Review in Advance first posted online on March 12, 2009. (Minor changes may still occur before final publication online and in print.)

# Field Experiments and the Political Economy of Development

Macartan Humphreys[1] and Jeremy M. Weinstein[2]

[1]Department of Political Science, Columbia University, New York, New York 10027; email: mh2245@columbia.edu

[2]Department of Political Science, Stanford University, Stanford, California 94305; email: jweinst@stanford.edu

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org by Columbia University on 03/23/09. For personal use only.

Annu. Rev. Polit. Sci. 2009. 12:367–78

The *Annual Review of Political Science* is online at polisci.annualreviews.org

This article's doi: 10.1146/annurev.polisci.12.060107.155922

Copyright © 2009 by Annual Reviews. All rights reserved

1094-2939/09/0615-0367$20.00

## Key Words

experimental methods, political institutions, social order

## Abstract

Social scientists have begun to work alongside developing country governments, nongovernmental organizations, and international organizations on experimental projects that address fundamental questions in the political economy of development. We describe the range of projects that have taken place or are currently under way, identify new and promising frontiers for research, and discuss some challenges that are particular to work in this area. The impact of this research will depend on the extent to which scholars can successfully link studies of experimental interventions to broader questions of social scientific interest.

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

## INTRODUCTION

In recent years, social scientists have begun to work alongside developing country governments, nongovernmental organizations (NGOs), and international organizations on experimental projects that provide development assistance to some communities but not others, that make the provision of aid conditional on past performance in some places and not others, and that provide voters with information about the performance of their elected representatives in some constituencies but not others. Such field experiments—and we describe many more examples below—are emerging as a key tool employed by scholars to study the political economy of development. This rapid growth is a reflection of changes in two distinct arenas: development practice and social science scholarship.

Development practitioners have recently called for sustained efforts to strengthen the evidence base on which interventions rest (Savedoff et al. 2006). Development agencies and implementing organizations find themselves under pressure—from a growing aid constituency and the politicians and philanthropists who control the purse strings for development funding—to identify the most effective strategies for promoting development.

This push for an evidence-based approach has been apparent for some time with respect to interventions in health, education, and agriculture, where, confronted by the difficulty of assessing the merits of various interventions using traditional methods, practitioners have embraced the use of randomized controlled trials. However, this pressure is now increasingly being brought to bear on agencies that implement interventions designed to change political processes (to promote what practitioners call "good governance"). These programs include, for example, investments in civil society organizations, the reform of bureaucracies, and the building of institutional capacity at the national and local levels—approaches that have become more central to the programming of development agencies over the past two decades. However, surprisingly little is known about the effectiveness of such interventions (Committee on the Evaluation of USAID Democracy Assistance Programs 2007). So, along with increased funding to promote better governance, there is a new demand for research that can figure out what works, when, and why.

Among social scientists, there is increased recognition (although not for the first time) of the merits of experimental methods for identifying causal effects (Oakley 2000). This is true in particular for scholars studying the politics of economic development, a subfield for which problems of endogeneity and unobserved selection loom large. In attempting to take seriously issues of econometric identification, scholars are increasingly conscious of the limitations of existing approaches to observational data, including the use of controls, instrumental variables, regression discontinuity, and propensity score matching when the process of assignment is unknown. At the same time, a small group of pioneers (many affiliated with the Poverty Action Lab at MIT) has demonstrated the potential of field experiments to address core questions about the development process.

There is also a growing appreciation among social scientists of the links between issues relevant to policy makers and questions of a more general, social scientific nature. In part, this is because consumers of policy analysis are demanding a stronger base of evidence for policy prescriptions. It is also the result of an accumulation of scholarship that shows how efforts to address policy concerns can create opportunities to answer questions at the core of social science.

This convergence of the interests of donors, implementing agencies, and social scientists represents a real opportunity for political scientists who study how political processes shape development. Funders want to know whether the valuable resources they invest in promoting good governance are paying off. Social scientists are optimistic about the potential of experimental approaches to yield valuable new insights on the origins and impact of political institutions. The result is a range of



ANNUAL REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

collaborations (and opportunities for collaboration) in which individual projects are designed to shed light on general problems in the political economy of development.

## WHAT FIELD EXPERIMENTS ARE AND HOW THEY ARE DONE

Cox & Reid (2000) define experiments as investigations in which an intervention, in all its essential elements, is under the control of the investigator. The key to distinguishing among the various families of experiments—natural experiments, lab experiments, and field experiments—is determining what or how much is under control.

We can usefully distinguish between two types of control: control over assignment to treatment and control over the treatment itself. The key goal of controlling assignment to treatment is to ensure that, insofar as possible, treated and untreated units are identical except that one group receives the treatment. This is important for isolating causal effects. In social science experiments, investigators typically seek to achieve this balance between treatment and control units through the random assignment of treatment status—a procedure that guarantees balance in expectation. The goals of controlling the treatment itself are to permit interventions of theoretical interest that cannot readily be examined in more natural settings, and, insofar as possible, to keep other factors fixed rather than simply balanced.

Traditionally, lab experiments have sought to maintain as much control as possible over both the treatment and the assignment to treatment. By contrast, in natural experiments, the researcher controls neither but seeks to find assignment processes that create comparable treatment and control groups "naturally" (in this sense, natural experiments are not experiments at all according to Cox & Reid's definition).[1]

Field experiments occupy an intermediate position: Researchers typically maintain control over the assignment to treatment while forgoing control over the treatment itself. Features such as the characteristics of subjects, the information available to them, and the precise manner and context in which the treatment is applied are more likely to take on values given by "nature" rather than being set at the discretion of the investigator. In the political economy of development, many field experiments go a step further and seek to study subject behavior in actual political processes, often implemented by someone other than the experimenter (these are sometimes referred to as "policy experiments"). These field experiments are designed to retain the advantage of control over the assignment of treatment to ensure internal validity (in particular, in establishing causality), while seeking greater external validity than lab experiments can generally achieve. But these gains come at a cost: The more that control over the treatment is lost, the more difficult it is to match experiments to quantities of theoretical interest and to get precise estimates of treatment effects.[2]

When field experiments are applied to questions about the political economy of development, a relatively simple formula is typically employed. Researchers forge a partnership with an implementing agency (or with politicians, NGOs, or other activists) before the implementation of a treatment. A set of units that

[1] Many studies that use the term natural experiment fall short of this ideal. For an example in which actual randomization processes justify the use of the term, see Chattopadhyay & Duflo (2004); for a discussion of a case with near-random assignment, see Hyde (2007).

[2] Whether or not this difference is properly thought of as a difference in "control" can be disputed (Harrison & List 2004). The key feature, however, is that in field experiments, many values are not freely selected by investigators. This is especially likely when experiments involve actual (rather than simulated) political processes. In this way, field experiments also differ from approaches in which investigators take lab experiments to the field but still seek to maintain control over the treatment (e.g., Henrich et al. 2004, Habyarimana et al. 2007), as well as studies in the growing field of survey experiments (Gibson 2008). This distinction between lab and field can however be easily overdrawn and, when treatments implemented by investigators have real-world effects, may be meaningless from the perspective of subjects.



REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

will form the treatment and control groups is then identified by the agency. Baseline measures are taken prior to randomization. Treatment units are selected in a lottery (sometimes public, sometimes private; sometimes involving blocking to help ensure balance, sometimes not; sometimes with treatment assigned to individuals, but often with treatment assigned to clusters). The intervention then takes place, after which final outcome measures are gathered. Data analysis tends to consist of simple difference in means tests, although models are often used to account for clustering, noncompliance with treatment, or, when necessary, the lack of balance that results from finite samples.

## OPPORTUNITIES

The central questions in the political economy of development focus on how politics affects development trajectories and how development reshapes politics. What are the political determinants of economic growth? What accounts for the different development trajectories of different areas? What are the political consequences of different types of economic change? These questions in turn require an understanding of more abstract questions. How is political order established? How does coercive capacity shape economic decisions? How are political decisions made in different contexts? What are the foundations of political accountability? We can use field experiments to explore many of these questions, provided we can map these issues of broad theoretical importance to the more specific questions that interest development practitioners. Finding such a mapping is not always straightforward.

The first column of **Table 1** lists policy questions that implementing agencies consistently ask about their own work. These range from the standard issue of whether assistance is actually driving improvements in human welfare to more nuanced questions about the relationship between political processes and development outcomes. Each of these questions of concern to policy makers corresponds to a question of concern to social scientists. Questions about aid effectiveness are related to underlying debates about the factors that drive economic growth. Concerns about corruption and the poor performance of government are linked to broader questions about how formal and informal institutions can shape and constrain political action. Efforts to ensure that development processes are participatory require an understanding of how preferences are aggregated in political systems. These are some of the central questions in the study of the political economy of development. Drawing out the connections between the issues that bedevil practitioners and the broader questions that drive social science research is the key to turning a field experiment from an exercise in impact evaluation to a more fundamental contribution to our understanding of development.

Although this field is relatively new, a number of field experiments have attempted to engage these broader debates. We describe only some of these efforts here, organizing our discussion around four of the major questions examined by scholars of the political economy of development. We begin with a caveat, however: In contrast to other areas of scholarship (such as the large literature on get-out-the-vote campaigns in American politics), there has not yet been a significant accumulation of knowledge from the use of field experiments in the political economy of development. Many projects are still in process, and those findings that exist have not been subjected to replication in other contexts. For this reason, we focus more on the promise of the field than on its achievements.

### Political Institutions and Economic Growth

For students of politics, one key question in the literature on economic growth is the extent to which political institutions matter in the development process. Cross-national (nonexperimental) work has struggled to convincingly identify the impact of political institutions on welfare given the difficulty of distinguishing cause and effect (Acemoglu et al. 2001). There is great scope, however, for



ANNUAL REVIEWS IN ADVANCE

**Table 1 Linking policy and basic research**

| Policy questions | Social scientific questions |
|---|---|
| Does aid work? What kind of aid is most effective? | What are the drivers of economic growth? |
| How can community or political structures be strengthened? | How do groups solve collective action problems? |
| How can politicians be held accountable? How can services be delivered in a way that minimizes leakage? | How do principals motivate agents? What is the effect of institutions, information structures, and incentives on political behavior? |
| What is the best way to distribute benefits to populations? | When are markets more efficient than governments for the provision of services? |
| What are the best strategies for building peace after civil war? | What are the origins of social order? |
| How can policy makers ensure that development processes are participatory? | How are preferences aggregated in political processes? |

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org by Columbia University on 03/23/09. For personal use only.

directly studying these effects by introducing random variation in the institutions that are used to make decisions and implement policies in different settings. Olken (2008), for example, examined the different effects of indirect decision making by elected bodies and referenda on budgetary choices through a project with the World Bank and the Government of Indonesia, which varied the way in which local funding decisions were made. An ongoing project in Afghanistan (Beath et al. 2008) varies the way in which community representatives are selected to administer a social development project to examine how single-member and multimember constituencies affect the welfare of the poor. Similar projects could be designed to examine the effects of institutions that protect property rights, decentralize power to lower levels of government, or put responsibility for making policy or spending resources in the hands of elected officials rather than leaders who assume power through hereditary succession.

## Weak States and Collective Action

In many developing areas, collective goods are not provided by states but depend instead on the ability of groups to work together, whether to provide public goods or to launch a rebellion. Theoretical work has focused on the concept of social capital, and empirical work has asked why some communities are able to act collectively while others are not. Much of the existing theoretical literature on the origins of social capital points to long-run evolutionary processes (Bowles & Gintis 2004), and more historical approaches emphasize key junctures at which major exogenous events reshaped social relations (Putnam 1994, Nunn 2008). Nonetheless, development practitioners intent on improving outcomes in the near term have moved quickly to implement interventions designed to increase a community's collective capacity. This set of investments opens the door to the study of the origins of social norms and informal institutions that may facilitate collective action. Fearon et al. (2009), for example, find that participation in a "community-driven development" program in postconflict Liberia increases a community's subsequent capacity to act collectively. They find that two years of exposure to a local institution-building project resulted in an increase in average contributions in a public-goods game approximately equivalent to a fourfold increase in the social rate of return of a private investment. In another study, a team of researchers partnered with an NGO to examine whether outside funding to indigenous women's organizations strengthens or undermines their capacity. Surprisingly, the major impact of the program was on the character of membership, leading to the displacement of more vulnerable, less-educated women from leadership posts (Gugerty & Kremer 2008). Other studies tackle issues of collective action more indirectly. Through a randomized intervention designed to improve educational attainment, Kremer et al. (2008) explore the



REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

determinants of individual political action. Paluck & Green (2008) examine how media broadcasts affect interethnic relations in a post-conflict context and find that exposure to a reconciliation program increases the propensity of listeners to take dissenting positions.

Potential next steps in this area include projects designed to assess (*a*) the conditions under which outside aid increases or undermines local capacity; (*b*) the importance of features such as community size, heterogeneity, and communication structures; and (*c*) which forms of collective action are most responsive to short-term interventions.

## Accountability

Social scientists and practitioners alike have focused attention on the principal-agent issues that plague many aspects of the development process—from service delivery to political representation. Theoretical work suggests that a variety of avenues can be pursued to ensure that agents act more in line with the interests of principals, but empirically, the effects of different strategies have been hard to disentangle from other factors that might explain why such strategies are pursued in the first place.

Field experiments are already being employed to test hypotheses about the impact of institutions, information, and incentives on political behavior. Olken (2007), for example, compares two competing anticorruption strategies in the context of a decentralized road-building project in Indonesia. He finds that the threat of an audit (a top-down strategy) reduces corruption; a change from a 4% to a 100% probability of an audit is associated with an 8% fall in missing expenditures. However, a mechanism for empowering the community to oversee the project (a bottom-up strategy) generally has no perceptible effect on corruption. Duflo et al. (2008) also find that top-down but impersonal monitoring of teachers' attendance (using cameras), coupled with financial incentives, was effective in increasing attendance at schools in India. In this case, monitoring resulted in a fall in absenteeism rates from 42% to 21%.

Other studies have found more support for bottom-up approaches. Bjorkman & Svensson (2009) identify a powerful effect of citizen oversight on the quality of local health care delivery in Uganda, a setting where top-down approaches have not been functioning. Areas in which communities were provided with the information needed to monitor a clinic's health services saw a subsequent 33% decline in child deaths. Many differences in the settings could account for these differences in outcomes. Bjorkman & Svensson, for example, point to the number of nonelites involved and the quality of information available to subjects as being important determinants of the efficacy of bottom-up approaches. As such studies multiply, research can shift from asking whether different interventions are effective to asking under what conditions one strategy may be more effective than another.

Ongoing studies examine related questions in the context of elite behavior. Humphreys & Weinstein (2007) are using the dissemination of performance scorecards to a random subset of the constituencies of Members of Parliament in Uganda to distinguish between selection-based and incentive-based accounts of legislator responsiveness. A project in Indonesia uses a randomized design to assess the efficacy of strategies that national governments can employ to elicit better performance from local governments, in the context of the country's efforts to achieve the Millennium Development Goals.

Many other issues remain to be addressed, including (*a*) how development should be financed in order to maximize performance while eliciting the contribution and cooperation of participating communities; (*b*) whether and how taxation shapes the relationship between constituents and their representatives; and (*c*) how accountability depends on time preferences and on the types of information available to each party.

## Order and Violence

Violent conflict poses a special challenge to economic development. National governments



REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

and international organizations seek strategies to prevent violence and to build peace in its aftermath. But such strategies depend on an understanding of fundamental social scientific questions. How is political order established? How does poverty affect the risk of conflict? How are violent organizations formed and dismantled? Some of these questions are now starting to be addressed by recent and ongoing field experiments. For example, scholars are examining how the organization of police forces in India affects civil–police relations and the incidence of crime (A. Banerjee, R. Chattopadhyay, E. Duflo, D. Keniston, unpublished; **http://www.povertyactionlab.org/projects/project.php?pid=81**). Another experiment examines how infusions of capital to youth organizations affect security in conflict-prone areas of Northern Uganda (Blattman et al. 2008). A third project exploits the implementation of an antiviolence intervention in Nigeria to examine the effects of intimidation on democratic practices (Collier & Vicente 2008). Although at an early stage, these projects demonstrate the potential for field experiments linked to important questions in an area that at first might seem nonconducive to experimentation.

Other unresolved issues include the role of third-party armed forces in peacekeeping, the links between natural resources and conflict, and the influence of different systems of representation on the propensity of actors to pursue violence.

These examples only scratch the surface of how field experiments are being used to advance the study of the political economy of development. Beyond those issues covered here, field experiments are also contributing to the study of political campaigning in the developing world (Wantchekon 2003, Vicente 2007), political preference formation and expression (Guan & Green 2006), and the role of leadership in determining social outcomes (Humphreys et al. 2006). Taken together, they suggest that the common concern about field experiments—that interesting manipulations lie beyond the reach of the experimenter—is no longer accurate. Big, unanswered questions remain, and field experiments are well placed to generate evidence that may begin to answer them.

## LIMITATIONS, CHALLENGES, AND NEW FRONTIERS

Field experiments are an excellent way to obtain a valid estimation of causal effects. However, the approach has substantial limitations. For a more general discussion of concerns with the experimental methods (with rejoinders), see Green & Gerber (2002). Here we focus on challenges that are particularly salient for studies in the political economy of development. We first identify five limitations that we believe are inherent in the method and that will likely continue to bound its usefulness in the field. We then describe four common challenges faced in the use of field experiments—practical constraints, spillovers, external validity, and ethics—before concluding with some thoughts about promising frontiers.

### Limitations Inherent in the Method

Many of the shortcomings of field experiments are being overcome with increasingly clever designs. Nevertheless, there are some constraints, we believe, on what can be done with the method.

**Real time.** Experiments take place in real time and, as a result, empirical studies take at least as long as the processes under study. Some processes of interest to students of development are likely to be rapid in nature, but many others, such as the evolution of cultures, the intergenerational transmission of ideas, and the effects of early education on political behavior, are naturally long in duration and may exceed the lifespan of researchers.

**History has happened.** Many research questions in the political economy of development are linked to historical events rather than to abstract propositions. For example, what was the effect of the 2004 tsunami on conflict in



REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

Southeast Asia? How did the French Revolution affect thinking on human rights? For such historical questions, the explanatory variable is fixed in time and not subject to randomization.

**Power.** For most experimental treatments, large numbers of units are required to generate precise estimates. For some questions, however, the number of units is sharply constrained by the subject of study. This is likely when the researcher cares about a single event or the treatment is something that takes place at a national or international level (e.g., the impact of freedom of the press on democracy in a particular country) or in a single institution (e.g., the consequence of increased salaries on corruption in the public works ministry).

**Variables as attributes.** Some explanatory variables of interest are attributes of subjects and cannot be manipulated without altering the subjects themselves (Holland 1986). For example, scholars of the political economy of development are interested in the effects of gender, religion, and ethnicity. However, these features, by most definitions at least, are not subject to manipulation (although their salience can be primed). Moreover, even if such aspects of an individual could in principle be assigned at random, this assignment would allow an examination of a *change* (e.g., a sex change or a religious conversion), not of the characteristic alone. The problem is that the history that goes along with membership in a given category cannot be transferred, and this history probably matters.

**Assignment to treatment.** In some cases, how a treatment is assigned may itself be a relevant feature of the treatment (Rubin 1986). It is likely to be of particular importance for the strategic interactions examined in political economy settings. For example, a message provided to a voter may have a different effect if the voter knows a politician strategically targeted him or her to receive that message than if it was randomly assigned. In such cases, randomization produces a treatment that is qualitatively different from the treatment of interest in the world.

## Practical Constraints

Many political processes and attributes are likely to remain (and likely ought to remain) unavailable for experimental manipulation: whether governments are authoritarian or democratic, whether regions secede, whether governments launch brutal counterinsurgency campaigns, whether a given individual adopts a given set of preferences, etc. Often such "treatments" are not under the control of a well-identified set of agents, and even political processes that are within the control of some group of actors typically can only be studied if those actors consent to the study. So far, there have been two responses to this constraint. First, for some of these questions, field experimentation is still possible through the use of random variation applied to analogous processes at lower levels. Indeed, most of the experiments cited above have been carried out with local governments, councils, and decision-making processes. Inferential validity depends on the extent to which processes examined at one level operate in the same way at a higher level. Second, even when the treatment of interest cannot be manipulated, the measurement of a causal effect may nevertheless be possible if interventions that affect the treatment are available. In such cases, experiments are used to create instruments for the explanatory variables of interest (and the quality of the instruments needs to be defended in the same way as in observational studies). Collier & Vicente (2008) pursue this strategy to identify the effects of violent intimidation on voting, not by controlling violence directly but by examining an antiviolence campaign implemented in randomly selected constituencies during a recent Nigerian election.

## Spillovers

In social science experiments, there is often a risk that a treatment intended for one unit might affect another. For example, delivery of



REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

an aid program to one community but not another may give rise to jealousy. Differences in outcomes may be interpreted as evidence for a positive program effect on the treated community even if all that has occurred is an adverse effect on the control community. Similarly, the control communities may benefit indirectly from the program, in which case the difference between outcomes in the two groups will underestimate the effect of the program both on the treated community and overall. This problem is especially germane when all the treated and control units are part of a single process of interest to researchers (such as individual participants in military demobilization) or engage in strategic interactions with one another. In such cases, field experiments violate the "stable unit treatment value assumption" (Rubin 1986), rendering the resulting inferences invalid. In many cases, steps can be taken to reduce the scope for spillovers of this form; however, the fact is that spillovers are likely an important part of the social processes under study. For this reason, a focus on directly estimating spillover effects could be an important area for new research (Miguel & Kremer 2004).

## External Validity

One of the most common concerns raised about field experiments (and experiments in general) is that they suffer from weak external validity: It is difficult to know whether the results found in one setting apply in another. However, this concern applies to any empirical examination of a general proposition. The reason cross-country regressions sometimes appear immune to this criticism is that the propositions being tested are often historical (e.g., good institutions produced growth after World War II) and can be tested on the universe of relevant cases rather than a subset. In contrast, general propositions (e.g., good institutions produce growth) are always tested on a subset of the universe of possible cases and therefore always require inferences to cases outside of a test sample. This is as true for case studies and cross-national research as it is for experiments.

At least two approaches can be used to improve external validity. First, variation within a case can be examined to estimate the dependence of treatment effects on characteristics of the units under study. This approach is, however, ineffective if the key variation takes places across rather than within cases. Second, studies can be replicated in a representative sample of the set of cases for which the inferences are to be made. This is the better approach, but it is rarely employed. Two challenges seem particularly important for establishing external validity using this approach. The first is that sites where field experiments are implemented may never be representative insofar as they are specifically drawn from the subset of cases in which randomization is possible; there is thus a need to assess the biases that may result from the selection of sites into the pool of studies. The second challenge is that the metrics used to estimate effects are often site specific, making a comparison of effects across studies difficult.

## Ethical Concerns

As researchers engage in field experimentation, the lines between observer and practitioner become blurred. In the field of the political economy of development, where interventions can have life-or-death consequences, this raises a new set of ethical considerations. To what extent are researchers responsible for outcomes that result from manipulations implemented by third parties? Under what conditions should researchers agree to work with third parties to study the impacts of their initiatives? Should collaboration with some kinds of organizations be off limits? At present, there are no clear guidelines for political scientists to follow. Human-subjects considerations address only some of these issues; moreover, the authority of institutional review boards in these matters is unclear when the organizations designing and implementing interventions are not affiliated with universities. In practice, researchers



ANNUAL REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

follow general do-no-harm principles, which involve avoiding deception or the propagation of false information, employing randomization only when there is equal need among subjects, not taking actions to prevent control groups from receiving benefits, etc.

But such principles are not clearly codified and are not universally endorsed, and in practice, researchers make different estimates of the tradeoffs between the quality of the measures that can be gathered and the ethical concerns that are raised by collecting those measures. The study of corrupt practices illustrates the point. To examine variation, some researchers use strategies that encourage subjects *not* to engage in bribery (Vicente 2007); others have provided incentives that indirectly induce subjects to engage in it (Bertrand et al. 2006); others have even assigned confederates to commit minor infractions in order to induce bribe-seeking by officials (Fried et al. 2008). These studies seek to answer important questions but confront difficult ethical issues on which greater clarity is needed if this research is to continue in a responsible manner.

## LOOKING AHEAD

The use of field experiments to study the political economy of development is still in its infancy. Studies have begun to appear; their reception suggests a new openness in the discipline to the power of field experiments for identifying causal effects in a field that, until now, has been dominated by cross-country regressions and comparative case studies. This new work has the potential both to shed light on practical questions of development effectiveness and to address broader questions at the heart of the development process. As yet, however, the number of completed experiments in the field is small, and there has not been a significant accumulation of knowledge.

As we look ahead to the likely role of field experiments in the future, we are conscious of the limitations of the approach and cognizant of the warning, voiced by Bates (2006), that field experiments must not "transform the field from a search for the underlying forces of development into a form of policy analysis." Four main challenges stand out.

First, the primary challenge, which we have highlighted throughout this discussion, is to ensure that studies which may serve as program evaluations for partners provide more general lessons for social scientists. Scholars will benefit from linking empirical studies to more general theoretical work, using that work to guide hypothesis formation.

Second, to engage with the major theories that inform the political economy of development, field experiments will need to seek greater nuance than is generally sought in clinical trials, focusing not simply on aggregate causal relationships but on the mechanisms that underlie them. Knowing, for example, whether aid tends to strengthen or weaken community cohesion is certainly important, but a more satisfactory account should seek to explain why this is the case.

Third, although it is rewarding to use these techniques to examine new and uncharted areas, the accumulation of knowledge will depend on replicating studies across sites and developing measures that are "portable," i.e., that possess an interpretation that can be useful beyond the study site.

Finally, and perhaps most difficult, studies will likely have to move beyond the very micro level at which many now operate. At present, most field experiments in the political economy of development examine the behavior of individual voters and constituents, community committees, or local governments. This is a level that, from an operational perspective, is well suited to field experimentation. However, much of the politics of development happens at an elite level, among ministers, members of parliament, leaders of business, and international donors. If this method is to contribute broadly to the study of how politics shapes development, field experiments will have to tackle such elite behavior as well.

ANNUAL REVIEWS IN ADVANCE

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We appreciate very helpful comments provided by Chris Blattman, James Fearon, Donald Green, Elisabeth King, Laura Paler, Cyrus Samii, and one anonymous reviewer.

## LITERATURE CITED

Acemoglu D, Johnson S, Robinson J. 2001. The colonial origins of comparative development: an empirical investigation. *Am. Econ. Rev.* 91(Dec.):1369–401

Bates R. 2006. Banerjee's approach might teach us more about impact but at the expense of larger matters. *Boston Rev. Books* July/Aug. 31(4). Available at **http://bostonreview.net/BR31.4/bates.html**. Accessed Jan. 1, 2009

Beath A, Enikolopov R, Fotini C. 2008. *Randomized impact evaluation of Afghanistan's National Solidarity Program (NSP): summary of evaluation methodology and activities*. Disc. pap. Washington, DC: World Bank

Bertrand M, Djankov S, Hanna R, Mullainathan S. 2006. *Does corruption produce unsafe drivers?* NBER Work. Pap. No. 12274

Bjorkman M, Svensson J. 2009. Power to the people: evidence from a randomized experiment of a citizen report card project in Uganda. *Q. J. Econ.* 124(May)

Blattman C, Fiala N, Martinez S. 2008. *Post-conflict youth livelihoods: an experimental impact evaluation of the Northern Uganda Social Action Fund (NUSAF)*. Concept note and design document, World Bank, Washington, DC

Bowles S, Gintis H. 2004. Persistent parochialism: trust and exclusion in ethnic networks. *J. Econ. Behav. Organ.* 55(1):1–23

Chattopadhyay R, Duflo E. 2004. Women as policy makers: evidence from a randomized policy experiment in India. *Econometrica* 72(Sep.):1409–43

Collier P, Vicente P. 2008. *Votes and violence: experimental evidence from a Nigerian election*. Unpublished manuscript, Dep. Econ., Oxford Univ.

Committee on the Evaluation of USAID Democracy Assistant Programs. 2007. *Improving Democracy Assistance: Building Knowledge Through Evaluations and Research*. Washington, DC: Natl. Acad. Press

Cox D, Reid N. 2000. *The Theory of the Design of Experiments*. Monographs on Statistics and Applied Probability 86. Boca Raton, FL: Chapman & Hall/CRC

Duflo E, Hanna R, Ryan S. 2008. *Monitoring works: getting teachers to come to school*. CEPR Disc. Pap. No. DP6682

Fearon J, Humphreys M, Weinstein J. 2009. Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict liberia. *Am. Econ. Rev. Papers Proc.* In press

Fried BJ, Lagunes P, Venkataramani A. 2008. *Corruption and inequality at the crossroad: a multi-method study of bribery and discrimination in Latin America*. Work. Pap. Dep. Polit. Sci., Yale Univ.

Gibson JL. 2008. Group identities and theories of justice: an experimental investigation into the justice and injustice of land squatting in South Africa. *J. Polit.* 70(July):700–16

Green D, Gerber A. 2002. Reclaiming the experimental tradition in political science. In *Political Science: State of the Discipline*, ed. I Katznelson, H Milner, pp. 805–32. New York: W.W. Norton

Guan M, Green DP. 2006. Noncoercive mobilization in state-controlled elections: an experimental study in Beijing. *Comp. Polit. Stud.* 39(Dec.):1175–93

Gugerty MK, Kremer M. 2008. Outside funding and the dynamics of participation in community associations. *Am. J. Polit. Sci.* 52(July):585–602

Habyarimana J, Humphreys M, Posner D, Weinstein J. 2007. Why does ethnic diversity undermine public goods provision? *Am. Polit. Sci. Rev.* 101(Nov.):709–25



REVIEWS IN ADVANCE

Harrison GW, List JA. 2004. Field experiments. *J. Econ. Lit.* XLII(Dec.):1009–55
Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. New York: Oxford Univ. Press
Holland P. 1986. Statistics and causal inference. *J. Am. Stat. Assoc.* 81(Dec.):945–60
Humphreys M, Masters WA, Sandbu ME. 2006. The role of leaders in democratic deliberations: results from a field experiment in Sao Tome and Principe. *World Polit.* 58(July):583–622
Humphreys M, Weinstein J. 2007. *Policing politicians: citizen empowerment and political accountability in Uganda*. Presented at Annu. Meet. Am. Polit. Sci. Assoc.
Hyde S. 2007. The observer effect in international politics: evidence from a natural experiment. *World Polit.* 60(October):37–63
Kremer M, Miguel E, Thornton R. 2008. *Does education change political attitudes? Evidence from a Kenyan school experiment*. Unpublished manuscript, Dep. Econ., Harvard Univ.
Miguel E, Kremer M. 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72:159–217
Nunn N. 2008. The long term effects of Africa's slave trades. *Q. J. Econ.* 123:139–76
Oakley A. 2000. A historical perspective on the use of randomized trials in social science settings. *Crime Delinquency* 46:315–29
Olken BA. 2007. Monitoring corruption: evidence from a field experiment in Indonesia. *J. Polit. Econ.* 115:200–49
Olken BA. 2008. *Direct democracy and local public goods: evidence from a field experiment in Indonesia*. NBER Work. Pap. #1123
Paluck EL, Green DP. 2008. *Deference, dissent, and dispute resolution: an experimental intervention using mass media to change norms and behavior in Rwanda*. Unpublished manuscript, Dep. Psychol., Princeton Univ.
Putnam R. 1994. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton Univ. Press
Rubin DB. 1986. Which ifs have causal answers? Discussion of Holland's "Statistics and Causal Inference." *J. Am. Stat. Assoc.* 81:961–62
Savedoff L, Levine R, Birdsall N, et al. 2006. *When will we ever learn? Improving lives through impact evaluation*. Rep. Evaluation Gap Work. Group, Cent. Global Dev., Washington DC
Vicente PC. 2007. *Is vote buying effective? Evidence from a randomized experiment in West Africa*. Unpublished manuscript, Dep. Econ., Oxford Univ.
Wantchekon L. 2003. Clientelism and voting behavior: evidence from a field experiment in Benin. *World Polit.* 55(April):399–422

Annu. Rev. Polit. Sci. 2009.12. Downloaded from arjournals.annualreviews.org
by Columbia University on 03/23/09. For personal use only.

REVIEWS IN ADVANCE