political economy of development that require a deeper understanding of the microfoundations of social processes. For example, what determines preferences over redistribution? Why do some individuals behave in a self-interested way while others seem to be altruistic? Why do some communities prefer private over public goods? Why is inequality tolerated more in some places than others? What triggers reciprocity?

Political scientists have found experimentation in the laboratory useful to study these and many other questions. The laboratory gives researchers complete control over assignment to treatment, the treatment itself, and -- perhaps most alluring -- control over the setting where subjects are exposed to the treatment. The price that researchers pay for the internal validity of experimental results produced in a laboratory is a well-known critique about external validity. Concerns about generalizability, however, are not a dismissal of laboratory experiments. Rather, they are an opportunity for creative

researchers (Camerer 2003). Indeed, recent studies have shown that lab-based experimentation needs not to be confined to universities.

Habyarimana et al. (2007), for example, take the experimental laboratory to Uganda to study the mechanisms that link high levels of ethnic diversity to low levels of public goods provision. In this study, subjects are naturally exposed to high ethnic diversity on a daily basis. Thus the conclusions drawn from the dictator, puzzle, network, and public goods games played by Ugandan subjects speak directly to the social phenomenon of interest.

The games in Uganda show that laboratory experimentation enables researchers to adjudicate among complex mechanisms that in less controlled settings would be confounded. For example, Habyarimana et al. find that ethnic diversity leads to lower provision of public goods, not because co-ethnics have similar tastes or are more altruistic, but because people from different ethnic groups are less linked in social networks. Therefore, the threat of social sanction for people that do not cooperate is less credible.

5. Challenges for experiments

Internal Validity

The advantage of experiments compared to observational research is that random assignment ensures that, in expectation, the treatment groups have the same observable and unobservable baseline characteristics. As the editors of this volume note in the introduction, however, random assignment alone does not guarantee that the experimental outcome will speak convincingly to the theoretical question at hand. The interpretation of

the experimental result is a matter of internal validity -- whether the treatment of interest was, in fact, responsible for changing outcomes. For example, in a pioneering field experiment, Levy Paluck and Green (2009) seek to gauge the causal effect of listening to a radio program aimed at discouraging blind obedience and reliance on direction from authorities, and promoting independent thought and collective action in problem solving in post-genocide Rwanda. Research assistants played the radio program on a portable stereo for the listener groups. The challenge of this experimental design in terms of internal validity is that listener groups often lingered to chat after the radio program finished. Therefore, the effect of the radio program could be conflated with the effect of socialization. Levy Paluck and Green successfully dealt with this challenge by recording on standardized observation sheets the lengths and subjects of discussions during and after the program. With this information, they could test whether groups exposed to a particular radio program socialized more than other groups.

The interpretation of experimental results also depends on what the control group receives as treatment. In the experiment in Rwanda, for example, the control group listened to an educational-entertainment radio soap opera, which aimed to change beliefs and behaviors related to reproductive health and HIV. The average treatment effect is therefore the relative influence of the different content of the radio programs. This comparison is different from a comparison between those who listen to a radio program and those who don't listen to anything at all. A comparison between a group of listeners and a control group, however, would be problematic in terms of internal validity because the treatment group would not only be exposed to radio program content but also to group meetings, interactions with research assistants, and so on.

More generally, researchers in political economy of development face three challenges. First, because of the nature of the subject, researchers in development and democracy need to forge alliances with relevant decision makers to study social phenomena. These alliances make research more realistic, but also more challenging. Policymakers, both in government and NGOs, are interested in maximizing the effect of a specific intervention and it is natural for them to endorse treatments that consist of a bundle of interventions. For example, Green et al. (2010), in partnership with the Sarathi Development Foundation, implemented a field experiment in India during the 2007 election to examine how voters in rural areas would respond to messages urging them not to vote on caste lines but to vote for development. The treatment consisted of puppet shows and posters. This bundle of interventions is attractive from the NGO perspective, but is challenging for researchers who want to estimate the average treatment effect of an educational campaign.

To make the challenge more explicit, assume that in the example of Green et al.'s Indian field experiment compliance with the research protocol was perfect. If the effects of posters and puppet shows are independent from each other, then the effect of the bundled intervention is equal to the sum of the effects of the individual components of the intervention. By contrast, if the effects of posters and puppet shows are not independent, then there are four possibilities: posters might magnify the effect of puppet shows and vice versa or, alternatively, posters might cancel out the effect of puppet shows (and vice versa). In this particular application, it might not be theoretically relevant to isolate the effects of the two components of the treatment. In other applications, however, the degree to which an experiment can shed light onto a

theoretical question will depend on how the individual components of bundled treatments map onto theoretically relevant variables.

The second challenge faced by experimental researchers is that logistical difficulties of working in the field oftentimes compromise compliance with research protocols. One form of noncompliance occurs when those assigned to the treatment group do not receive the treatment. In this case, the randomly assigned groups remain comparable, but the difference in average outcomes does not measure the average treatment effect. For example, De La O et al. (2010) design an informational campaign in Mexico where households in randomly selected polling precincts receive a flyer with information about their municipal government's use of a federal transfer scheme aimed at improving the provision of public services. Complying with the research protocol was more challenging in some of the experimental sites than in others because some of the polling precincts were more isolated. Naturally, easy-to-access precincts are different than harder-to-access precincts – easy-to-access precincts are more urban and wealthier than the other precincts. These sociodemographic differences are directly correlated to partisanship. Thus, in this example, noncompliance in the form of failure-to-treat could greatly compromise the experimental design. De La O et al. circumvent the problem of noncompliance by including several mechanisms of supervision in the distribution of flyers, including the use of GPS receivers and unannounced audits.

An alternative form of noncompliance occurs when a treatment intended for one unit inadvertently treats a unit in another group. The risk of spillover effects is prevalent in the study of politics of development. In the Rwanda experiment, for example, the radio program was also being nationally broadcasted, so listeners in both treatment groups

could listen to the program independent of the study. To minimize spillover effects, Levy Paluck and Green use strategies, such as offering to give participants in both groups the cassettes containing the radio program they were not suppose to listen to at the end of the study. An alternative strategy to deal with problems generated by spillovers is for researchers to choose a unit of analysis that enables them to estimate overall treatment effects. For example, Miguel and Kramer (2004) design a field experiment in Kenya where de-worming drugs are randomly phased into schools, rather than provided to individuals. With this design, Miguel and Kramer can take into account the fact that medical treatment at the individual level has positive externalities for nontreated individuals in the form of reduced disease transmission.⁴

External Validity

Field experiments are valuable tools for the study of development and democracy, but designing and executing an experiment that speaks convincingly to theoretical questions of interest to the field presents some challenges, in addition to the ones discussed in the previous section. Just like field researchers, experimental researchers face a tradeoff between the depth of knowledge that comes from studying a particular population and the generalizability of their findings (Wood 2007).

In order to address challenges to external validity, researchers must design their experiments with four things in mind. First, it is often the case that researchers need to exert great effort to include in a study the subset of the population worth studying, rather than the subset of the population that is most readily available to participate in a randomized trial. For example, Habyarimana et al. (2007) recruit their subjects from an

⁴ For more details on Miguel and Kramer (2004) experiment please see Nickerson in this volume.

area in Uganda characterized by high levels of ethnic diversity and low levels of public goods provision. In the Rwandan experiment, Levy Paluck and Green (2009) include two genocide survivor communities and two prisons in their fourteen experimental sites.

Fearon, Humphrey, and Weinstein's (2009) study includes communities in post-conflict Liberia where the majority of the population had been affected by war either because they experienced violence or were displaced.

Second, the context of an experiment must resemble the context of the social phenomenon of interest. For example, in the experiment in Mexico, De La O et al. (2010) distribute to households the information about municipal spending of the infrastructure fund close to the election day. An alternative design would be to recruit individuals for a study where similar information would be distributed in informational meetings directed by the researchers. This design, however, comes less naturally than that of flyer distribution – a widely used communication technique in developing countries.

Third, researchers must find creative ways to design treatments that resemble the variables of interest in the real world. In this sense, not only the treatment but the scale of a field experiment must be taken into account when thinking about external validity.

Consider the recent trend in the field, where researchers collaborate with policymakers to evaluate an intervention in its pilot phase. Within these partnerships, policymakers welcome researchers' interventions in small-scale versions of larger policy projects. Yet, as Deaton (2009) explains:

small scale projects may operate substantially different than their large scale version. A project that involves a few villagers or a few villages may not attract the attention of corrupt public officials because it is not worth their while to

undermine or exploit them, yet they would do so as soon as any attempt were made to scale up. So that there is no guarantee that the policy tested by the randomized controlled trial will have the same effects as in the trial, even on the subjects included in the trial (42).

Finally, researchers must find ways to measure outcomes that resemble the actual outcomes of theoretical interest. Indeed, experiments have in some cases started to revolutionize the field by presenting alternative measures of key concepts, such as corruption and vote buying. Consider Olken's (2007) field experiment in 608 Indonesian villages where treatments were designed to test the effectiveness of top-down and bottom-up monitoring mechanisms to reduce corruption. Unlike much of the empirical work that measures corruption based on perceptions, Olken measured corruption more directly, by comparing two measures of the same quantity, one before and one after corruption. With this innovative measure, Olken found that bottom-up interventions were successful in raising participation levels. However, when compared to the top-down intervention, the bottom-up interventions proved to be less successful at reducing corruption.

Nickerson et al. (2010) present another example where a field experiment innovates the measurement of a critical concept on the field. Numerous qualitative studies of vote buying have concluded that the exchange of votes for gifts or cash is a prevalent practice around the world. Yet studies based on survey research have consistently found surprisingly little evidence of vote buying. Nickerson et al. measured the frequency of vote buying in the 2008 Nicaraguan municipal elections using a survey-based list experiment. All respondents were asked how many activities from a list were

carried out by candidates and party operatives during the elections. The control group was given a list of four activities, including typical campaign activities like hanging posters, visiting homes, placing advertisements in the media, as well as not-so-typical activities like making threats. The treatment group was given the same list of activities, with the addition of vote buying. Since respondents were not asked which of the activities they witnessed but rather how many, a certain degree of anonymity when reporting vote buying was guaranteed. The proportion of respondents receiving a gift or favor in exchange for their vote was then measured as the difference in responses between the treatment and the control group. Based on the list experiment, the authors estimated that nearly a quarter of respondents received a gift or favor in exchange for their vote. In contrast, less than three percent of respondents reported that they had received a gift or favor when asked directly.⁵

Moving forward, researchers will be confronted with the challenge of designing field experiments in a way that enables the accumulation of knowledge. According to Martel Garcia and Wantchekon (2010), there are two ways to achieve this goal. One option is to replicate as much as possible the relationship between two variables under different conditions (the robustness approach). The ongoing research on the role of information in community development projects illustrates this approach. Banerjee et al. (2010) find that in India a randomly assigned information campaign was not effective at fostering community involvement in Village Education Committees and, ultimately, had no impact on teacher effort or student learning outcomes. By contrast, a similar study in Uganda reveals that, as a result of an informational campaign, people became more engaged in Community-Based Organizations and began to monitor the health units more

⁵ For more details on the origins of the list experiment see Sniderman in this volume.

extensively. This community-based monitoring increased the quality and quantity of primary health care provision (Bjorkman and Svensson 2007).

The examples provided in this section show that, even in cases where similar experiments are executed across two different populations, contextual differences could cause the same intervention to have different effects. An alternative to replicating similar treatments in different contexts is to use an analytical approach that makes the theoretical foundations of an experiment more explicit (Martel Garcia and Wantchekon 2010). This analytical approach brings front and center the mechanisms that link a causal variable to an outcome. By being explicit about mechanisms, researchers can develop trajectories of experiments that are suitable to test theoretically informed hypotheses.

Consider, for example, the Benin electoral experiments (see Wantchekon 2003, 2009). One of the findings of the 2001 experiment is that voters are more likely to react positively to a public goods message when it comes from a coethnic candidate. A possible explanation for this finding is that voters trust a candidate from their ethnic group more than they trust a candidate from another group. This means that the mediating variable between ethnic ties and votes is trust, or the credibility of the candidate. By testing the relationship between credibility of candidates and voting behavior in a follow-up experiment in 2006, Wantchekon (2009) improves the external validity of the results of the 2001 experiment. As the Benin electoral experiments illustrate, to make scientific progress in this field, new experimental designs should not only take into consideration the context of current experiments, but should also focus on testing various aspects of a theory in a coherent way.

On the Ethics of getting involved in elections

One of the most striking features of experiments on democracy is that they require researchers to work directly with policymakers, politicians or government officials and to get involved in, in many cases, with running elections, government programs, or education campaigns. Embedding experiments in the context of real elections and programs brings a great degree of realism to the treatments. However, what is gained in terms of the external validity of the experimental results may not sufficiently offset ethical concerns.

We are far from having a consensus on where to draw the line between interventions that are ethical and interventions that are not. Nevertheless, there are several guidelines that researchers can follow when designing an experiment. First of all, an intervention will raise fewer ethical concerns if the units under study are exposed to a treatment they would ordinarily seek. In the Benin experiments, for example, the clientelist treatment could at first glance be a source of concern. Candidates in Benin, however, typically run campaigns based on clientelist appeals, regardless of researchers' presence. In such experiments, the researcher was merely acting as an unpaid campaign advisor to the candidate or civic educator. The researcher's main contribution was to suggest random assignment of campaign messages to districts. If anything, random assignment of messages is more ethical than the standard opportunistic tailoring of messages to what voters want to hear.

A similar concern is raised by experimental designs where subjects in one group are denied a treatment that they would ordinarily seek. For example, a study undertaken to examine the effect of international aid, where some villages are randomly selected to

receive aid and some equally needy villages are randomly selected to be denied aid, is bound to raise ethical questions. Practical considerations, however, can help researchers mitigate these concerns. For example, in most cases, NGOs and governments have limited budgets that force them to make decisions regarding where to start an educational campaign, a social policy, or any other intervention of interest. Random assignment in these cases provides policymakers with a transparent and fair way to decide the order in which subjects are, for example, enrolled in a program.

An ongoing field experiment in Uganda illustrates this empirical strategy. Annan et al. (2010), in collaboration with Innovations for Poverty Action and the Association of Volunteers in International Service (AVIS), are evaluating the Women's Income Generating Support (WINGS) program that provides women with grants and business training. To find whether small grants empower women and shape their political participation, Annan et al. will enroll women to the program in different phases over the course of three years. The order of enrollment is randomly assigned. This design enables causal inferences, but no vulnerable household contacted by researchers will be left out of the program.

A second way to think about ethical issues is to ask: what are the costs to subjects of participating in an experiment? In the Benin examples, if there were a cost to voters for being exposed to clientelist messages, this cost is already routinely incurred in all elections. In fact, the whole purpose of the experiment was to lower the cost of this campaign strategy for voters in future elections. More generally, experimental designs must take into account the costs of exposing subjects to treatments including, but not

limited to, material costs (e.g., the opportunity costs of spending time in the study), psychological costs, and even physical costs.

A third set of ethical issues that researchers must take into account is the degree to which interventions alter the outcomes and the costs associated with such departures. For example, in the experimental study of elections, one common concern is that researchers change the result of an election. A 2002 *New York Times* article commenting on the 2001 Benin experiment stated: “There are some major ethical concerns with field experiments in that they can affect election results and bring up important considerations of informed consent” (Browning 2002). Wantchekon, however, suppressed this possibility by including in the experiment only safe districts, where candidates collaborating in the study had a stronghold.

In this particular example, the subset of districts where ethical concerns are manageable coincided with the subset of districts that were theoretically relevant to study, because clientelism is more resilient in districts where one political machine has a monopoly than in districts where there is more political competition. In other applications, restricting the experiment to certain subpopulations where ethical concerns are manageable may compromise the external validity of the experiment’s results.

Finally, many research questions in the political economy of development, like the effect of violence on development, involve interventions that are difficult to study through experimentation without raising ethical concerns. Creative experimental designs, however, can enable researchers to study social phenomena that at first glance seem out of reach. For example, Vicente (2007) conducted a field experiment in São Tomé and Príncipe to study vote buying. As in many other countries, buying votes is illegal in São

Tomé. Thus, Vicente randomly assigned subjects to be exposed to an anti-vote buying campaign, which was sponsored by the National Electoral Commission.

6. Concluding Remarks

The rise of experiments as one of the most prominent empirical strategies has led to new advances in the study of democracy and development. So far, some experimental results have confirmed previous arguments, such as the effectiveness of clientelism as a mobilization strategy and the prevalence of political and social inequalities despite institutional innovations. Other experiments have revealed relationships that only a randomized control trial could uncover, like the fact that clientelist appeals reinforce ethnic voting and not the other way around. Finally, some experiments are revolutionizing the measurement of core concepts in the field. For example, we now know that vote buying measured experimentally is more prevalent than what observational studies suggested.

Going forward, field experiments in collaboration with policymakers, governments, and NGOs are a promising line of research. The next round of experiments, however, faces considerable challenges, including those we have highlighted throughout this chapter. First, researchers must find creative ways to design interventions that are attractive to potential partners but that still speak convincingly to theoretically relevant questions. In doing so, researchers must pay special attention to internal validity issues. Second, a more analytical approach would help guide researchers to design experiments that enable significant accumulation of knowledge to take place. Finally, as the scope of

experimentation expands, the tradeoff between external validity and ethical concerns will become more salient.

Despite these challenges, experimental research on development and democracy is a productive and exciting endeavor. As insightful as the experimental research has been up until now, numerous substantive questions remain unanswered. Hopefully the selection of studies covered in this chapter illustrate how experiments can be used as a research tool to study broader and more central questions about the relationship between democracy and development.

References

- Annan, Jeannie, Christopher Blattman, Eric Green, and Julian Jamison. 2010. "Uganda: Enterprises for Ultra-poor Women after War." Concept note. Yale University.
Retrieved from: <http://chrisblattman.com/projects/wings/> (6/29/10)
- Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence From a Randomized Evaluation in Education in India." [*American Economic Journal: Economic Policy* 2](#): 1-30.
- Bertrand, Marianne, Simeon Djankov, Rema Hanna, and Sendhil Mullainathan. 2007. "Obtaining a driver's license in India: An experimental approach to studying corruption." *Quarterly Journal of Economics* 122: 1639-76.
- Besley, Timothy, and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117: 1415-1452.

- Blattman Christopher, Nathan Fiala, and S. Martinez. 2008. "Post-conflict youth livelihoods: an experimental impact evaluation of the Northern Uganda Social Action Fund (NUSAF)." Concept note and design document. Washington, DC: World Bank.
- Boix, Carles, and Susan Stokes. 2003. "Endogenous Democratization." *World Politics* 55: 517-549.
- Bjorkman, Martina, and Jakob Svensson. 2007. "Power to the People: Evidence from a Randomized Field Experiment of a Community-Based Monitoring Project in Uganda." World Bank Policy Research Working Paper No. 4268.
- Browning Lynnley. 2002, August 31. "Professor Offer A Reality Check for Politicians" The New York Times. <http://www.nytimes.com/2002/08/31>
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Chattopadhyay, Raghendra, and Esther Duflo. 2004. "Women as Policymakers: Evidence from a Randomized Policy Experiment in India." *Econometrica* 72: 1409-1443.
- Collier P, Vicente P. 2008. "Votes and violence: experimental evidence from a Nigerian election." Households in Conflict Network (HiCN) Working Paper No. 50.
- De La O, Ana L. 2008. "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico." Typescript. Yale University.
- De La O, Ana L., Alberto Chong, Dean Karlan, and Leonard Wantchekon. 2010. "Information Dissemination and Local Governments' Electoral Returns, Evidence

- from a Field Experiment in Mexico.” Paper presented at conference on Redistribution, Public Goods and Political Market Failures, Yale University.
- Deaton, Angus. 2009. “Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development.” Unpublished manuscript. Princeton University.
- Duflo, Esther, and Petia Topalova. 2004. “Unappreciated Service: Performance, Perceptions, and Women Leaders in India.” Faculty working paper. Massachusetts Institute of Technology, Department of Economics.
- Fearon James D., Macarthan Humphreys, and Jeremy Weinstein. 2009. “Can Development Aid Contribute to Social Cohesion After Civil War? Evidence from a Field Experiment in Post-conflict Liberia.” *American Economic Review: Papers and Proceedings* 99: 287-291.
- Ferraz, Cláudio, and Frederico Finan. 2008. “Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes” *Quarterly Journal of Economics* 123: 703-745.
- Green, Jennifer, Abhijit Banerjee, Donald Green, and Rohini Pande. 2010. “Political Mobilization in Rural India: Three Randomized Field Experiments.” Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL.
- Habyarimana, James, Macarthan Humphreys, Daniel N. Posner, and Jeremy Weinstein. 2007. “Why Does Ethnic Diversity Undermine Public Goods Provision?” *American Political Science Review* 101: 709-25.

- Humphreys, Macartan, William A. Masters, and Martin E. Sandbu. 2006. "The role of leaders in democratic deliberations: results from a field experiment in Sao Tome and Principe." *World Politics* 58: 583-622.
- Humphreys, Macartan, and Jeremy Weinstein. 2009. "Field Experiments and the Political Economy of Development." *Annual Review of Political Science* 12: 367-78.
- Keefer, Philip. 2007. "The Poor Performance of Poor Democracies." In *The Oxford Handbook of Comparative Politics*, eds. Carles Boix and Susan Stokes. New York: Oxford University Press.
- Levy Paluck, Elizabeth and Donald P. Green. 2009. "Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda." *American Political Science Review* 103: 622-644.
- 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.
- 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. 2008. "Tune in to governance: An experimental investigation of radio campaigns in Africa." Paper presented at conference on Field Experiments in Comparative Politics and Policy, University of Manchester, UK.
- Moehler, Devra C. 2010. "Democracy, Governance, and Randomized Development Assistance." *The ANNALS of the American Academy of Political and Social Science* 628: 30-46.

- Nickerson, David, Ezequiel Gonzalez-Ocantos, Chad Kiewiet de Jonge, Carlos Meléndez, and Javier Osorio. 2010. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." Unpublished manuscript. University of Notre Dame.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115: 200-249.
- Olken, Benjamin. 2010. "Direct Democracy and Local Public Goods, Evidence from a Field Experiment in Indonesia." *American Political Science Review*, forthcoming.
- Przeworski, Adam, and Limongi Neto, Fernando. 1997. "Modernization: Theories and Facts." *World Politics* 49: 155-183.
- Ravallion, Martin. 2008. "Evaluation in the Practice of Development." Policy Research Working Paper No, 4547. Washington, DC: World Bank.
- Shepsle, K. 2006, "Old Questions and New Answers about Institutions: The Riker Objection Revisited." In *The Oxford Handbook of Political Economy*, eds. Barry R. Weingast and Donald A. Wittman. New York: Oxford University Press.
- Stokes, Susan. 2007. "Political Clientelism." In *The Oxford Handbook of Comparative Politics*, eds. Carles Boix and Susan Stokes. New York: Oxford University Press.
- Stromberg, David. 2004. "Radio's Impact on Public Spending." *Quarterly Journal of Economics* 119: 189-221.
- Vicente, Pedro C. 2007. "Is vote buying effective? Evidence from a randomized experiment in West Africa." Economics Series Working Papers No. 318. Oxford, UK: University of Oxford, Department of Economics.

Wantchekon, Leonard. 2003. "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin." *World Politics* 55: 399-422.

Wantchekon, Leonard. 2009. "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin. Working Paper. New York University.

Wood, Elizabeth. 2007. "Field Research." In *The Oxford Handbook of Comparative Politics*, eds. Carle Boix and Susan Stokes. New York: Oxford University Press.